Accepted Manuscript

Dynamic pricing in the presence of individual learning

Xi Weng

PII: S0022-0531(14)00175-6
DOI: 10.1016/j.jet.2014.11.016
Reference: YJETH 4353

To appear in: *Journal of Economic Theory*

Received date: 7 December 2013
Revised date: 15 November 2014
Accepted date: 26 November 2014

Please cite this article in press as: X. Weng, Dynamic pricing in the presence of individual learning, *J. Econ. Theory* (2014), http://dx.doi.org/10.1016/j.jet.2014.11.016

Dynamic Pricing in the Presence of Individual Learning*

Xi Weng†

Department of Applied Economics, Guanghua School of Management, Peking University, 100871 Beijing, China

Abstract

This paper studies price dynamics in a setting in which a monopolist sells a new experience good over time to many buyers, and the seller can neither price discriminate among the buyers nor commit to a price rule. Buyers learn from their own experiences about the effectiveness of the product. Individual learning generates *ex post* heterogeneity, which affects the buyers' purchasing decisions, the monopolist's pricing strategy, and efficiency. When learning occurs through good news signals, buyers receive a rent because of the possible advantageous belief caused by short-lived deviations. If a good news signal arrives, the price can instantaneously increase or decrease depending on the arrival time of this signal. The equilibrium is inefficient because the monopolist's incentive to exploit known buyers leads to inefficient early termination of exploration. When learning occurs through bad news signals, *ex post* heterogeneity has no such effect, since only homogeneous unknown buyers purchase the experience good.

JEL classification: D83; C02; C61; C73

Keywords: Learning; Experimentation; Strategic pricing; Exponential bandit; Good news case; Bad news case

*I am especially grateful to George Mailath for continuous encouragement, guidance and support. I also thank Hongbin Cai, David Dillenberger, Jan Eeckhout, Hanming Fang, Qingmin Liu, Steve Matthews, Guido Menzio, Tymofiy Mylovanov, Mallesh Pai, Andrew Postlewaite, Jeroen Swinkels and seminar participants at Kellogg, Michigan, UCL, Penn, Toulouse, University of Hong Kong and UPF for helpful comments and suggestions. I also acknowledge financial support from the National Natural Science Foundation of China (Grant No. 71303014) and Guanghua Leadership Institute (Grant No. 12-02), as well as support from the Spanish Ministry of the Economy and Competitiveness (Project ECO2012-36200) and the Key Laboratory of Mathematical Economics and Quantitative Finance (Peking University), Ministry of Education, China.

†Corresponding author. Fax: +86-10 6275 1463. E-mail address: wengxi125@gsm.pku.edu.cn.

1 Introduction

In many markets for new experience goods, buyers face huge uncertainty about the effectiveness or possible side effects of the product. Take as an example the market for new drugs. The effectiveness of a new drug critically depends on whether there is an appropriate match between this drug and each patient's particular problems ([16]). Patients learn from their own experiences (*individual learning*) about this effectiveness.

This paper investigates how a monopolist sells a new experience good to many buyers over time in the presence of individual learning. The monopolist and the buyers initially are equally unsure about the effectiveness of the product. How will this monopolist price strategically if she observes each buyer's past actions and outcomes? Without having seen the effectiveness of the product, potential purchasers becomes increasingly pessimistic. In order to keep buyers purchasing the product, the price must be reduced. How will the monopolist react when the product is revealed to be effective for one buyer? Will strategic pricing achieve efficient allocation?

In this paper, dynamic monopoly pricing is modeled as an infinite-horizon, continuous-time process. The monopolist sells a perishable experience good. She can neither price-discriminate across buyers nor commit to a price rule. At each instant of time, the monopolist first posts a spot price, which is contingent on the available public information about the experiences of the buyers. Each buyer then decides to either buy one unit of the experience good or take an outside option (modeled as another good of known characteristics). The experience good generates random lump-sum payoffs according to independent Poisson processes. The arrival rate of the lump-sum payoffs depends on an unknown individual attribute, which is binary and uncorrelated across buyers. For tractability, we assume that the public arrival of lump-sum payoffs immediately resolves the idiosyncratic uncertainty of the receiver. A key feature of the model is that buyers can become *ex post heterogeneous* in two ways: heterogeneity can be induced by either different outcomes or different actions.

We consider two different cases. In the *good news case*, the experience good generates positive lump-sum payoffs; in the *bad news case*, it generates negative lump-sum damages (e.g., side effects of new drugs). This paper fully characterizes the symmetric Markov perfect equilibrium for both cases. If a monopolist sells to a single buyer, the equilibrium price is set such that that buyer is indifferent between purchasing the experience good and taking the outside option. The buyer's purchasing decision is purely myopic since her continuation value is independent of the learning outcomes. This leads to an efficient outcome since the monopolist fully internalizes the social surplus.

We first characterize the symmetric Markov perfect equilibrium for the good news case in which there are two buyers. In phase S, i.e., when no lump-sum payoff has yet arrived, the monopolist sells to both unknown buyers before quitting the market (an "unknown" buyer refers to a buyer whose valuation of the good has not been revealed); in phase I, i.e., after one buyer has received a lump-sum payoff, the critical tradeoff is whether to sell to both buyers and or to sell only to the known buyer. In both phases, the equilibrium purchasing behavior is determined by a cutoff in the posterior belief about the unknown buyer's individual attribute. Each unknown buyer makes a purchase when the posterior belief is above this cutoff and takes the outside option otherwise.

In phase I, the unknown buyer's purchasing decision is purely myopic as in the single buyer case. The key reason for this is that if the monopolist sells to both buyers, the equilibrium price is set to make the more pessimistic unknown buyer indifferent. In phase S, however, the presence of ex post heterogeneity has two important implications for the equilibrium price.

First, consider the situation where two *ex ante* identical unknown buyers make different purchasing decisions. One buyer continues experimentation by purchasing the experience good, while the other buyer deviates to take the outside option for a small amount of time. If the experimenter does not receive any lump-sum payoffs during that period, she becomes more pessimistic about her individual attribute. Without price discrimination, if the monopolist sells to two different buyers, the optimal price is set to make the more pessimistic buyer indifferent between the alternatives. The deviator, who is more optimistic about the experience good, pays less than her willingness to pay. This implies that it is always profitable for the deviator to cause asymmetric beliefs between the other buyer and herself, and the monopolist must provide extra subsidy to deter such a short-lived deviation.[1] Because of this *deterrence effect*, the equilibrium price is lowered such that each unknown buyer myopically prefers purchasing the experience good than taking the outside option.

Second, there is another positive *continuation value effect* on the price in phase S. This is driven by ex post heterogeneity in phase I. In phase I, the monopolist faces a trade-off between exploiting the buyer who is known to be good, and exploring the unknown buyer. As long as this trade-off is resolved in favor of exploration, the known buyer receives a rent by paying a relatively low price. This feeds back to phase S: the monopolist can charge a relatively high price due to this extra incentive to experiment. The combination of the deterrence effect and the continuation value effect implies that the price response to the first lump-sum payoff is ambiguous. When the first lump-sum payoff arrives relatively early, there is an instantaneous drop in the price; whereas, when the first lump-sum payoff arrives relatively late, there is an instantaneous jump in the price.

The equilibrium purchasing behavior in the good news case is characterized for an arbitrary number of buyers. It turns out that the equilibrium experimentation level is always lower than the socially efficient one when at least one buyer has received a lump-sum payoff. This is due to the existence of ex post heterogeneity: known buyers are willing to pay more than unknown buyers. Without price discrimination, the tradeoff between exploitation and exploration leads to inefficient

[1] This is similar to findings in the dynamic moral hazard model with experimentation (see, e.g., [3], [9] and [20]). In these papers, it is shown that a misalignment of beliefs between the principal and the agent is always profitable for the agent, and must be dissuaded by providing more high-powered incentives. However, the mechanism is due to unobservable effort, and hence is different from the one discussed in this paper.

early termination of experimentation.

We then characterize the symmetric Markov perfect equilibrium for the bad news case. It is shown that the equilibrium is always efficient. The key insight is that although buyers become heterogeneous, the buyers who have received lump-sum damages will never repurchase the experience good. The only potential buyers are the unknown ones, who are ex post homogeneous in a symmetric equilibrium. Another important difference between the good and bad news cases is that no extra subsidy is needed in the bad news case since deviations by an unknown buyer make the deviator more pessimistic. As a result, there is no deterrence effect and no continuation value effect.

Related Literature

[4] and [19] are two early papers that analyze the impact of price competition on experimentation. They show that if there is only one buyer, dynamic duopoly competition with vertically differentiated products can achieve efficiency. However, [5] show that in the presence of social learning, dynamic duopoly competition cannot achieve efficiency. [6] and [12] allow *ex ante heterogeneity* in the sense that buyers are different in their willingness to pay.[2] Both papers assume a continuum of buyers. At each instant of time, an individual buyer makes a myopic optimal choice and strategic interactions between buyers do not exist.

[7] also consider a dynamic monopoly pricing problem, but with a continuum of buyers. The difference in crucial modeling assumptions leads them to investigate different properties of the equilibrium price path. When the framework of a continuum of buyers is used, it implies that no individual buyer can affect the price path of the monopolist, and hence the analysis of deviation incentives is greatly simplified. However, it is also impossible to discuss the impact of a single good news signal on price.[3] Instead, [7] are more concerned with would follow a downward trajectory or whether it would eventually go up in equilibrium. [14] and [15] develop a different model for dynamic monopoly pricing under social learning. Their model is closer to that found in the herding literature: each short-lived buyer makes a purchasing decision following a pre-determined sequence. In contrast, in our model, all buyers are long-lived and make purchasing decisions repeatedly.

This paper is also closely connected to the continuous-time strategic experimentation literature. A nonexhaustive list of related papers includes [11], [21], [22] and [23].[4] The analysis of our model

[2] [27] also investigates a duopoly model with *ex ante heterogeneity* along a location. This paper considers a two-period model and is mainly concerned about consumer loyalty, i.e., whether in the second period, buyers return to the seller that they originally purchased from.

[3] In the Web Appendix, we consider an extended good news model with a continuum of buyers. In equilibrium, the monopolist will initially decrease the price to encourage experimentation. It turns out that the deterrence effect and the continuation effect cancel each other out, and the price is set such that the unknown buyers are myopically indifferent between purchasing the experience good and taking the outside option. However, due to the option value of becoming a known buyer, each unknown buyer strictly prefers purchasing the experience good.

[4] The strategic experimentation framework is also used as a building block to investigate broader issues. For example, [26] investigates voting in a strategic experimentation environment; [28] considers a war-of-attrition game

setting is greatly simplified by the use of exponential bandits, building on [23]. Most research in the strategic experimentation literature assumes a common value environment, in which the players' payoffs are perfectly correlated. This enables us to use a uni-dimensional posterior belief as the unique state variable to characterize the value functions. By considering individual learning, we introduce multi-dimensional posterior beliefs and develop methods to solve this dimensionality problem.

In addition to the theoretical body of work, several empirical studies attempt to quantify the importance of learning considerations on consumers' dynamic purchasing behavior. However, most existing works have exclusively focused on modeling individual consumer behavior and analyzing the impact of idiosyncratic uncertainty (see, e.g., [2], [16], [18], etc). This paper complements this literature by considering how the monopolist seller dynamically adjusts the price in the presence of individual learning.

The remainder of this paper is organized as follows. Section 2 introduces the model and defines the solution concept. Section 3 and Section 4 solve a symmetric Markov perfect equilibrium and discuss the efficiency of the equilibrium for the good news case and the bad news case, respectively. Section 5 concludes the paper.

2 Model Setting

We consider a continuous-time model with $t \in [0, +\infty)$ and a positive discount rate $r > 0$. The market consists of n buyers indexed by $i = 1, 2, \cdots, n$ and one monopolist, who are all risk-neutral. A monopolist with a zero cost of production sells a *risky product* with unknown value.[5] At each point in time, a buyer can either buy one unit of the risky product or take a safe outside option/product.

If a buyer purchases the safe product, she receives a known deterministic flow payoff $s > 0$.[6] The value of the risky product to a buyer i consists of two components: a deterministic flow payoff $\xi_f \geq 0$ and a random lump-sum payoff ξ_l. The presence of lump-sum payoffs depends on the quality of the match between the product and the specific buyer: it is either *relevant* ($\kappa_i = 1$) or *irrelevant*($\kappa_i = 0$). The arrival of random lump-sum payoffs ξ_l is independent across buyers and modeled as a Poisson process with intensity $\lambda_H \kappa_i$, with $\lambda_H > 0$. The common priors are such that $\rho_0 = \Pr(\kappa_i = 1)$ for each buyer i. The product characteristic and the match qualities are initially unobservable to all players (seller and buyers), but the parameters λ_H, ξ_f, ξ_l, and ρ_0 are common knowledge.

in which the players are learning about their private payoffs; [3], [13], [20] and [24] all investigate moral hazard problems when effort affects the speed of learning.

[5]The zero cost assumption is simply a normalization. The model with production cost $c > 0$ is equivalent to another one with zero production cost and the flow payoff of the safe product being $s' = s + c$.

[6]Alternatively, we can assume that the flow payoff is random but drawn from a commonly known distribution with expectation $s > 0$.

We consider two cases for the above setting. In the *good news case*, $\xi_l > 0$ and the arrival of lump-sum payoffs makes the risky product more attractive than the safe one. We assume that the risky product is superior to the safe one only when the buyers can receive lump-sum payoffs:

Assumption 1 *(Good News Case) In the good news case,* $\xi_l > 0$ *and* $\xi_f < s < \xi_f + \lambda_H \xi_l$.

In the *bad news case*, $\xi_l < 0$ and the arrival of lump-sum payoffs makes the risky product less attractive than the safe one. We impose the requirement that the risky product is superior to the safe one only when the buyers cannot receive lump-sum payoffs:

Assumption 2 *(Bad News Case) In the bad news case,* $\xi_l < 0$ *and* $\xi_f > s > \xi_f + \lambda_H \xi_l$.

All players observe each buyer's past actions and outcomes. As a result, both the seller and the buyers hold common posterior beliefs about any given buyer's match quality. In both cases, if one buyer receives a lump-sum payoff from the risky product, every player immediately knows that that buyer's match is relevant. The absence of lump-sum payoffs makes it very likely that a match is irrelevant.

At each instant of time t, the monopolist first announces a spot price based on previous history and then each buyer decides which product to purchase conditional on previous history and the announced price. It is assumed that the monopolist can neither price-discriminate across buyers nor commit to a price rule.

2.1 Belief Updating

Denote by N_{it} the total number of lump-sum payoffs received by buyer i before time t. Let P_t be the price charged by the monopolist at time t. Set $a_{it} = 1$ if buyer i purchases the risky product at time t; and $a_{it} = 0$ if buyer i purchases the safe product at time t. Public history before time t is defined as:

$$h_t \triangleq (\{a_{i\tau}, N_{i\tau}\}_{i=1}^n, P_\tau)_{0 \le \tau < t}\,.$$

Posterior belief that buyer i's match is relevant is defined as:

$$\rho_{it} \triangleq \Pr[\kappa_i = 1 \mid h_t].$$

Given prior ρ_0, the posteriors ρ_{it} evolve according to Bayes' rule. A buyer i who has not received any lump-sum payoff before time t expects to receive a lump-sum payoff from the risky product with arrival rate $\lambda_H a_{it} \rho_{it}$. If a lump-sum payoff is received, ρ_{it} immediately jumps to 1; otherwise, ρ_{it} obeys the following differential equation at those times t when a_{it} is right continuous:[7]

$$\dot{\rho}_{it} = -\lambda_H a_{it} \rho_{it} (1 - \rho_{it}). \tag{1}$$

[7] If buyer i has not received good news within the time period $[t, t+h)$, then the posterior belief $\rho_{i,t+h}$ can be

2.2 Strategies and Payoffs

Throughout the paper, we focus on symmetric Markov perfect equilibria. The natural state variables are the posteriors $\boldsymbol{\rho}$. The state variable $\boldsymbol{\rho}_t$ is required to be feasible in the sense that

$$\boldsymbol{\rho}_t \in \Sigma = \{\boldsymbol{\rho} \in [0,1]^n : \text{ either } \rho_i = 1 \text{ or } \rho_i \leq \rho_0 \text{ for all } i\}.$$

Purchasing Decision Buyer i's acceptance policy is a function of states $\boldsymbol{\rho}$ and price P

$$\alpha_i : \Sigma \times \mathbb{R} \to \{0,1\}.$$

Since lump-sum payoffs arrive at rate $\rho_{it}\lambda_H$, the expected flow of utility associated with purchasing decision a_{it} is

$$a_{it}\rho_{it}\lambda_H\xi_l + a_{it}(\xi_f - P_t) + (1 - a_{it})s.$$

The choice of a_{it} affects not only flow utility but also how beliefs $\boldsymbol{\rho}_t$ are updated. Given beliefs $\boldsymbol{\rho} \in \Sigma$, monopolist's strategy P and other buyers' strategies α_{-i}, buyer i's value (sum of normalized expected discounted utility) from purchasing strategy α_i is

$$U_i(\alpha_i, P, \alpha_{-i}; \boldsymbol{\rho}) = \mathbb{E}\int re^{-rt}\left\{\alpha_i(\boldsymbol{\rho}_t, P_t)\left(\rho_{it}\lambda_H\xi_l + \xi_f - P_t\right) + (1 - \alpha_i(\boldsymbol{\rho}_t, P_t))s\right\}dt$$

where the expectation is taken over $\{\boldsymbol{\rho}_t : t \in [0,\infty)\}$ with $\boldsymbol{\rho}_0 = \boldsymbol{\rho}$.

Pricing Decision Given prior ρ_0, the monopolist's price is a function of states $\boldsymbol{\rho}$

$$P : \Sigma \to \mathbb{R}.$$

Given buyers' strategies $\{\alpha_i\}_{i=1}^n$, the flow profits associated with price P_t are $\sum_{i=1}^n \alpha_i(\boldsymbol{\rho}_t, P_t)P_t$. The choice of P_t affects not only flow profits but also purchasing decisions and as a result how beliefs are updated. Given beliefs $\boldsymbol{\rho}$ and buyers' strategies $\{\alpha_i\}_{i=1}^n$, the monopolist's value (sum of normalized expected discounted profits) from the pricing policy P is

$$J(P, \alpha; \boldsymbol{\rho}) = \mathbb{E}\int re^{-rt}\sum_{i=1}^n \alpha_i(\boldsymbol{\rho}_t, P(\boldsymbol{\rho}_t))P(\boldsymbol{\rho}_t)dt$$

where the expectation is taken over $\{\boldsymbol{\rho}_t : t \in [0,\infty)\}$ with $\boldsymbol{\rho}_0 = \boldsymbol{\rho}$.

written as:

$$\rho_{i,t+h} = \frac{\rho_{it}e^{-\lambda_H\int_0^h a_{i,t+\tau}d\tau}}{\rho_{it}e^{-\lambda_H\int_0^h a_{i,t+\tau}d\tau} + 1 - \rho_{it}}.$$

Since $a_{i\tau}$ is right continuous with respect to time at time t, there exists some $\bar{h} > 0$ such that $a_{i,t+\tau} = a_{i,t}$ for all $\tau \leq \bar{h}$. Hence by definition:

$$\dot{\rho}_{it} = \lim_{h\to 0}\frac{\rho_{i,t+h} - \rho_{i,t}}{h} = -\lambda_H a_{it}\rho_{it}(1 - \rho_{it}).$$

Admissible Strategies A critical issue associated with continuous time model setting is that a well-defined strategy profile need not yield a well-defined outcome ([8] and [25]). Some restrictions on strategies must be imposed to overcome this issue. One requirement is that the Markovian strategy profile (P, α) be admissible. The formal definition can be found in the appendix. If a strategy profile satisfies this requirement, the induced outcome is well behaved in the sense that the purchasing decisions a_{it} and pricing decisions P_t are right continuous functions in the absence of lump-sum payoffs.

2.3 Symmetric Markov Perfect Equilibrium

We consider a symmetric Markov perfect equilibrium. The formal definition of our solution concept is the following:

Definition 1 *Given prior ρ_0, an admissible Markovian strategy profile $\{P^*, \alpha^*\}$ is a* Markov perfect equilibrium *if for all i, feasible beliefs $\boldsymbol{\rho}$ and all admissible strategies $\tilde{P}$ and $\tilde{\alpha}_i$:*[8]

$$J(P^*, \alpha^*; \boldsymbol{\rho}) \geq J(\tilde{P}, \alpha^*; \boldsymbol{\rho}) \quad \textit{and} \quad U_i(\alpha_i^*, P^*, \alpha_{-i}^*; \boldsymbol{\rho}) \geq U_i(\tilde{\alpha}_i, P^*, \alpha_{-i}^*; \boldsymbol{\rho}).$$

Moreover, $\{P^*, \alpha^*\}$ is symmetric *if for all permutations $\pi : \{1, \cdots, n\} \to \{1, \cdots, n\}$, $P(\tilde{\boldsymbol{\rho}}) = P(\boldsymbol{\rho})$ where $\tilde{\rho}_i = \rho_{\pi^{-1}(i)}$ and $\alpha_i(\boldsymbol{\rho}, P) = \alpha_{\pi(i)}(\tilde{\boldsymbol{\rho}}, P)$.*

3 Equilibrium in the Good News Case

In the good news case, $\xi_l > 0$ and the arrival of a lump-sum payoff makes the risky product more valuable to the receiver of this payoff. In this section, we normalize $\xi_f = 0$ and $\xi_l = v > 0$. Assumption 1 implies $g \triangleq \lambda_H v > s > 0$.

With two players, there are only two situations to consider: 1) phase S, in which neither buyer has received good news; and 2) phase I, in which one buyer has received good news. Obviously, when both buyers have received good news, the monopolist should charge price $g - s$ to extract the full surplus.

3.1 Socially Efficient Allocation

Before solving for a symmetric Markov perfect equilibrium, we first solve for the socially efficient allocation. The linear utility function enables us to obtain the efficient allocation policy by solving a specific multi-armed bandit problem, in which payoffs are given by the aggregate surplus.

Suppose $k \leq n$ buyers have received good news; then it is socially optimal for them to keep purchasing the risky product following assumption 2 and the social surplus function is

[8] Strategies $\tilde{P}$ and $\tilde{\alpha}_i$ need not be Markovian. The definition of admissible non-Markovian strategies can also be found in the appendix.

$$\Omega_k(\rho) = kg + (n-k)W(\rho)$$

where

$$W(\rho) = \sup_{\{\alpha_t\}_{t\in\mathbb{R}_+}\in\Gamma} \mathbb{E}\int_{t=0}^{\infty} re^{-rt}[\alpha_t\rho_t g + (1-\alpha_t)s]dt$$

is the optimal value for an unknown buyer with posterior belief ρ. Γ in the above definition denotes the set of sequences $\{\alpha_t\}_{t\in\mathbb{R}_+}$ satisfying $\alpha_t \in \{0,1\}$ for all $t \in \mathbb{R}_+$, and α_t being right continuous in t.

Since the unknown buyers are facing a standard independent two-armed bandit problem, previous research (see [23]) has characterized the optimal cutoff and value function W. It is efficient for the unknown buyers to stop purchasing the risky product once the posterior belief ρ reaches

$$\rho^e = \frac{rs}{(r+\lambda_H)g - \lambda_H s}$$

if no lump-sum payoff has yet been received. The efficient cutoff ρ^e does not depend on the priors ρ_0.

3.2 Equilibrium Characterization for $n = 2$

In the two-buyer case, there are three situations to consider. In phase S, denote U_S as the value function for each unknown buyer; and J_S as the value function for the monopolist. In phase I, denote U_I as the value function for the unknown buyer; V_I as the value function for the known buyer; and J_I as the value function for the monopolist. When both buyers have received lump-sum payoffs, denote V_2 as the value function for the known buyers, and J_2 as the value function for the monopolist.

For $\zeta = S, I$, denote α^0_ζ (α^1_ζ) as the strategy used by the known (unknown) buyers. Let P_ζ be the price charged by the monopolist. Then definition 1 implies that a triple of $(P_\zeta, \alpha^0_\zeta, \alpha^1_\zeta)$ is a symmetric Markov perfect equilibrium if the following conditions are satisfied:

- for $\zeta = I$, $\alpha^0_\zeta = 1$ if $P \le g - s$ and $= 0$ otherwise;
- for $\zeta = S$, the unknown buyers choose acceptance policy α^1_ζ to maximize (α^1_ζ should vary with time, and the same is true for $\zeta = I$):

$$U_\zeta(\rho) = \sup_{\alpha^1_\zeta} \mathbb{E}\left\{\int_{t=0}^{\tau} re^{-rt}\left[\alpha^1_\zeta(\rho_t g - P_\zeta(\rho_t)) + (1-\alpha^1_\zeta)s\right]dt + e^{-r\tau}\left(\frac{1}{2}V_I(\rho_\tau) + \frac{1}{2}U_I(\rho_\tau)\right)\right\}$$

and given α^1_ζ, the monopolist chooses price $P_\zeta(\rho_t)$ to maximize

$$J_\zeta(\rho) = \sup_{P_\zeta(\cdot)} \mathbb{E}\left\{\int_{t=0}^{\tau} 2re^{-rt}\alpha^1_\zeta(\rho_t, P_\zeta(\rho_t))P_\zeta(\rho_t)dt + e^{-r\tau}J_I(\rho_\tau)\right\},$$

where τ is the first (possibly infinite) time at which a new unknown buyer receives good news;

- for $\zeta = I$, the unknown buyer chooses acceptance policy α^1_ζ to maximize:

$$U_\zeta(\rho) = \sup_{\alpha^1_\zeta} \mathbb{E}\left\{\int_{t=0}^{\tau} re^{-rt}\left[\alpha^1_\zeta(\rho_t g - P_\zeta(\rho_t)) + (1-\alpha^1_\zeta)s\right]dt + e^{-r\tau}V_2(\rho_\tau)\right\}$$

and given $(\alpha^0_\zeta, \alpha^1_\zeta)$, the monopolist chooses price $P_\zeta(\rho_t)$ to maximize

$$J_\zeta(\rho) = \sup_{P_\zeta} \mathbb{E}\left\{\int_{t=0}^{\tau} re^{-rt}\left[\alpha^0_\zeta(\rho_t, P_\zeta(\rho_t)) + \alpha^1_\zeta(\rho_t, P_\zeta(\rho_t))\right]P_\zeta(\rho_t)dt + e^{-r\tau}J_2(\rho_\tau)\right\};$$

- beliefs update according to Bayes' rule: ρ_t satisfies the law of motion, i.e., equation (1);

- when both buyers have received lump-sum payoffs, the price is $g - s$, such that $J_2 = 2(g - s)$ and $V_2 = s$.

First, it is straightforward to see that the known buyers always buy the risky product if the price is lower than $g - s$ and do not buy otherwise. Second, when both unknown buyers purchase the risky product, the conditional probability that any given unknown buyer becomes good is simply $1/2$, since the two unknown buyers' payoff distributions are identical. Finally, if both buyers have received lump-sum payoffs, it is optimal for the monopolist to charge price $g - s$ to extract all of the surplus.

3.2.1 Equilibrium in phase I

A backward procedure is used to characterize the equilibrium. In phase I, the equilibrium cutoff $\rho^\star_I$ and the various value functions are provided by the following proposition. The proofs of the following and all subsequent results can be found in Appendix B.

Proposition 1 *Fix any symmetric Markov perfect equilibrium. In phase I, the unknown buyer purchases the risky product if and only if the posterior belief ρ is larger than*

$$\rho^\star_I \triangleq \frac{r(g+s)}{2rg + \lambda_H(g-s)}.$$

The equilibrium price is $P_I(\rho) = g\rho - s$ and the unknown buyer receives value $U_I(\rho) = s$; the known buyer receives value

$$V_I(\rho) = \max\left\{s, s + g(1-\rho)(1 - [\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}]^{r/\lambda_H})\right\}; \tag{2}$$

and the monopolist receives value

$$J_I(\rho) = \begin{cases} 2(g\rho - s) + (g + s - 2g\rho_I^\star)\frac{1-\rho}{1-\rho_I^\star}[\frac{(1-\rho)\rho_I^\star}{(1-\rho_I^\star)\rho}]^{r/\lambda_H} & \textit{if } \rho > \rho_I^\star \\ g - s & \textit{otherwise.} \end{cases}$$

It is straightforward to see that the equilibrium cutoff $\rho_I^\star$ is strictly larger than the efficient cutoff ρ^e. This is because ex post heterogeneity means that the known buyer is willing to pay more than the unknown buyer. In the absence of price discrimination, the monopolist faces a tradeoff between exploiting the buyer who is known to be good, and exploring the unknown buyer. The incentive to charge a high price and extract the full surplus from the known buyer causes inefficient early termination of exploration. Another remark is that the unknown buyer is myopically indifferent between purchasing the risky product or not since there is no learning value attached to the purchasing behavior (the unknown buyer always receives value s regardless of whether she receives a lump-sum payoff).

3.2.2 Equilibrium in phase S

Now consider the situation in which none of the buyers have received lump-sum payoffs yet. Assume that the posterior belief ρ is large enough that both buyers purchase the risky product in equilibrium. Since $\rho_I^\star > \rho^e$, it is natural to construct an equilibrium such that $\rho_I^\star > \rho_S^\star$. As a result, there are two cases to consider: $\rho \geq \rho_I^\star$ and $\rho < \rho_I^\star$.[9]

To characterize the equilibrium price and cutoff, we proceed as follows. First, we use the incentive compatibility constraint to derive the value function of the experimenting buyers. Second, we derive expressions of equilibrium price and the monopolist's value function based on the experimenting buyers' value function derived in the first step. Finally, we apply value matching and smooth pasting conditions (see, e.g., [17]) to pin down the equilibrium cutoff.

To ensure that both unknown buyers continue to experiment, a necessary condition requires both i) each buyer has an incentive to participate (i.e., the value is larger than the outside option s); and ii) neither buyer should benefit from the following deviations: stopping experimentation for a very small amount of time and then switching back to the specified equilibrium behavior.[10]

[9] It is shown in the proof of Proposition 3 that it is impossible to have $\rho_I^\star \leq \rho_S^\star$ in equilibrium.

[10] Here we consider only the continuous-time analog of one-shot deviation, because it has been proved that the lack of profitable one-shot deviations is sufficient to rule out profitable deviations by the definition of admissible strategies (see Appendix B.4).

The deviations described in constraint ii) are similar to one-shot deviations in discrete time models. Formally, this implies that for any $\rho > \rho_S^\star$, there exists $\bar{h}$ such that for all $h \leq \bar{h}$,

$$U_S(\rho) \geq \hat{U}(\rho; h) = \int_{t=0}^{h} re^{-rt} s dt + \rho(1 - e^{-\lambda_H h})e^{-rh}U_I(\rho) + [1 - \rho(1 - e^{-\lambda_H h})]e^{-rh}U^D(\rho, \rho_h) \quad (3)$$

where $\hat{U}(\rho; h)$ denotes the value for a deviator that deviates for h length of time. The deviator receives a deterministic payoff s within the h length of time. After the deviation, with probability $\rho(1 - e^{-\lambda_H h})$, the non-deviator has received lump-sum payoffs and the continuation value for the deviator is $U_I(\rho) = s$; with the complementary probability, the non-deviator has not received a lump-sum payoff and the two unknown buyers become asymmetric. In the latter situation, the deviator receives a continuation value $U^D(\rho, \rho_h)$ where superscript D stands for "deviator." The non-deviator ρ_h is more pessimistic than the deviator ρ since $\rho_h = \frac{\rho e^{-\lambda_H h}}{\rho e^{-\lambda_H h} + (1-\rho)} < \rho$. Obviously, equation (3) is a tighter constraint than the participation constraint since $U_I(\rho) = s$ and $U^D(\rho, \rho_h) \geq s$.

The most important technical result in this paper is to evaluate $\lim_{h \to 0} \frac{U_S(\rho) - \hat{U}(\rho; h)}{h}$. The main difficulty is to evaluate the off-equilibrium value function $U^D(\rho, \rho_h)$. We explain here intuitively how to derive the off-equilibrium value for a deviating buyer.

First notice that $\rho > \rho_S^\star$ means that it is optimal for the monopolist to sell to both unknown buyers on the equilibrium path. Then, for sufficiently small h, it is still optimal for the monopolist to sell to both unknown buyers after an h-deviation.

In other words, given a sufficiently small h, there exists some $\bar{h}'$ such that for all $h' \leq \bar{h}'$, we have:

$$\begin{aligned} U^D(\rho, \rho_h) = \mathbb{E} \int_{t=0}^{h'} & re^{-rt}(\rho_t g - \tilde{P}_t)dt \\ & + \rho(1 - e^{-\lambda_H h'})e^{-rh'}V_I(\rho_{h+h'}) + \rho_h(1 - e^{-\lambda_H h'})e^{-rh'} s \\ & + [1 - \rho(1 - e^{-\lambda_H h'}) - \rho_h(1 - e^{-\lambda_H h'})]e^{-rh'}U(\rho_{h'}, \rho_{h+h'}). \quad (4) \end{aligned}$$

In the above expression, ρ_t is the posterior about the deviator's match quality and starts from ρ; and $\tilde{P}_t$ is the off-equilibrium price set by the monopolist after an h-deviation. Obviously, the value function $U^D(\rho, \rho_h)$ depends on the off-equilibrium price and cannot be evaluated directly.

Meanwhile, notice that the non-deviator's value can be expressed as:

$$U^{ND}(\rho,\rho_h) = \mathbb{E}\int_{t=0}^{h'} re^{-rt}(\rho'_t g - \tilde{P}_t)dt + \rho(1-e^{-\lambda_H h'})e^{-rh'}s + \rho_h(1-e^{-\lambda_H h'})e^{-rh'}V_I(\rho_{h'}) + [1-\rho(1-e^{-\lambda_H h'}) - \rho_h(1-e^{-\lambda_H h'})]e^{-rh'}U(\rho_{h+h'},\rho_{h'}), \quad (5)$$

where ρ'_t is the posterior about the non-deviator's match quality and starts from ρ_h.

The key step is to decompose $U^D(\rho,\rho_h)$ as:

$$U^D(\rho,\rho_h) = U^{ND}(\rho,\rho_h) + (U^D(\rho,\rho_h) - U^{ND}(\rho,\rho_h)).$$

The reason for doing this decomposition is that the off-equilibrium price is cancelled when we subtract $U^{ND}(\rho,\rho_h)$ from $U^D(\rho,\rho_h)$, Hence, $Z(\rho,\rho_h) \triangleq U^D(\rho,\rho_h) - U^{ND}(\rho,\rho_h)$ is independent of the off-equilibrium price $\tilde{P}$ and can be evaluated directly.

Buyer ρ_h's value $U^{ND}(\rho,\rho_h)$ can be computed without using the off-equilibrium price. If the non-deviator has not received a lump-sum payoff during an h-deviation, she becomes more pessimistic than the deviator. If the monopolist wants to make a sale to both buyers, the optimal price is set according to the reservation value of the more pessimistic buyer. An expression of $U^{ND}(\rho,\rho_h)$ can be derived from the ρ_h buyer's incentive compatibility constraint. In the appendix, we show that this implies a first-order ordinary differential equation for $U^{ND}(\rho,\rho_h)$, which can be solved by imposing the boundary condition that $U(\rho_h,\rho_h) = U_S(\rho_h)$.

Second, given any $t < h'$, notice equations (4) and (5) also hold for posteriors $(\rho(t),\rho_h(t))$ where

$$\rho(t) = \frac{\rho e^{-\lambda_H t}}{\rho e^{-\lambda_H t} + (1-\rho)}, \quad \text{and} \quad \rho_h(t) = \frac{\rho_h e^{-\lambda_H t}}{\rho_h e^{-\lambda_H t} + (1-\rho_h)}.$$

Redefine

$$Z(t) = Z(\rho(t),\rho_h(t)) = U(\rho(t),\rho_h(t)) - U(\rho_h(t),\rho(t))$$

to be a function of time t . A first-order ordinary differential equation about $Z(t)$ can be obtained by subtracting equation (5) from equation (4) and letting the length of time interval converge to zero. Solving the ordinary differential equation, the expression for $Z(\rho,\rho_h)$ can be recovered by substituting time t as functions of $\rho(t)$ and $\rho_h(t)$. The boundary condition is such that $Z = 0$ once ρ_h reaches $\rho_S^\star$.

The details of the derivation can be found in Lemma 1 in the appendix. Moreover, Lemma 2 in the appendix implies that in equilibrium, a profit-maximizing monopolist should always make the incentive constraints to be "binding" in the sense that $\lim_{h\to 0}\frac{U_S(\rho)-\hat{U}(\rho;h)}{h} = 0$. Lemma 1 and Lemma 2 together give an important characterization of the equilibrium value function U_S:

Proposition 2 *Fix the monopolist's strategy such that $\rho_S^\star$ is the equilibrium cutoff in phase S. If, in equilibrium, both unknown buyers continue to experiment at posterior belief $\rho > \rho_S^\star$, the value $U_S(\rho)$ satisfies differential equation (6) for $\rho \le \rho_I^\star$, and differential equation (7) for $\rho > \rho_I^\star$. Meanwhile, if the value $U_S(\rho)$ satisfies the differential equations below, then the unknown buyers will keep experimenting in equilibrium.*

As shown in the appendix, differential equation (6) is given by

$$0 = 2(r+\lambda_H\rho)(U_S(\rho)-s) + \lambda_H\rho(1-\rho)U_S'(\rho) + \frac{r\lambda_H g}{r+\lambda_H}\frac{(1-\rho)^2\rho_S^\star}{1-\rho_S^\star}\left(\frac{(1-\rho)\rho_S^\star}{\rho(1-\rho_S^\star)}\right)^{r/\lambda_H} - \frac{rg}{r+\lambda_H}\lambda_H\rho(1-\rho); \quad (6)$$

and differential equation (7) is given by

$$0 = 2(r+\lambda_H\rho)(U_S(\rho)-s) + \lambda_H\rho(1-\rho)U_S'(\rho) + (r+\lambda_H\rho)g(1-\rho)\left(\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}\right)^{r/\lambda_H} - \lambda_H g\rho(1-\rho) - r\left[\frac{r+\lambda_H+\lambda_H\rho_I^\star}{(r+\lambda_H)(1-\rho_I^\star)}\left(\frac{\rho_I^\star}{1-\rho_I^\star}\right)^{r/\lambda_H} - \frac{\lambda_H}{r+\lambda_H}\left(\frac{\rho_S^\star}{1-\rho_S^\star}\right)^{1+r/\lambda_H}\right]g(1-\rho)^2\left(\frac{1-\rho}{\rho}\right)^{r/\lambda_H}. \quad (7)$$

The necessity of Proposition 2 comes from combining Lemma 1 and Lemma 2. In the appendix, we prove the sufficiency of this result as well: given the equilibrium value function $U_S(\rho)$ and off-equilibrium value function $U^D(\rho, \rho_h)$, an experimenting buyer does not profit from any deviation. The proof proceeds in two steps. First, we show that there are no profitable "one-shot" deviations: there exists $\bar{h} > 0$ such that it is suboptimal to deviate for $h \le \bar{h}$ length of time both on and off equilibrium path. Second, similar to the one-shot deviation principle in discrete time, we prove that the non-existence of profitable one-shot deviations is sufficient to rule out profitable deviations by the definition of admissible strategies.

Since a learning value is attached to purchasing behavior, the unknown buyer is not behaving myopically. The monopolist must provide an extra subsidy to deter deviations because the deviator gains a rent from the possible advantageous belief caused by short-lived deviations: $U_S(\rho) > s$.

Fig. 1 depicts the buyer's equilibrium value function $U_S(\rho)$ as computed from Proposition 2. The equilibrium can be divided into two cases. In the "niche" case, $\rho_S^\star < \rho \le \rho_I^\star$ and hence the monopolist only sells to the known buyer after the arrival of the first lump-sum payoff; in the "mass" case, $\rho > \rho_I^\star$ and the monopolist sells to both buyers after the arrival of the first lump-sum payoff. The buyer's equilibrium value is strictly higher than the outside option even in the niche case, in which the arrival of the first lump-sum brings no rent to the receiver of the payoff ($V_I(\rho) = s$). The shape of the value function changes at $\rho_I^\star$ because of the switch from the mass case to the niche case. However, the derivative of the value function is checked to be continuous at $\rho_I^\star$.

Denote the equilibrium price in phase S to be $P_S(\rho)$. Then, the value for a buyer from purchasing

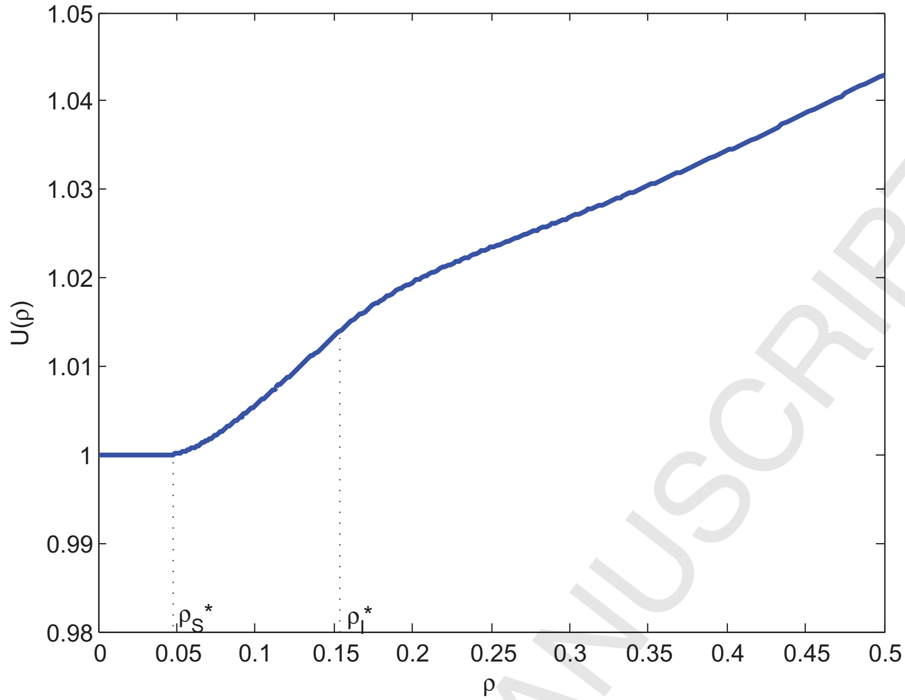

Figure 1: Buyer's Equilibrium Value Function

the risky product can be characterized by the following HJB equation:

$$rU_S(\rho) = r(\rho g - P_S(\rho)) + \lambda_H \rho(U_I(\rho) - U_S(\rho)) + \lambda_H \rho(V_I(\rho) - U_S(\rho)) - \lambda_H \rho(1-\rho)U_S'(\rho) \qquad (8)$$

where $U_I(\rho) = s$; and $V_I(\rho)$ is given by equation (2).

Meanwhile, by selling the products, the monopolist's value can be characterized as follows:

$$rJ_S(\rho) = 2rP_S(\rho) + 2\lambda_H \rho(J_I(\rho) - J_S(\rho)) - \lambda_H \rho(1-\rho)J_S'(\rho). \qquad (9)$$

where $J_I(\rho)$ is given by Proposition 1.

Equations (8) and (9) are value functions if both unknown buyers purchase the risky product. The RHS of equation (8) can be decomposed into four elements: i) the expected payoff rate from purchasing the risky product $r(\rho g - P_S(\rho))$; ii) the jump of the value function to V_I if a given buyer receives a lump-sum payoff; iii) the drop of the value function to $U_I = s$ if the other buyer receives a lump-sum payoff; and iv) the effect of Bayesian updating on the value function when no lump-sum is received. Equation (9) can be interpreted similarly.

The equilibrium price $P_S(\rho)$ can be derived by combining Proposition 2 and equation (8). By Proposition 1, in the niche case, upon the arrival of the first lump-sum payoff, the monopolist immediately shuts down experimentation and charges price $g-s$. This greatly reduces the unknown buyers' incentives to experiment. However, the monopolists will adjust the price such that $P_S(\rho)$ is still continuous at $\rho_I^\star$, since the buyer's equilibrium value is continuously differentiable at $\rho_I^\star$.

We can then substitute the price expression $P_S(\rho)$ into equation (9) and characterize the equilibrium cutoff $\rho_S^\star$ by applying value matching and smooth pasting conditions:

$$U_S(\rho_S^\star) = s,\ J_S(\rho_S^\star) = 0,\ J_S'(\rho_S^\star) = 0.$$

Proposition 3 *(Characterize the symmetric Markov perfect equilibrium) In phase S, the unknown buyers purchase the risky product at posterior belief ρ if and only if*

$$\rho > \rho_S^\star = \rho^e = \frac{rs}{rg + \lambda_H(g-s)}.$$

$\rho_S^\star = \rho^e$ implies that the equilibrium stopping rule is efficient in phase S. This result can be understood by rewriting $\rho_S^\star$ as

$$\rho = \frac{rs}{rg + \lambda_H(V_I(\rho) + J_I(\rho)) - \lambda_H s}, \tag{10}$$

where $V_I(\rho)$ and $J_I(\rho)$ are the continuation values when one buyer receives a lump-sum payoff. Since $\rho_S^\star < \rho_I^\star$, $V_I(\rho_S^\star) = s$ and $J_I(\rho_S^\star) = g - s$. When no buyer has received any lump-sum payoff, as there is no temptation to exploit the known buyer, the monopolist fully extracts the social surplus of the unknown buyer. Therefore, in equilibrium, experimentation is terminated efficiently in phase S.

It is straightforward to check that at $\rho_S^\star$, the smooth pasting condition for $U_S(\cdot)$ is also satisfied: $U_S'(\rho_S^\star) = 0$. Explicitly, the monopolist is solving an optimal stopping problem given the price that she has to charge in order to keep the unknown buyers experimenting. Implicitly, given the equilibrium pricing strategy $P_S(\cdot)$, the unknown buyers are facing an optimal stopping problem, as well. At the equilibrium cutoff, the smooth pasting condition for $U_S(\cdot)$ should also be satisfied. This fact is useful when we discuss efficiency for any $n \geq 2$ buyers because it enables us to characterize the equilibrium cutoff without solving for the value functions.

3.2.3 Equilibrium Price Path

Fig. 2 depicts different price paths in the symmetric Markov perfect equilibrium depending on how many buyers have received lump-sum payoffs. The presence of idiosyncratic uncertainty has two important implications for the equilibrium price.

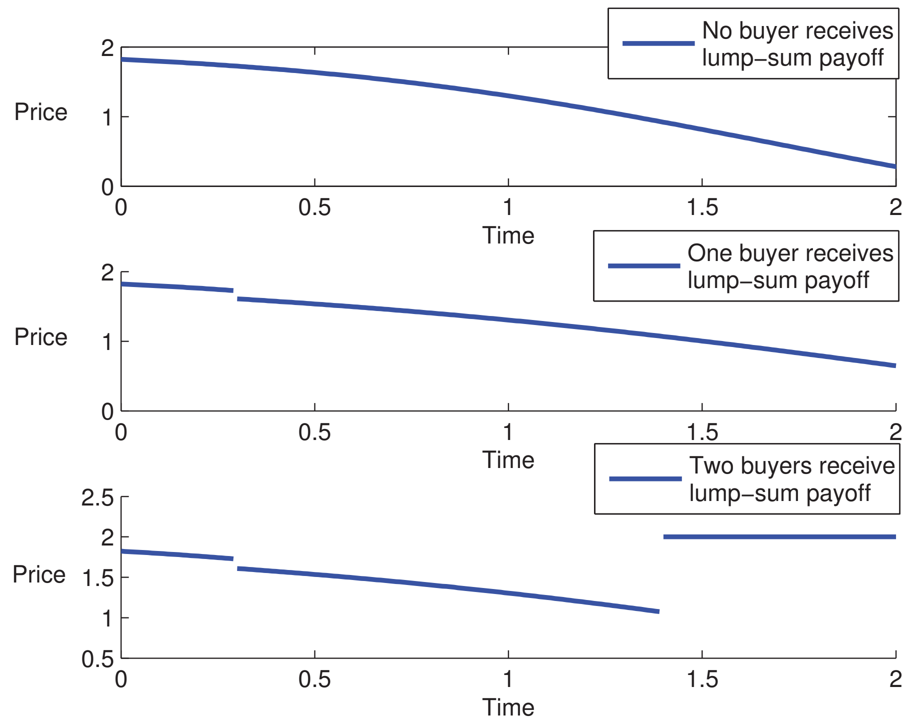

Figure 2: Equilibrium Price Dynamics

First, in phase S, assume instead that the equilibrium value for each unknown buyer is exactly s. Then the equilibrium price should be:

$$\tilde{P}_S(\rho) = \rho g - s + \frac{\lambda_H}{r}\rho(V_I(\rho) - s).$$

To deter the buyers from taking the outside option, the equilibrium value for each unknown buyer must be strictly larger than s. The actual equilibrium price $P_S(\rho)$ is strictly less than $\tilde{P}_S(\rho)$ because of this deterrence effect. Fig. 3 compares the equilibrium price path with and without the deterrence effect. This shows that the price reduction caused by the deterrence effect is quite significant. Moreover, the deterrence effect decreases over time and reaches zero when $\rho = \rho_S^\star$.

Second, there is another positive continuation value effect on the price in phase S. This is driven by ex post heterogeneity in phase I. To understand this effect, we compare the equilibrium price in phase I, $P_I(\rho)$ and the price without the deterrence effect $\tilde{P}_S(\rho)$. In the mass case ($\rho > \rho_I^\star$),

$$\tilde{P}_S(\rho) - P_I(\rho) = \frac{\lambda_H}{r}\rho(V_I(\rho) - s) > 0.$$

After the arrival of the first lump-sum payoff, the known buyer receives a rent by paying a relatively low price. Thus, in phase S, the monopolist can charge a relatively high price due to the extra

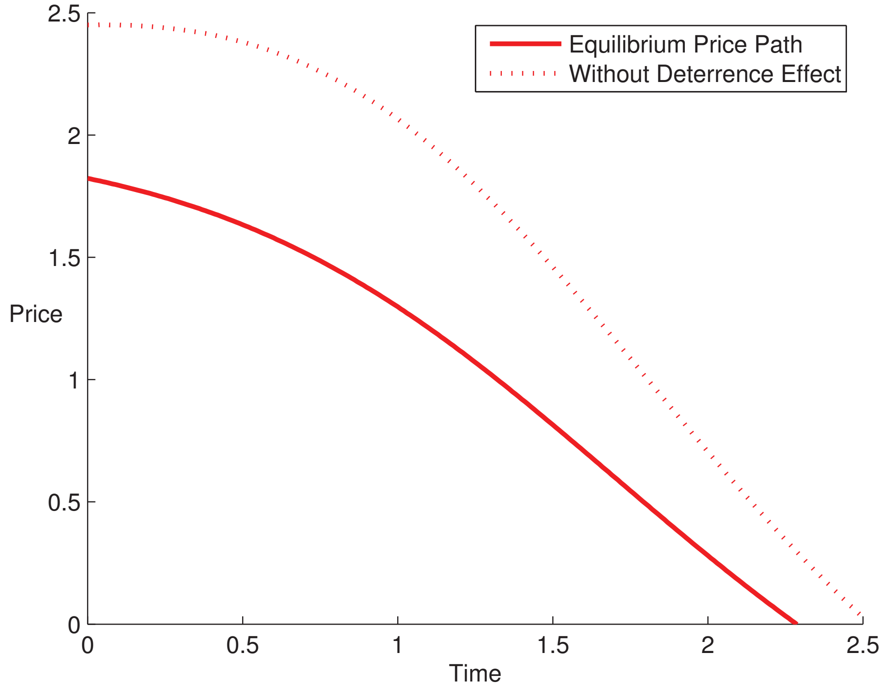

Figure 3: Deterrence Effect

incentive to experiment as a result of being the first to receive a lump-sum payoff gives access to an additional rent in phase I.

The combination of the deterrence effect and the continuation value effect implies that the instantaneous price reaction to the arrival of the first lump-sum payoff is ambiguous. In particular, when the first lump-sum payoff arrives early, there could be an instantaneous drop in price in order to encourage the unknown buyer to experiment as shown by Fig. 2.

Corollary 1 *There exists* $\bar{\rho} > \rho_I^\star$*, such that* $P_S(\rho) > P_I(\rho)$ *for* $\rho > \bar{\rho}$ *and* $P_S(\rho) < P_I(\rho)$ *for* $\rho < \bar{\rho}$*.*

The above corollary implies: in the early days of the market, ρ is higher and it is more likely to have $P_S(\rho) > P_I(\rho)$; in the late days of the market, ρ is lower and it is more likely to have $P_S(\rho) < P_I(\rho)$. Fig. 4 describes a situation in which with the same prior, the price might either drop or jump depending on when the first lump-sum payoff arrives.

3.3 Efficiency

This section discusses the efficiency property of the symmetric Markov perfect equilibrium for an arbitrary number of buyers.

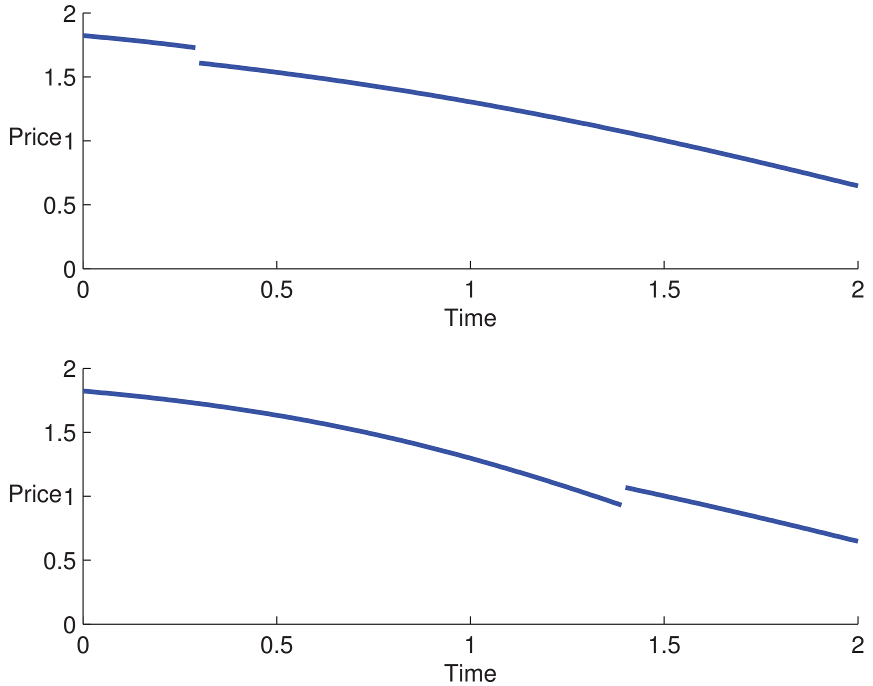

Figure 4: Instantaneous Price Response to the First Arrival of Good News

Theorem 1 *Consider a market with any $n \geq 2$ buyers. The symmetric Markov perfect equilibrium is inefficient. Moreover, the equilibrium experimentation terminates too early except when no buyer has received a lump-sum payoff.*

The result of Theorem 1 is intuitive. As in Proposition 1, when more than one buyer has received a lump-sum payoff, the monopolist faces a trade-off between exploiting those buyers who are known to be good, and exploring those buyers who have not yet determined their valuation. This trade-off always leads to inefficient early termination of exploration. Moreover, as shown in the proof of Theorem 1, the equilibrium cutoff when k buyers are known to be good satisfies

$$\rho_k^{\star} = \frac{nrs + kr(g-s)}{nrg + (n-k)\lambda_H(g-s)},$$

which is obviously increasing in k: the monopolist is more likely to stop experimentation as the temptation to exploit current known buyers increases, and is deceasing in n: the incentive to exploration becomes higher as the number of buyers grows.

We are now in a position to summarize the roles played by ex post heterogeneity. First, in phase S, ex post heterogeneity means that there is a future benefit for the deviator by becoming more optimistic than the non-deviators. The monopolist must provide an extra subsidy to deter deviations. Second, in the individual learning phase, ex post heterogeneity implies that the receivers of lump-sum payoffs are more optimistic than the unknown buyers. If the monopolist wishes to serve all buyers, the known buyers extract rents. This continuation value effect has a positive impact on the price in phase S: the monopolist can charge a relatively high price due to this extra incentive to experiment. The interaction of deterrence and continuation value effects leads to an ambiguous instantaneous price reaction to the arrival of the first lump-sum payoff. Finally, ex post heterogeneity generates a tradeoff between exploitation and exploration for the monopolist. The equilibrium experimentation level is lower than the socially efficient level as we have seen in the two-buyer case.[11]

One important implicit assumption is that the monopolist cannot commit to a certain price rule. It is common in the dynamic monopoly pricing literature to assume lack of commitment (see, e.g., the literature on the Coase conjecture). With commitment, efficiency can be easily restored by using the following "trial" contract. The monopolist offers a contract (p, T) where T is the length of the trial period and p is the lump-sum price charged at the beginning of the game. The monopolist commits that by paying p, the consumers can purchase the product for T length of time for free. After the trial period, the monopolist charges price $g - s$ and the known buyers will purchase the product.

By setting T such that the posterior belief at T, ρ_T, is the same as $\rho^e = \frac{rs}{rg+\lambda_H(g-s)}$, the monopolist achieves expected profits from each buyer:

$$\pi = (1 - e^{-rT})(\rho_0 g - s) + \rho_0(1 - e^{-\lambda_H T})e^{-rT}(g - s),$$

where $(1 - e^{-rT})(\rho_0 g - s)$ is the price charged at the beginning of the game to incentivize buyers to purchase. After time T, with probability $\rho_0(1 - e^{-\lambda_H T})$, the buyer receives at least one lump-sum payoff, and will pay $g - s$ to purchase the product.

Obviously, $\rho_T = \rho^e$ implies that experimentation stops efficiently. Meanwhile, the monopolist is willing to offer this trial contract because it extracts the full surplus received by the buyers and, hence, achieves the highest possible expected profits. One attractive property of this contract is that no buyer is incentivized to deviate given that the deviation has no impact on future price.

4 Equilibrium in the Bad News Case

In the bad news case, the arrival of lump-sum payoffs (referred to as lump-sum damages hereafter) would immediately reveal that the risky product is unsuitable for the buyer. Denote $\xi_f = A$ and $\lambda_H \xi_l = -B < 0$. Condition $A - B < s < A$ is imposed such that the risky product is superior to

[11] In the Web Appendix, it is shown that if the buyers' valuations are perfectly correlated instead of independent, there is no deterrence or continuation value effect. Furthermore, the equilibrium is always efficient since the monopolist can fully internalize the social surplus.

the safe one only when the buyers cannot receive lump-sum damages.[12]

4.1 Socially Efficient Allocation

Unlike in the good news case, a large prior ρ_0 means that the probability of receiving lump-sum damages is high and this discourages the buyers from purchasing the risky product. Therefore, instead of solving an optimal stopping problem (i.e., terminating experimentation when belief reaches a certain cutoff), in the bad news case, we solve an optimal starting problem, i.e., beginning experimentation when belief is lower than a certain cutoff.

As in the good news case, we discuss socially efficient allocation when k out of n buyers have received lump-sum damages. The social surplus function could be written as (known buyers will take the outside option and are guaranteed to receive s)

$$\Omega_k(\rho) = ks + (n-k)W(\rho)$$

where

$$W(\rho) = \sup_{\{\alpha_t\}_{t\in\mathbb{R}_+}\in\Gamma} \mathbb{E}\int_{t=0}^{\infty} re^{-rt}[\alpha_t(A-\rho_t B) + (1-\alpha_t)s]dt$$

defines the optimal control problem for the unknown buyer. Γ again denotes the set of sequences $\{\alpha_t\}_{t\in\mathbb{R}_+}$ satisfying $\alpha_t \in \{0,1\}$ for all $t\in\mathbb{R}_+$, and α_t being right continuous in t. The corresponding HJB equation is

$$W(\rho) = \max\left\{s, A-\rho B + \frac{1}{r}\left[\lambda_H\rho(s-W(\rho)) - \lambda_H\rho(1-\rho)W'(\rho)\right]\right\}. \tag{11}$$

Solve the optimal starting problem defined by equation (11) and we obtain the following result: if $k\geq 0$ buyers are known to receive lump-sum damages, it is socially efficient for those buyers to always purchase the safe product. For the remaining $n-k$ unknown buyers, it is socially efficient to start experimentation if and only if[13]

$$\rho \leq \rho^e = \frac{(r+\lambda_H)(A-s)}{\lambda_H A + rB - \lambda_H s}.$$

[12] This way of modeling the bad news case is similar to the models in [1] and [10].

[13] The solution to the differential equation

$$W(\rho) = A-\rho B + \frac{1}{r}\left[\lambda_H\rho(s-W(\rho)) - \lambda_H\rho(1-\rho)W'(\rho)\right]$$

is

$$W(\rho) = A - \frac{\lambda_H(A-s)+rB}{r+\lambda_H}\rho + C(1-\rho)(\frac{1-\rho}{\rho})^{r/\lambda_H}.$$

C must be zero since $\rho=0$ is included in the domain. ρ^e can be directly solved from $W(\rho)=s$.

4.2 Equilibrium

In any symmetric equilibrium, buyers can be divided into two groups: known buyers and unknown buyers. Let α_k^0 (α_k^1) be the strategy for the known (unknown) buyers, in which subscript k indicates the number of buyers who have received lump-sum damages. Let V_k, U_k and J_k be value functions for the known buyers, the unknown buyers and the monopolist, respectively, when k buyers have received lump-sum damages. Finally, let P_k denote the price charged by the monopolist. Definition 1 implies that the triple of $(P_k, \alpha_k^0, \alpha_k^1)$ is a symmetric Markov perfect equilibrium if:

- $\alpha_k^0 = 1$ if $P \leq A - B - s$ and $= 0$ otherwise;
- for any $k < n$, given P_k, the unknown buyers choose acceptance policy α_k^1 to maximize:

$$\begin{aligned} U_k(\rho) &= \sup_{\alpha_k^1} \mathbb{E} \int_{t=0}^{\tau} re^{-rt}[\alpha_k^1(A - \rho_t B - P_k(\rho_t)) + (1 - \alpha_k^1)s]dt \\ &+ e^{-r\tau}\left(\frac{1}{n-k}V_{k+1}(\rho_\tau) + \frac{n-k-1}{n-k}U_{k+1}(\rho_\tau)\right) \end{aligned}$$

  where τ is the first (possibly infinite) time at which a new unknown buyer receives bad news;

- given (α_k^0, α_k^1), the monopolist chooses price $P_k(\rho_t)$ to maximize

$$J_k(\rho) = \sup_{P_k} \mathbb{E}\left\{\int_{t=0}^{\tau} re^{-rt}\left[k\alpha_k^0(\rho_t, P_k(\rho_t)) + (n-k)\alpha_k^1(\rho_t, P_k(\rho_t))\right]P_k(\rho_t)dt + e^{-r\tau}J_{k+1}(\rho_\tau)\right\}$$

- beliefs update according to Bayes' rule: ρ_t satisfies the law of motion, i.e., equation (1);
- for $k = n$, the monopolist will not serve any buyer such that $J_n = 0$ and $V_n = s$.

First, a profit-maximizing monopolist should never set the price lower than $A - B - s < 0$ in order to sell to the known buyers. This implies that V_k is always s. Second, when $n - k$ unknown buyers purchase the risky product, the conditional probability that any given unknown buyer receives lump-sum damages is simply $1/(n-k)$, since the $n-k$ unknown buyers' payoff distributions are identical. Finally, once the monopolist starts to sell to the unknown buyers, she will continue to sell as long as no lump-sum damage is received.

In a symmetric Markov perfect equilibrium, when experimentation takes place on the equilibrium path, the price is set such that both the participation constraint and the no profitable deviation constraint are satisfied. In the bad news case, the deviations do not impose more restrictions than the participation constraint, because the most pessimistic unknown buyer's value is always s in equilibrium. This is different from the good news case. In the good news case, a one-shot deviation makes the non-deviators more pessimistic if they have not received any lump-sum payoffs during the deviation period. In that situation, the price charged by the monopolist is lower than the price the deviator is willing to pay. However, in the bad news case, a deviation makes the deviator more pessimistic. After the deviation, if the monopolist wishes to serve all unknown buyers, the optimal price is determined by the price the deviator is willing to pay; if the monopolist does not wish to serve all unknown buyers, the deviator is the first buyer to be excluded. In both cases, the deviator cannot gain more than the outside option after a deviation. Therefore, setting the equilibrium value to be s is adequate to deter deviations.

Given the equilibrium value for each unknown buyer is always s (since they are equally pessimistic), the equilibrium price can easily be derived: the monopolist will always charge $P_I(\rho) = A - \rho B - s$ no matter how many buyers have received lump-sum payoffs. By solving the monopolist's optimal starting problem, we obtain the following theorem:

Theorem 2 *Consider a market with $n \geq 2$ buyers. The symmetric Markov perfect equilibrium is always efficient.*

The above theorem is intuitive: the equilibrium is always efficient as the monopolist is able to extract the full surplus. The comparison of the good news and the bad news cases highlights the important role played by ex post heterogeneity. Indeed, in the bad news case, buyers can also become ex post heterogeneous by taking different actions or achieving different outcomes. As explained earlier, the deterrence effect does not exist, because taking the outside option puts the deviator in a disadvantageous position. Moreover, there is no tradeoff between exploitation and exploration because the buyers who have received lump-sum damages will never purchase the risky product. This rules out the continuation value effect as well. As a result, although buyers become ex post heterogeneous, the potential buyers of the risky product are always the unknown ones, who are ex post homogeneous in a symmetric equilibrium. Hence, the equilibrium is always efficient.

5 Conclusion

This paper considers a dynamic monopoly pricing environment in which the monopolist cannot price-discriminate among the buyers. Individual learning generates ex post heterogeneity both from different actions and different outcomes. If the monopolist wishes to sell to several buyers, the optimal price is set to make the most pessimistic buyer indifferent between the alternatives. In the good news case, this has significant implications both on the equilibrium path and off the equilibrium path. On the equilibrium path, the receivers of lump-sum payoffs become more optimistic than the non-receivers. This implies: i) the arrival of the first good news signal reduces the continuation value for the unknown buyers, and this effect might lead to an instantaneous drop in price; and ii) the monopolist faces different buyers after the arrival of lump-sum payoffs and the absence of price discrimination leads to an inefficient termination of experimentation.

There is another subtle off-equilibrium implication. By taking short-lived deviations (switching to the safe option for an instant), each buyer can extract a rent if she becomes more optimistic than other buyers after the deviation. This generates a future benefit from deviation. If the monopolist wishes to make a sale to several unknown buyers, she must provide an extra subsidy to deter deviations. This paper establishes a way to analyze off-equilibrium behavior when the buyers are asymmetric after a deviation. This analysis allows us to explicitly derive the deviator's value function and the equilibrium price path. It is shown that such a deterrence effect leads to a significant reduction in the equilibrium price.

However, in the bad news case, the above implications do not exist. This can be explained by two facts. First, on the equilibrium path, the receivers of lump-sum damages immediately take the outside option and the buyers who stay in the experience good market are still ex post homogeneous. Second, off the equilibrium path, a buyer cannot benefit from deviations because the deviator becomes more pessimistic.

In the Web Appendix, we consider the correlated case by introducing another dimension of uncertainty on product quality. A lump-sum payoff can occur only when both the product quality is high and the match quality is relevant. The analysis of the correlated case is more complicated since there is one more state variable: the belief about the product characteristic. As the belief about the product characteristic can be expressed as a function of the belief about individual match quality, we can apply similar techniques to analyze the problem. It is shown that adding this common uncertainty does not change the main results as long as idiosyncratic uncertainty still exists.

There are several extensions to consider in the future. For tractability, we have assumed that the arrival of lump-sum payoffs immediately resolves the idiosyncratic uncertainty of the receiver. We could consider a model where the arrival of lump-sum payoffs cannot immediately resolve the idiosyncratic uncertainty of the receiver as in [22]. As long as ex post heterogeneity exists, the resulting equilibrium would be inefficient as well.

Another natural extension of the current model is to consider a dynamic duopoly pricing environment. This issue is partially investigated by [6], who consider a model with a continuum of buyers such that buyers choose according to their myopic preferences at each instant. It would be interesting to consider a model with a finite number of buyers such that each buyer's choice has non-trivial effects on learning and future prices.

Appendix

A Admissible Strategies

Before formally defining admissible Markovian strategies, we define admissibility for general strategies. First denote an outcome h to be

$$h \triangleq (\{a_{it}, N_{it}\}_{i=1}^{n}, P_t)_{0\le t<\infty} ;$$

and H is the set of all possible outcomes. A sub-outcome $h^- \subset h$ only includes information about purchasing decisions and lump-sum payoffs:

$$h^- \triangleq (\{a_{it}, N_{it}\}_{i=1}^{n})_{0\le t<\infty} ;$$

and H^- is the set of all possible sub-outcomes.

In general, a strategy can be viewed as a map from the set of outcomes to actions. We focus on strategies that are independent of previous prices; this is because allowing pricing as a function of previous prices may generate more complicated problems.[14] The monopolist's pricing decision is given by the mapping:

$$P : H^- \times [0,\infty) \to \mathbb{R};$$

and the buyers' acceptance decision is given by the mapping:

$$\alpha_i : H \times [0,\infty) \to \{0,1\}.$$

$P(h^-,t)$ is the price charged by the monopolist at time t; and $\alpha_i(h,t)$ is the purchasing decision made by buyer i at time t. Assumptions A1 and A2 stated below guarantee that the strategies are well defined.

Denote vector $a = (a_1, \cdots, a_n)$ and vector $N = (N_1, \cdots, N_n)$. A metric on the sets of outcomes is defined as:

$$D^-(\hat{h}_t^-, \tilde{h}_t^-) = \int_0^t \left[d(\hat{a}_\tau, \tilde{a}_\tau) + d(\hat{N}_\tau, \tilde{N}_\tau)\right] d\tau$$

and

$$D(\hat{h}_t, \tilde{h}_t) = \int_0^t \left[d(\hat{a}_\tau, \tilde{a}_\tau) + d(\hat{N}_\tau, \tilde{N}_\tau)\right] d\tau + |\hat{P}_t - \tilde{P}_t|$$

where d is the Euclidean norm. In particular, the previous prices do not enter in the definition of $D(\hat{h}_t, \tilde{h}_t)$; only the current price matters. The metric D (D^-) determines a Borel σ-algebra $\mathcal{B}_H$ ($\mathcal{B}_{H^-}$). The first restriction on strategies is that:

A1. P is a $\mathcal{B}_{H^-} \times \mathcal{B}_{[0,\infty)}$ measurable function and α_i is a $\mathcal{B}_H \times \mathcal{B}_{[0,\infty)}$ measurable function.

The second restriction requires that the strategies take the same actions if two histories are almost the same:

A2. For all t, and $\hat{h}, \tilde{h} \in H$ such that $D(\hat{h}_t, \tilde{h}_t) = 0$, then $P(\hat{h}^-, t) = P(\tilde{h}^-, t)$ and $\alpha_i(\hat{h}, t) = \alpha_i(\tilde{h}, t)$.

A1 and A2 are two natural restrictions on strategies. Additional conditions must be imposed to guarantee that the induced outcome is unique. Before doing that, we define an outcome h to

[14] For example, any decreasing price path is consistent with the pricing function $P(h,t) = \inf_{\tau<t} P_\tau$.

be *compatible* with a given strategy profile $\{P, \alpha\}$ if h satisfies: $P(h^-, t) = P_t$ and $\alpha_i(h, t) = a_{it}$. A straightforward modification of the argument in [8] shows the following:

Proposition 4 *A strategy profile (P, α) generates a unique distribution over compatible outcomes if it satisfies:*

1. *for any outcomes $\hat{h}$ and $\tilde{h}$ and any time t such that $D(\hat{h}_t, \tilde{h}_t) = 0$ and $\hat{N}_t = \tilde{N}_t$,*

$$\lim_{\epsilon \searrow 0} P(\hat{h}, t + \epsilon) = \lim_{\epsilon \searrow 0} P(\tilde{h}, t + \epsilon);$$

   and

2. *for any $\hat{h}$ and $\tilde{h}$ and any t such that $D(\hat{h}_t, \tilde{h}_t) = 0$, $\hat{N}_t = \tilde{N}_t$ and $\lim_{\epsilon \searrow 0} \hat{P}_{t+\epsilon} = \lim_{\epsilon \searrow 0} \tilde{P}_{t+\epsilon}$, then there exists $\epsilon > 0$ and $a \in \{0, 1\}$ such that $\alpha_i(\hat{h}, \tilde{t}) = \alpha_i(\tilde{h}, \tilde{t}) = a$ for any $\tilde{t} \in (t, t + \epsilon)$.*

We say a strategy profile (P, α) is *weakly admissible* if it satisfies conditions 1 and 2 in Proposition 4. In Proposition 4, condition 2 is the key condition. This condition is slightly different from the inertia condition proposed in [8]. The modification is needed to handle the possible situation in which the arrival of a lump-sum payoff at time t results in the purchasing decisions a_t not being right continuous in time.

Any Markovian strategy profile (P, α) which induces a weakly admissible strategy profile generates a unique distribution over compatible outcomes. However, the notion of weak admissibility does not guarantee that the induced outcome allows us to use equation (1) to update beliefs.

Definition 2 *A Markovian strategy profile (P, α) is* strongly admissible *in the good news case if it satisfies:*[15]

1. *$P(\boldsymbol{\rho})$ is left continuous and non-decreasing when it is continuous: for each $\boldsymbol{\rho} \in \Sigma$ and $\delta > 0$, there exists some $\epsilon > 0$ s.t. $P(\boldsymbol{\rho}') \leq P(\boldsymbol{\rho})$ and $|P(\boldsymbol{\rho}') - P(\boldsymbol{\rho})| \leq \delta$ for all feasible $\boldsymbol{\rho}' \leq \boldsymbol{\rho}$ such that $||\boldsymbol{\rho}' - \boldsymbol{\rho}|| \leq \epsilon$;*[16]

2. *$\alpha_i(\boldsymbol{\rho}, P)$ is left continuous: for each $\boldsymbol{\rho} \in \Sigma$ and $\delta > 0$, there exists some $\epsilon' > 0$ s.t. $\alpha_i(\boldsymbol{\rho}', P') = \alpha_i(\boldsymbol{\rho}, P)$ for all feasible $(\boldsymbol{\rho}', P') \leq (\boldsymbol{\rho}, P)$ such that $||(\boldsymbol{\rho}', P') - (\boldsymbol{\rho}, P)|| \leq \epsilon'$; and*

3. *if h is a history compatible with (P, α), $\mathcal{C}(t; h) < \infty$ for $t < \infty$, where $\mathcal{C}(t; h)$ denotes the number of times τ before t such that purchasing behavior a_τ is discontinuous.*

It is straightforward to check that conditions 1 and 2 in Definition 2 are sufficient to guarantee that (P, α) induces a weakly admissible strategy profile. In addition, these two conditions imply that any outcome induced by the Markovian strategy profile (P, α) is well behaved in the sense that the purchasing decisions a_{it} and pricing decisions P_t are right continuous functions when there is no arrival of lump-sum payoffs. This enables us to use equation (1) to update beliefs. In the good news case, condition 1 implies P_t is decreasing when it is continuous but we also allow jumps in the price path. Condition 3 requires that each buyer can change actions no more than a finite number of times in a finite time interval, since condition 2 does not preclude the possibility of an

[15] For the bad news case, condition 1 should be changed to require that P is piecewise non-increasing.

[16] We write $(x_1, \cdots, x_n) \leq (y_1, \cdots, y_n)$ if $x_i \leq y_i$ for $i = 1, \cdots, n$, and $|| \cdot ||$ is the Euclidean norm.

infinite number of changes on any time interval. This additional condition is needed to simplify the analysis of the equilibrium.

Definition 2 is too strong in the sense that even cutoff strategies may not be strongly admissible.[17] We use the completion argument in [8] to overcome this issue. First define a metric on the space of strongly admissible strategies. A Markovian strategy profile (P, α) is *admissible* if there exists strongly admissible Markovian strategy profiles $\{(P_k, \alpha_k)\}_{k=1}^{\infty}$ such that $\lim_{k\to\infty}(P_k, \alpha_k) = (P, \alpha)$. An outcome h is *consistent* with an admissible strategy profile (P, α) if there exists strongly admissible Markovian strategy profiles $\{(P_k, \alpha_k)\}_{k=1}^{\infty}$ and outcomes $\{h_k\}_{k=1}^{\infty}$ satisfying the following three conditions: i) for each k, h_k is compatible with (P_k, α_k), ii) $\lim_{k\to\infty}(P_k, \alpha_k) = (P, \alpha)$ and iii) $\lim_{k\to\infty} h_k = h$. An admissible Markovian strategy profile (P, α) may not generate a unique distribution over compatible outcomes. However, the proof of Theorem 2 in [8] applies here, as well, to show that each admissible Markovian strategy profile (P, α) is identified with a unique distribution over consistent outcomes. When referring to outcomes generated by an admissible Markovian strategy profile (P, α), we restrict attention to the consistent outcomes.

In our definition of Markov perfect equilibrium, we allow the deviating strategies to be non-Markovian. Additional conditions on the non-Markovian strategies are also needed to make sure that the induced outcome will be well behaved even off the equilibrium path. The conditions imposed are counterparts of conditions 1-3 in Definition 2.

Definition 3 *Define time t as a regular time for outcome h if there is no arrival of lump-sum payoffs at time t. A weakly admissible strategy profile (P, α) is* strongly admissible *in the good news case if it satisfies:*

1. *P is right continuous and non-increasing when continuous at any regular time: for any outcomes h and any regular time t,*

$$\lim_{\epsilon \searrow 0} P(h, t + \epsilon) = P(h, t);$$

   and there exists $\bar{\epsilon}_1 > 0$ such that $P(h, t + \epsilon) \leq P(h, t)$ for all $\epsilon \leq \bar{\epsilon}_1$;

2. *for any h and any regular t such that P_t is right continuous and non-increasing at time t, there exists $\bar{\epsilon}_2 > 0$ and $a \in \{0, 1\}$ such that $\alpha_i(h, \tilde{t}) = \alpha_i(h, t)$ for any $\tilde{t} \in (t, t + \bar{\epsilon}_2)$; and*

3. *if h is a history compatible with (P, α), $\mathcal{C}(t; h) < \infty$ for $t < \infty$.*

A non-Markovian strategy profile (P, α) is *admissible* if there exists strongly admissible non-Markovian strategy profiles $\{(P_k, \alpha_k)\}_{k=1}^{\infty}$ such that $\lim_{k\to\infty}(P_k, \alpha_k) = (P, \alpha)$. For an admissible non-Markovian strategy profile (P, α), we also restrict attention to those consistent outcomes that can be similarly defined.

B Proofs of Results

B.1 Proof of Proposition 1

Proof. In phase I, denote ρ to be the common posterior belief about the unknown buyer's idiosyncratic uncertainty. Denote $P_I(\rho)$ as the price set by the monopolist for $\rho > \rho_I^\star$, where $\rho_I^\star$ is the

[17] For example, consider a cutoff strategy such that the cutoff price for buyer i is strictly increasing in beliefs and buyer i takes the risky product at the cutoff price. This strategy violates the condition that α_i is left continuous in beliefs.

equilibrium cutoff. Then, the value function for the unknown buyer satisfies

$$rU_I(\rho) = r(g\rho - P_I(\rho)) + \rho\lambda_H(s - U_I(\rho)) - \lambda_H\rho(1-\rho)U_I'(\rho).$$

Certainly, a profit-maximizing monopolist always sets prices $P_I(\rho) = g\rho - s$ such that $U_I(\rho) = s$. The monopolist's problem is to choose between charging a low price $g\rho - s$ in order to keep experimenting and charging a high price $g - s$ to extract the full surplus from the known buyer. Obviously, this is an optimal stopping problem with HJB equation

$$rJ_I(\rho) = \max\left\{r(g-s), 2r(g\rho - s) + \rho\lambda_H(2(g-s) - J_I(\rho)) - \lambda_H\rho(1-\rho)J_I'(\rho)\right\}. \qquad (12)$$

On the RHS of equation (12), $g - s$ is the value if the monopolist only sells to the good buyer by charging $g - s$; if the monopolist decides to continue experimentation, she not only receives instantaneous revenue $2(g\rho - s)$ by selling to both buyers, but may also receive a future value of $2(g - s)$ if the unknown buyer receives a lump-sum payoff. From the value matching and smooth pasting conditions, it is straightforward to characterize the equilibrium cutoff as

$$\rho_I^\star = \frac{r(g+s)}{2rg + \lambda_H(g-s)}.$$

The equilibrium value function $J_I(\rho)$ could be solved as:

$$J_I(\rho) = \begin{cases} 2(g\rho - s) + (g + s - 2g\rho_I^\star)\frac{1-\rho}{1-\rho_I^\star}[\frac{(1-\rho)\rho_I^\star}{(1-\rho_I^\star)\rho}]^{r/\lambda_H} & \text{if } \rho > \rho_I^\star \\ g - s & \text{otherwise.} \end{cases}$$

The known buyer only needs to pay $P_I(\rho) = g\rho - s < g - s$ before ρ reaches $\rho_I^\star$, but has to pay $g - s$ afterwards. The value function for this buyer is given by differential equation

$$rV_I(\rho) = r(g(1-\rho) + s) + \rho\lambda_H(s - V_I(\rho)) - \lambda_H\rho(1-\rho)V_I'(\rho) \qquad (13)$$

for $\rho > \rho_I^\star = \frac{r(g+s)}{2rg+\lambda_H(g-s)}$ and $V_I(\rho) = s$ for $\rho \le \rho_I^\star = \frac{r(g+s)}{2rg+\lambda_H(g-s)}$. Equation (13) is an ordinary differential equation with boundary condition: $V_I(\rho_I^\star) = s$. This gives us the expression of $V_I(\rho)$ in the proposition. ■

B.2 Characterize $\lim_{h\to 0}\frac{U_S(\rho)-\hat{U}(\rho;h)}{h}$

Lemma 1 *Fix prior ρ_0 and let $\rho_S^\star$ be the equilibrium cutoff in phase S. For $\rho_S^\star < \rho \le \rho_I^\star$,*

$$\begin{aligned}\lim_{h\to 0}\frac{U_S(\rho) - \hat{U}(\rho;h)}{h} &= 2(r + \lambda_H\rho)(U_S(\rho) - s) + \lambda_H\rho(1-\rho)U_S'(\rho) \\ &\quad - \frac{rg}{r+\lambda_H}\lambda_H\rho(1-\rho) + \frac{r\lambda_H g}{r+\lambda_H}\frac{\rho_S^\star(1-\rho)^2}{1-\rho_S^\star}(\frac{(1-\rho)\rho_S^\star}{\rho(1-\rho_S^\star)})^{r/\lambda_H}; \qquad (14)\end{aligned}$$

and for $\rho > \rho_I^\star$,

$$\lim_{h\to 0}\frac{U_S(\rho)-\hat{U}(\rho;h)}{h} = 2(r+\lambda_H\rho)(U_S(\rho)-s)+\lambda_H\rho(1-\rho)U_S'(\rho)$$
$$+(r+\lambda_H\rho)g(1-\rho)(\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)})^{r/\lambda_H}-\lambda_H g\rho(1-\rho)$$
$$-r\left[\frac{r+\lambda_H+\lambda_H\rho_I^\star}{(r+\lambda_H)(1-\rho_I^\star)}(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H}-\frac{\lambda_H}{r+\lambda_H}(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H}\right]g(1-\rho)^2(\frac{1-\rho}{\rho})^{r/\lambda_H}. \quad (15)$$

Proof. First notice that if $\lim_{h\to 0}\frac{U_S(\rho)-U^D(\rho,\rho_h)}{h}$ exists, $\lim_{h\to 0}\frac{U_S(\rho)-\hat{U}(\rho;h)}{h}$ can be written as:

$$\lim_{h\to 0}\frac{U_S(\rho)-\hat{U}(\rho;h)}{h}=(r+\lambda_H\rho)(U_S(\rho)-s)+\lim_{h\to 0}\frac{U_S(\rho)-U^D(\rho,\rho_h)}{h}. \quad (16)$$

The main issue is to evaluate $U^D(\rho,\rho_h)$ for $\rho>\rho_h$. We proceed in the following three steps: first, we decompose U^D as the sum of the non-deviator's value U^{ND} and the difference $Z=U^D-U^{ND}$; second, we derive an ordinary differential equation about U^{ND}, and explicitly solve it; finally, we derive an ordinary differential equation about Z with respect to time t, and explicitly solve it.

1. **Decompose off-equilibrium value function**

   Fix $h>0$ to be sufficiently small and the monopolist will still sell to both buyers after an h-deviation.[18] Therefore, there exists $\bar{h}'$ such that for all $h'\le\bar{h}'$, we have:

$$U^D(\rho,\rho_h)=\mathbb{E}\int_{t=0}^{h'}re^{-rt}(\rho_t g-\tilde{P}_t)dt$$
$$+\rho(1-e^{-\lambda_H h'})e^{-rh'}V_I(\rho_{h+h'})+\rho_h(1-e^{-\lambda_H h'})e^{-rh'}s$$
$$+[1-\rho(1-e^{-\lambda_H h'})-\rho_h(1-e^{-\lambda_H h'})]e^{-rh'}U(\rho_{h'},\rho_{h+h'}). \quad (17)$$

   In the above expression, ρ_t is the posterior about the deviator's match quality, and starts from ρ; and $\tilde{P}_t$ is the off-equilibrium price set by the monopolist after an h-deviation.

   By purchasing the risky product, the non-deviator gets value

$$U^{ND}(\rho,\rho_h)=\mathbb{E}\int_{t=0}^{h'}re^{-rt}(\rho_t' g-\tilde{P}_t)dt$$
$$+\rho(1-e^{-\lambda_H h'})e^{-rh'}s+\rho_h(1-e^{-\lambda_H h'})e^{-rh'}V_I(\rho_{h'})$$
$$+[1-\rho(1-e^{-\lambda_H h'})-\rho_h(1-e^{-\lambda_H h'})]e^{-rh'}U(\rho_{h+h'},\rho_{h'}), \quad (18)$$

   where ρ_t' is the posterior about the non-deviator's match quality, and starts from ρ_h.

   Obviously, the off-equilibrium value function $U^D(\rho,\rho_h)$ can be decomposed as

$$U^D(\rho,\rho_h)=U^{ND}(\rho,\rho_h)+Z(\rho,\rho_h)$$

[18] If the monopolist only sells to the deviator, the loss from not selling to the non-deviator is proportional to the equilibrium value $J_S(\rho_h)>0$, but the gain is proportional to $\rho-\rho_h$. As h goes to zero, the loss always dominates the gain.

where $Z(\rho, \rho_h) = U^D(\rho, \rho_h) - U^{ND}(\rho, \rho_h)$.

The fact that the ρ_h buyer purchases the risky product means that it is not profitable for her to have "one-shot" deviations:

$$U^{ND}(\rho, \rho_h) \geq \tilde{U}(h') = \int_{t=0}^{h'} re^{-rt} s dt + \rho(1 - e^{-\lambda_H h'})e^{-rh'} s + [1 - \rho(1 - e^{-\lambda_H h'})]e^{-rh'} U(\rho_h, \rho_{h'}). \quad (19)$$

Since the ρ_h buyer is more pessimistic about the probability of receiving lump-sum payoffs, the optimal off-equilibrium price $\tilde{P}$ is set such that the ρ_h buyer has incentives to experiment. Denote $\tilde{U}(\rho; \rho_h)$ as $U^{ND}(\rho, \rho_h)$ for a fixed ρ_h since ρ_h does not change in the expression of $\tilde{U}(h')$. The fact that

$$\lim_{h' \to 0} \frac{U^{ND}(\rho, \rho_h) - \tilde{U}(h')}{h'} = (r + \lambda_H \rho)\tilde{U}(\rho; \rho_h) - (r + \lambda_H \rho)s + \lambda_H \rho(1 - \rho)\tilde{U}'(\rho; \rho_h)$$

is left-continuous in ρ and ρ_h implies that in equilibrium, the following equation is satisfied:[19]

$$\lim_{h' \to 0} \frac{U^{ND}(\rho, \rho_h) - \tilde{U}(h')}{h'} = 0.$$

Thus we derive an ordinary differential equation for $\tilde{U}(\rho; \rho_h)$

$$(r + \lambda_H \rho)\tilde{U}(\rho; \rho_h) = (r + \lambda_H \rho)s - \lambda_H \rho(1 - \rho)\tilde{U}'(\rho; \rho_h). \quad (20)$$

The off-equilibrium value function $U^D(\rho, \rho_h)$ can be further decomposed as:

$$U^D(\rho, \rho_h) = \tilde{U}(\rho; \rho_h) + Z(\rho, \rho_h).$$

2. **Solve for the off-equilibrium value function $\tilde{U}(\rho; \rho_h)$.**

   Equation (20) is an ordinary differential equation with general solution:

$$\tilde{U}(\rho; \rho_h) = s + C_h \times (1 - \rho)(\frac{1 - \rho}{\rho})^{r/\lambda_H}.$$

   When $\rho = \rho_h$, the two buyers are identical and it returns to the equilibrium path: $\tilde{U}(\rho_h; \rho_h) = U_S(\rho_h)$. This boundary condition implies:

$$C_h = \frac{U_S(\rho_h) - s}{(1 - \rho_h)(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H}}. \quad (21)$$

   Since on the equilibrium path, experimentation stops at $\rho_S^\star$, the unknown buyer receives a value less than the outside value ($U_S(\rho) < s$) for $\rho < \rho_S^\star$. Equation (21) implies that the

[19] The proof is similar to the proof of Lemma 2. If it is strictly larger than zero, we can find a neighborhood of beliefs to increase price $\tilde{P}(\rho, \rho_h)$ but the buyers will still purchase the risky product. This constitutes a profitable deviation for the monopolist.

posterior ρ_h will never be lower than $\rho_S^\star$ no matter how large h is. In other words, the monopolist always stops selling to both buyers if $(\rho, \rho_h) = (f(\rho_S^\star; h), \rho_S^\star)$, where

$$f(\rho_S^\star; h) = \frac{\rho_S^\star}{\rho_S^\star + e^{-\lambda_H h}(1 - \rho_S^\star)}$$

corresponds to the the posterior about the deviator's match quality when the posterior ρ_h drops to $\rho_S^\star$.

3. **Solve for the off-equilibrium value function** $Z(\rho, \rho_h)$**.**

Denote

$$Z(t) = Z(\rho(t), \rho_h(t)) = U(\rho(t), \rho_h(t)) - U(\rho_h(t), \rho(t))$$

where $\rho(t)$ is the posterior about the deviator's match quality after t length of time beginning from ρ and ρ_h (given that no lump-sum payoff is received during this period). Similarly, $\rho_h(t)$ is the posterior about the non-deviator's match quality. The posteriors can be expressed as:

$$\rho(t) = \frac{\rho e^{-\lambda_H t}}{\rho e^{-\lambda_H t} + (1 - \rho)}, \ \rho_h(t) = \frac{\rho_h e^{-\lambda_H t}}{\rho_h e^{-\lambda_H t} + (1 - \rho_h)}.$$

Given any $t < h'$, the monopolist will sell to both buyers $\rho(t)$ and $\rho_h(t)$. Subtract equation (18) from (17) yields:

$$\begin{aligned} Z(t) = \mathbb{E} \int_0^{h''} & re^{-r\tau}(\rho_\tau g - \rho'_\tau g) d\tau \\ & + e^{-rh''}(1 - e^{-\lambda_H h''}) \left\{\rho(t)[V_I(\rho_h(t + h'')) - s] + \rho_h(t)[s - V_I(\rho(t + h''))]\right\} \\ & \qquad + e^{-rh''} \left[1 - \rho(t)(1 - e^{-\lambda_H h''}) - \rho_h(t)(1 - e^{-\lambda_H h''})\right] Z(t + h''). \end{aligned} \tag{22}$$

Let h'' go to 0 and we get an ordinary differential equation about $Z(t)$:

$$(r + \lambda_H \rho(t) + \lambda_H \rho_h(t))Z(t) - \dot{Z}(t) = H(t) \tag{23}$$

where

$$H(t) = r(\rho(t) - \rho_h(t))g + \lambda_H \rho(t)(V_I(\rho_h(t)) - s) - \lambda_H \rho_h(t)(V_I(\rho(t)) - s).$$

The value function Z can be derived by a backward procedure.

First, if both $\rho(t)$ and $\rho_h(t)$ are smaller than $\rho_I^\star$, then both $V_I(\rho(t))$ and $V_I(\rho_h(t))$ are s and $H(t) = r(\rho(t) - \rho_h(t))g$. It is straightforward to solve differential equation (23):

$$Z(t) = \frac{rg}{r + \lambda_H}(\rho(t) - \rho_h(t)) + Ce^{rt}(1 - \rho(t))(1 - \rho_h(t)). \tag{24}$$

Repeating the above procedure yields

$$Z_3(\rho, \rho_h) = \frac{rg}{r + \lambda_H}(\rho - \rho_h) + D_{h3}(1 - \rho)(1 - \rho_h)[(\frac{1 - \rho_h}{\rho_h})^{r/\lambda_H} - (\frac{1 - \rho}{\rho})^{r/\lambda_H}], \tag{25}$$

where

$$D_{h3} = -\frac{rg}{r+\lambda_H}\frac{e^{\lambda_H h}-1}{1-e^{-rh}}(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H}.$$

Second, if $\rho(t) > \rho_I^\star$ and $\rho_h(t) \leq \rho_I^\star$, then

$$H(t) = r(\rho(t)-\rho_h(t))g - \lambda_H \rho_h(t) g(1-\rho(t))(1-[\frac{(1-\rho(t))\rho_I^\star}{\rho(t)(1-\rho_I^\star)}]^{r/\lambda_H}).$$

Similarly, we solve Z as:

$$Z_2(\rho,\rho_h) = \frac{rg}{r+\lambda_H}(\rho-\rho_h) - \frac{\lambda_H g}{r+\lambda_H}\rho_h(1-\rho) + \rho_h(1-\rho)g[\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}]^{r/\lambda_H} + D_{h2}(1-\rho)(1-\rho_h)(\frac{1-\rho}{\rho})^{r/\lambda_H}. \quad (26)$$

D_{h2} is determined such that Z_2 and Z_3 coincide when $\rho = \rho_I^\star$. This gives us

$$D_{h2} = -\frac{rg}{r+\lambda_H}\left[(e^{(r+\lambda_H)h}-e^{rh})(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H} + e^{-\lambda_H h}(\frac{\rho_I^\star}{1-\rho_I^\star})^{1+r/\lambda_H}\right].$$

Finally, if both $\rho(t)$ and $\rho_h(t)$ are larger than $\rho_I^\star$, then we have already solved

$$Z_1(\rho,\rho_h) = (\rho-\rho_h)g - [(1-\rho_h)(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (1-\rho)(\frac{1-\rho}{\rho})^{r/\lambda_H}]g(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H} + D_{h1}(1-\rho)(1-\rho_h)[(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (\frac{1-\rho}{\rho})^{r/\lambda_H}]. \quad (27)$$

D_{h1} is determined such that Z_1 and Z_2 coincide when $\rho_h = \rho_I^\star$:

$$D_{h1} = \left[\frac{1}{\rho_I^\star} + \frac{(r+\lambda_H)e^{-rh}-\lambda_H - re^{-(r+\lambda_H)h}}{(r+\lambda_H)(1-e^{-rh})} + \frac{r(e^{\lambda_H h}-1)}{(r+\lambda_H)(1-e^{-rh})}\right](\frac{\rho_I^\star}{1-\rho_I^\star})^{1+r/\lambda_H} + D_{h3}.$$

After solving for $U^D(\rho,\rho_h)$, $\lim_{h\to 0}\frac{U_S(\rho)-U^D(\rho,\rho_h)}{h}$ can be evaluated directly. Substitute the results into equation (16) and we get the equations stated in Lemma 1. ■

B.3 "Binding" Incentive Constraint

Lemma 2 *Fix a prior ρ_0 such that $\rho_S^\star$ is the equilibrium cutoff in phase S. For $\rho > \rho_S^\star$, any best response by the monopolist entails:*

$$\lim_{h\to 0}\frac{U_S(\rho)-\hat{U}(\rho;h)}{h} = 0.$$

Proof. First, it is obvious that

$$\lim_{h\to 0} \frac{U_S(\rho) - \hat{U}(\rho; h)}{h} \geq 0$$

since $U_S(\rho) \geq \hat{U}(\rho; h)$ for $h \leq \bar{h}$. Suppose by contradiction that there exists ρ_1 such that

$$F(\rho_1) \triangleq \lim_{h\to 0} \frac{U_S(\rho_1) - \hat{U}(\rho_1; h)}{h} = c > 0.$$

From Lemma 1, $F(\rho)$ is left continuous in ρ, which implies that if $F(\rho_1) = c > 0$, then there exists $h^\dagger$ and ϵ_1 such that for all $h < h^\dagger$ and $\rho_1 - \epsilon_1 < \rho' < \rho_1$:

$$U_S(\rho') - \hat{U}(\rho'; h) > hc/2.$$

Choose ϵ_2 to satisfy

$$\rho_1 - \epsilon_2 = \frac{\rho_1 e^{-\lambda_H h^\dagger}}{\rho_1 e^{-\lambda_H h^\dagger} + (1 - \rho_1)}$$

and define $\hat{\epsilon} = \min\{\epsilon_1, \epsilon_2\}$. Now define a new pricing strategy such that

$$\tilde{P}_S(\rho) = \begin{cases} P_S(\rho) + \frac{c}{2} \text{ if } \rho_1 - \hat{\epsilon} < \rho \leq \rho_1 \\ P_S(\rho) \text{ otherwise.} \end{cases}$$

Obviously, under this new pricing strategy, the unknown buyer will still purchase the risky product since

$$U_S(\rho') - \hat{U}(\rho'; h) > hc/2.$$

However, the monopolist obtains a higher profit and, hence, this constitutes a profitable deviation for the monopolist. Therefore, it is impossible to have

$$\lim_{h\to 0} \frac{U_S(\rho) - \hat{U}(\rho; h)}{h} > 0$$

in equilibrium. ■

B.4 Proof of Proposition 2

Proof. The necessity part directly comes from Lemma 1 and Lemma 2. To prove the sufficiency part, we proceed in two steps. First, Lemma 3 and Lemma 4 show that there do not exist profitable continuous time analogs of one-shot deviations both on and off the equilibrium path. Second, using the definition of admissible strategy, we prove the continuous time analog of the one-shot deviation principle: the non-existence of profitable one-shot deviations is sufficient to guarantee that any admissible deviation is unprofitable.

Lemma 3 *The value functions derived are sufficient to deter one-shot deviations: there exists $\bar{h} > 0$ such that it is not profitable for an experimenting buyer to deviate for any $h \leq \bar{h}$ length of time.*

Proof. After a buyer deviates h length of time, the monopolist can either make a sale to both buyers or sell only to the deviator. If the latter is the continuation play, $U^D(\rho, \rho_h) = s$ since the optimal price only needs to satisfy the deviator's participation constraint. Since $U_S(\rho) > s$, it is immediately clear that it is not profitable to deviate. Therefore, an interesting case occurs when

the monopolist makes a sale to both buyers after an h-deviation. We must consider the following two sub-cases.
Case 1. $\rho \leq \rho_I^\star$. In this case, it is straightforward to show

$$\hat{U}(\rho;h) - s = \left[\frac{r\lambda_H e^{-(2r+\lambda_H)h}}{(2r+\lambda_H)(r+\lambda_H)} + \frac{re^{-rh}(1-e^{-\lambda_H h})}{r+\lambda_H}\right] g\rho(1-\rho)$$
$$- \frac{\left[e^{-rh}\lambda_H + r(e^{\lambda_H h}-1)\right] g}{r+\lambda_H} \frac{(1-\rho)^2\rho_S^\star}{1-\rho_S^\star}[\frac{(1-\rho)\rho_S^\star}{\rho(1-\rho_S^\star)}]^{r/\lambda_H} + D(1-\rho)^2(\frac{1-\rho}{\rho})^{2r/\lambda_H}$$

and

$$U_S(\rho) - s = \frac{r\lambda_H}{(2r+\lambda_H)(r+\lambda_H)} g\rho(1-\rho) - \frac{\lambda_H g}{r+\lambda_H} \frac{(1-\rho)^2\rho_S^\star}{1-\rho_S^\star}[\frac{(1-\rho)\rho_S^\star}{\rho(1-\rho_S^\star)}]^{r/\lambda_H} + D(1-\rho)^2(\frac{1-\rho}{\rho})^{2r/\lambda_H}.$$

In order to show $\hat{U}(\rho;h) \leq U(\rho)$, it suffices to prove for all $h \geq 0$, both

$$S(h) \triangleq \frac{\lambda_H(1-e^{-(2r+\lambda_H)h})}{2r+\lambda_H} - e^{-rh}(1-e^{-\lambda_H h})$$

and

$$T(h) \triangleq (e^{\lambda_H h} - 1 - \frac{\lambda_H(1-e^{-rh})}{r})$$

are larger than zero. Notice $S(0) = 0$, $S'(0) = 0$ and $S''(h) > 0$. Therefore, $S(h)$ is a convex function which achieves its minimum at $h = 0$. As a result, $S(h) \geq 0$ for all $h \geq 0$. Similarly, it can be shown that $T(0) = 0$, $T'(0) = 0$ and $T''(h) > 0$. Therefore, $T(h) \geq 0$ as well. Hence, there is no profitable one-shot deviation for $h > 0$.
Case 2. $\rho > \rho_I^\star$. In this case, $\rho_h > \rho_I^\star$ for sufficiently small h, and we have:

$$U_S(\rho) - \hat{U}(\rho;h) = \left[\frac{\lambda_H(1-e^{-(2r+\lambda_H)h})}{2r+\lambda_H} - e^{-rh}(1-e^{-\lambda_H h})\right] g\rho(1-\rho)$$
$$+ \left(\frac{r(e^{\lambda_H h}-1) - \lambda_H(1-e^{-rh})}{r+\lambda_H}\right)[\frac{(1-\rho)\rho_S^\star}{\rho(1-\rho_S^\star)}]^{1+r/\lambda_H} g\rho(1-\rho)$$
$$- \left[\frac{(r+\lambda_H)e^{-rh} - \lambda_H - re^{-(r+\lambda_H)h}}{r+\lambda_H} + \frac{r(e^{\lambda_H h}-1) - \lambda_H(1-e^{-rh})}{r+\lambda_H}\right][\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}]^{1+r/\lambda_H} g\rho(1-\rho).$$

Notice $\rho_h > \rho_I^\star$ implies that $[\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}]^{1+r/\lambda_H} < (e^{-\lambda_H h})^{1+r/\lambda_H}$. Hence, $U_S(\rho) - \hat{U}(\rho;h) \geq 0$ if

$$X(h) \triangleq S(h)e^{(r+\lambda_H)h} + \frac{rT(h)}{r+\lambda_H}\left([\frac{(1-\rho_I^\star)\rho_S^\star}{\rho_I^\star(1-\rho_S^\star)}]^{1+r/\lambda_H} - 1\right) - \frac{(r+\lambda_H)e^{-rh} - \lambda_H - re^{-(r+\lambda_H)h}}{(r+\lambda_H)} \geq 0.$$

It is straightforward to check that at $h = 0$,

$$e^{(r+\lambda_H)h}S(h) - \frac{rT(h)}{r+\lambda_H} - \frac{(r+\lambda_H)e^{-rh} - \lambda_H - re^{-(r+\lambda_H)h}}{r+\lambda_H} = 0,$$

implying that $X(0) > 0$. Therefore, there must exist $\bar{h}$ such that it is not profitable to deviate for $h \leq \bar{h}$ length of time as well. ■

The next step is to show after some deviations, that neither the deviator nor the non-deviator wants another one-shot deviation.

Lemma 4 *Given that the deviator has deviated h length of time in total, it is not profitable for either buyer to have another deviation: for any posterior beliefs ρ and ρ_h, there exists $\bar{h} > 0$ such that it is not profitable for a buyer to deviate for any $h' < \bar{h}$ length of time. .*

Proof. For any posterior beliefs ρ and ρ_h, the seller can either sell only to the deviator or sell to both buyers. Completely characterizing the seller's optimal decision is complicated. However, we aim to prove that no matter what the seller's optimal choice is, it is always unprofitable to have a one-shot deviation.

First, we consider the deviating incentive of the deviator. Assume that after the h-deviation, the monopolist sells only to the deviator. Then setting $U^D(\rho, \rho_h) = s$ is sufficient to deter deviations. If the monopolist sells to both buyers, then if the deviator deviates another h' length of time, the value function $\hat{U}^h(\rho; h')$ satisfies:

$$\hat{U}^h(\rho; h') - s = e^{-rh'}[1 - \rho_h(1 - e^{-\lambda_H h'})](U^D(\rho, \rho_{h+h'}) - s);$$

while if the deviator does not deviate, the value function is $U^D(\rho, \rho_h)$. Given the expressions of off the equilibrium path value function $U^D(\rho, \rho_h)$, we are also able to show that there exists $\bar{h} > 0$ such that it is not profitable to deviate for $h' \leq \bar{h}$ length of time. The intuition is clear: after the deviation, the price is set such that the non-deviator is indifferent and, hence, the deviator should strictly prefer purchasing. The proof is similar to the tedious proof of Lemma 3 and is omitted.

Second, we need to show that the non-deviator has no interest in taking on the role of the deviating buyer by deviating for another period of time. If the monopolist sells only to the deviator, it is obviously non-profitable for the more pessimistic non-deviator to purchase the risky product. We thus only need to show, if the monopolist sells to both buyers, that the ρ_h buyer will not deviate for any h' length of time. Notice that it suffices to consider $h' \leq h$ because Lemma 4 already implies that it is not optimal to deviate beyond the point when h' exceeds h. The value associated with an h'-deviation is provided by:

$$\tilde{U}(h') = \int_{t=0}^{h'} re^{-rt} s dt + \rho(1 - e^{-\lambda_H h'})e^{-rh'} s + [1 - \rho(1 - e^{-\lambda_H h'})]e^{-rh'} U^{ND}(\rho_h, \rho_{h'}).$$

Given

$$U^{ND}(\rho, \rho_h) = s + C_h \times (1 - \rho)(\frac{1 - \rho}{\rho})^{r/\lambda_H},$$

it is straightforward to show: $U^{ND}(\rho, \rho_h) \geq \tilde{U}(h')$ for all $h' \leq h$. ■

Finally, we are in a position to show that any admissible deviation is not profitable. Suppose on the contrary, there exists another admissible strategy $\tilde{\alpha}_1$ (could be non-Markovian) for buyer 1 such that the value under this strategy is higher than the equilibrium value for some ρ

$$U_1(\tilde{\alpha}_1, P^*, \alpha_2^*; \rho) - U_S(\rho) = \epsilon > 0.$$

Notice by the definition of admissible strategies, $\tilde{\alpha}_1$ can be written as the limit of a sequence of strongly admissible strategies $\tilde{\alpha}_1^k$. Take sufficiently large T and define a new strategy $\hat{\alpha}_1$ as:

$$\hat{\alpha}_1 = \begin{cases} \tilde{\alpha}_1 & \text{if } t < T; \\ \alpha_1^* & \text{if } t \geq T. \end{cases}$$

For sufficiently large T, this new strategy also generates a value higher than $U_S(\rho)$.[20] Similarly define $\hat{\alpha}_1^k$ such that ϵ'_k in Definition 3 is $\frac{1}{k}$ and, obviously, $\hat{\alpha}_1$ is the limit of $\hat{\alpha}_1^k$. For each $\hat{\alpha}_1^k$, there can be at most a finite number of deviations in a finite time interval $[0, T)$. Since

$$U_1(\hat{\alpha}_1, P^*, \alpha_2^*; \rho) = \lim_{k\to\infty} U_1(\hat{\alpha}_1^k, P^*, \alpha_2^*; \rho) > U_S(\rho),$$

$U_1(\hat{\alpha}_1^k, P^*, \alpha_2^*; \rho) > U_S(\rho)$ for sufficiently large k. This implies that for sufficiently large k, there must exist profitable one-shot deviations. However, this contradicts Lemma 3 and Lemma 4. ∎

B.5 Proof of Proposition 3

Proof. Solving the ordinary differential equations in Proposition 2 yields:

$$\begin{aligned} U_S(\rho) &= s + \frac{r\lambda_H}{(2r+\lambda_H)(r+\lambda_H)} g\rho(1-\rho) - \frac{\lambda_H g}{r+\lambda_H}\frac{\rho_S^\star(1-\rho)^2}{1-\rho_S^\star}\left(\frac{(1-\rho)\rho_S^\star}{\rho(1-\rho_S^\star)}\right)^{r/\lambda_H} \\ &+ D(1-\rho)^2\left(\frac{1-\rho}{\rho}\right)^{2r/\lambda_H}, \end{aligned} \tag{28}$$

for any $\rho_S^\star < \rho \leq \rho_I^\star$. $U_S(\rho_S^\star) = s$ then implies

$$D = \frac{\lambda_H g}{2r+\lambda_H}\left(\frac{\rho_S^\star}{1-\rho_S^\star}\right)^{1+2r/\lambda_H}.$$

Substituting this expression into equation (8) yields

$$P_S(\rho_S^\star) = \rho_S^\star g - s.$$

Then boundary conditions

$$J_S(\rho_S^\star) = 0 \quad \text{and} \quad J'_S(\rho_S^\star) = 0$$

immediately imply that $\rho_S^\star$ should satisfy equation

$$\rho = \frac{rs}{rg + \lambda_H g - \lambda_H s} = \frac{rs}{rg + \lambda_H(V_I(\rho) + J_I(\rho)) - \lambda_H s}.$$

We also need to rule out the possibility that $\rho_S^\star \geq \rho_I^\star$. Suppose, on the contrary, that this can be true. Therefore, after receiving a lump-sum payoff, the monopolist always keeps selling to both buyers. Using the same procedure, we are able to show that in this case, the value $U_S(\rho)$ satisfies differential equation

$$\begin{aligned} 0 = 2(r+\lambda_H\rho)(U_S(\rho)-s) + \lambda_H\rho(1-\rho)U'_S(\rho) + (r+\lambda_H\rho)g(1-\rho)\left(\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}\right)^{r/\lambda_H} - \lambda_H g\rho(1-\rho) \\ - \left[\frac{r+\lambda_H\rho_S^\star}{1-\rho_S^\star}\left(\frac{\rho_I^\star}{1-\rho_I^\star}\right)^{r/\lambda_H} - \lambda_H\left(\frac{\rho_S^\star}{1-\rho_S^\star}\right)^{1+r/\lambda_H}\right] g(1-\rho)^2\left(\frac{1-\rho}{\rho}\right)^{r/\lambda_H}. \end{aligned} \tag{29}$$

[20] Notice that the value that each buyer is able to get cannot exceed g. Therefore, we can choose T such that $e^{-rT}g = \epsilon/2$.

By some algebra, we obtain $\rho_S^\star$, which should also satisfy equation:

$$\rho = \frac{rs}{rg + \lambda_H (V_I(\rho) + J_I(\rho)) - \lambda_H s}.$$

For $\rho > \rho_I^\star$,

$$V_I(\rho) + J_I(\rho) > V_I(\rho_I^\star) + J_I(\rho_I^\star) = g.$$

Therefore, we have

$$\rho_S^\star = \frac{rs}{rg + \lambda_H (g - s)} < \rho_I^\star,$$

which leads to a contradiction. ■

B.6 Proof of Corollary 1

Proof. From Proposition 2 and equation (8), we can solve the equilibrium price as:

$$P_S(\rho) = \rho g - s - \frac{\lambda_H}{2r + \lambda_H} g\rho(1-\rho) + D(1-\rho)^2 (\frac{1-\rho}{\rho})^{2r/\lambda_H} \tag{30}$$

for $\rho_S^\star < \rho \leq \rho_I^\star$, and

$$P_S(\rho) = \rho g - s + \frac{\lambda_H}{2r + \lambda_H} g\rho(1-\rho) + (D - \frac{2\lambda_H g}{2r + \lambda_H} (\frac{\rho_I^\star}{1-\rho_I^\star})^{1+2r/\lambda_H})(1-\rho)^2 (\frac{1-\rho}{\rho})^{2r/\lambda_H} \tag{31}$$

for $\rho > \rho_I^\star$. The constant D is given in the proof of Proposition 3:

$$D = \frac{\lambda_H g}{2r + \lambda_H} (\frac{\rho_S^\star}{1-\rho_S^\star})^{1+2r/\lambda_H}.$$

Obviously, when $\rho_S^\star < \rho \leq \rho_I^\star$,

$$P_S(\rho) - P_I(\rho) = -g(1-\rho) - \frac{\lambda_H}{2r + \lambda_H} g\rho(1-\rho) + D(1-\rho)^2 (\frac{1-\rho}{\rho})^{2r/\lambda_H} < 0.$$

And, for $\rho > \rho_I^\star$,

$$P_S(\rho) - P_I(\rho) = \left[\frac{\lambda_H g}{2r + \lambda_H} (\frac{\rho}{1-\rho})^{1+2r/\lambda_H} + D - \frac{2\lambda_H g}{2r + \lambda_H} (\frac{\rho_I^\star}{1-\rho_I^\star})^{1+2r/\lambda_H}\right] (1-\rho)^2 (\frac{1-\rho}{\rho})^{2r/\lambda_H}.$$

As ρ approaches $\rho_I^\star$, $P_S(\rho) - P_I(\rho) < 0$ since $\rho_I^\star > \rho_S^\star$. The term

$$\frac{\lambda_H g}{2r + \lambda_H} (\frac{\rho}{1-\rho})^{1+2r/\lambda_H} + D - \frac{2\lambda_H g}{2r + \lambda_H} (\frac{\rho_I^\star}{1-\rho_I^\star})^{1+2r/\lambda_H}$$

is increasing in ρ, and is positive for ρ close to 1. Therefore, there must exist $\bar{\rho} > \rho_I^\star$ such that $P_S(\rho) - P_I(\rho) > 0$ for $\rho > \bar{\rho}$, and $P_S(\rho) - P_I(\rho) < 0$ for $\rho < \bar{\rho}$. ■

B.7 Proof of Theorem 1

Proof. Denote $\rho_k^\star$ to be the equilibrium cutoff such that at this belief, the monopolist would stop selling to the unknown buyers when $k \geq 1$ buyers have received lump-sum payoffs. Let V_k, U_k and J_k be the equilibrium value functions for the known buyers, the unknown buyers and the monopolist, respectively, when $k \geq 1$ buyers have received lump-sum payoffs. Finally, let P_k denote the price charged by the monopolist. From a backward procedure, it could be shown that:

Lemma 5 *The equilibrium cutoffs satisfy*

$$\rho_k^\star = \frac{nrs + kr(g-s)}{nrg + (n-k)\lambda_H(g-s)}$$

and

$$\rho^e < \rho_k^\star < \rho_{k+1}^\star$$

for all $0 \leq k \leq n-1$.

Proof. The lemma is proved by induction. If all of the buyers turn out to be good, then it is optimal for the monopolist to charge $g-s$ and fully extract the total surplus. If all but one buyer has already received a lump-sum payoff, the monopolist faces the same tradeoff between exploitation and exploration as in the two-buyer case. The monopolist has to charge $g\rho - s$ to keep the unknown buyer experimenting. Her value function from selling to the unknown buyer satisfies equation

$$(r + \rho\lambda_H)J_{n-1}(\rho) = nr(g\rho - s) + n\rho\lambda_H(g-s) - \lambda_H\rho(1-\rho)J'_{n-1}(\rho);$$

with boundary conditions

$$J_{n-1}(\rho_{n-1}^\star) = (n-1)(g-s) \quad \text{and} \quad J'_{n-1}(\rho_{n-1}^\star) = 0.$$

It is straightforward to see that:

$$\rho_{n-1}^\star = \frac{rs + (n-1)rg}{\lambda_H(g-s) + nrg}$$

and

$$J_{n-1}(\rho) = \max\left\{(n-1)(g-s), n(g\rho - s) + \left[(n-1)g + s - ng\rho_{n-1}^\star\right]\frac{1-\rho}{1-\rho_{n-1}^\star}\left[\frac{(1-\rho)\rho_{n-1}^\star}{(1-\rho_{n-1}^\star)\rho}\right]^{r/\lambda_H}\right\}.$$

Meanwhile, the value for the known buyers is given by:

$$V_{n-1}(\rho) = \max\left\{s, s + g(1-\rho)(1 - [\frac{(1-\rho)\rho_{n-1}^\star}{\rho(1-\rho_{n-1}^\star)}]^{r/\lambda_H})\right\}.$$

If all but two buyers have received lump-sum payoffs, the value function for the monopolist becomes:

$$J_{n-2}(\rho) = \max\left\{(n-2)(g-s), nP_{n-2}(\rho) + \frac{2\rho\lambda_H}{r}[J_{n-1}(\rho) - J_{n-2}(\rho)] - \frac{\lambda_H\rho(1-\rho)}{r}J'_{n-2}(\rho)\right\}.$$

If the monopolist sells to the unknown buyers, the price P_{n-2} is set such that the unknown buyers have an incentive to keep experimenting:

$$rP_{n-2}(\rho) = r(\rho g - U_{n-2}(\rho)) + \lambda_H\rho(s - U_{n-2}(\rho)) + \lambda_H\rho(V_{n-1}(\rho) - U_{n-2}(\rho)) - \lambda_H\rho(1-\rho)U'_{n-2}(\rho).$$

Value matching and smooth pasting conditions mean that at the equilibrium cutoff $\rho^\star_{n-2}$,

$$U_{n-2}(\rho^\star_{n-2}) = s,\ U'_{n-2}(\rho^\star_{n-2}) = 0,\ J_{n-2}(\rho^\star_{n-2}) = (n-2)(g-s) \text{ and } J'_{n-2}(\rho^\star_{n-2}) = 0.$$

The above equations imply that $\rho^\star_{n-2}$ satisfies equation

$$(n-2)(g-s) = n\left\{\rho^\star_{n-2}g - s + \frac{\rho^\star_{n-2}\lambda_H}{r}\left[V_{n-1}(\rho^\star_{n-2}) - s\right]\right\} + \frac{2\rho^\star_{n-2}\lambda_H}{r}\left[J_{n-1}(\rho^\star_{n-2}) - (n-2)(g-s)\right].$$

If $\rho^\star_{n-2} > \rho^\star_{n-1}$, then $V_{n-1}(\rho^\star_{n-2}) > s$ and $J_{n-1}(\rho^\star_{n-2}) > (n-1)(g-s)$. But this implies

$$\begin{aligned}(n-2)(g-s) &> n(\rho^\star_{n-2}g - s) + \frac{2\rho^\star_{n-2}\lambda_H}{r}(g-s)\\ &\Longrightarrow \rho^\star_{n-2} < \frac{2rs + (n-2)rg}{2\lambda_H(g-s) + nrg} < \rho^\star_{n-1} = \frac{rs + (n-1)rg}{\lambda_H(g-s) + nrg}.\end{aligned}$$

This contradicts the assumption that $\rho^\star_{n-2} > \rho^\star_{n-1}$. Therefore, it must be the case that $\rho^\star_{n-2} \leq \rho^\star_{n-1}$ such that $V_{n-1}(\rho^\star_{n-2}) = s$ and $J_{n-1}(\rho^\star_{n-2}) = (n-1)(g-s)$. It is straightforward to see

$$\rho^\star_{n-2} = \frac{2rs + (n-2)rg}{2\lambda_H(g-s) + nrg}.$$

For general $0 \leq j \leq n-1$, assume

$$\rho^\star_k = \frac{nrs + kr(g-s)}{nrg + (n-k)\lambda_H(g-s)}$$

for $k \geq j+1$. At $\rho^\star_j$,

$$j(g-s) = n\left[(\rho^\star_j g - s) + \frac{\lambda_H\rho^\star_j}{r}(V_{j+1}(\rho^\star_j) - s)\right] + \frac{(n-j)\lambda_H\rho^\star_j}{r}\left[J_{j+1}(\rho^\star_j) - j(g-s)\right].$$

It is similar to show by contradiction that it is impossible to have $\rho^\star_j > \rho^\star_{j+1}$ and hence the equilibrium cutoff can be solved as

$$\rho^\star_j = \frac{nrs + jr(g-s)}{nrg + (n-j)\lambda_H(g-s)}.$$

A standard induction argument then implies that for all $0 \leq k \leq n-1$, we would have

$$\rho^\star_k = \frac{nrs + kr(g-s)}{nrg + (n-k)\lambda_H(g-s)}$$

and it is trivial to check that

$$\rho^e < \rho^\star_k < \rho^\star_{k+1}$$

for all $0 \leq k \leq n-2$. ■

Lemma 5 means that the equilibrium stopping rule is inefficient for $k \geq 1$. However, as in Proposition 3, the equilibrium cutoff $\rho^\star_0$ is always the same as the socially efficient cutoff ρ^e. ■

B.8 Proof of Theorem 2

Proof. When k buyers have already received lump-sum damages, the monopolist chooses to sell to the unknown buyers if:

$$J_k(\rho) = (n-k)(A - \rho B - s) + \frac{1}{r}\left[(n-k)\lambda_H \rho (J_{k+1}(\rho) - J_k(\rho)) - \lambda_H \rho(1-\rho)J_k'(\rho)\right] \geq 0.$$

An induction argument is used to solve the equilibrium cutoffs. First,

$$J_{n-1}(\rho) = A - s - \frac{\lambda_H(A - s + \frac{rB}{\lambda_H})}{r + \lambda_H}\rho \geq 0$$

if and only if $\rho \leq \rho^\star_{n-1} = \rho^e$. We can guess that

$$J_k(\rho) = (n-k)\left[A - s - \frac{\lambda_H(A - s + \frac{rB}{\lambda_H})}{r + \lambda_H}\rho\right].$$

Suppose this is true for $j = k+1, \cdots, n-1$, then solving differential equation

$$J_k(\rho) = (n-k)(A - \rho B - s) + \frac{1}{r}\left[(n-k)\lambda_H \rho (J_{k+1}(\rho) - J_k(\rho)) - \lambda_H \rho(1-\rho)J_k'(\rho)\right]$$

yields

$$J_k(\rho) = (n-k)\left[A - s - \frac{\lambda_H(A - s + \frac{rB}{\lambda_H})}{r + \lambda_H}\rho\right].$$

The conjecture about $J_k(\rho)$ hence is justified by induction.

Obviously,

$$J_k(\rho) = (n-k)\left[A - s - \frac{\lambda_H(A - s + \frac{rB}{\lambda_H})}{r + \lambda_H}\rho\right] \geq 0$$

if and only if $\rho \geq \rho^e$ for all $k \geq 1$. Therefore, the symmetric Markov perfect equilibrium is efficient for any $k \leq n$. ■

References

[1] Abreu, D., P. Milgrom, and D. Pearce (1991): "Information and Timing in Repeated Partnerships," *Econometrica*, 59(6), 1713–1733.

[2] Ackerberg, D. (2003): "Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination," *International Economic Review*, 44(3), 1007–1040.

[3] Bergemann, D., and U. Hege (2005): "The Financing of Information: Learning and Stopping," *The RAND Journal of Economics*, 36(4), 719–752.

[4] Bergemann, D., and J. Välimäki (1996): "Learning and Strategic Pricing," *Econometrica*, 64(5), 1125–1149.

[5] ——— (2000): "Experimentation in Markets," *Review of Economic Studies*, 67(2), 213–234.

[6] ——— (2002): "Entry and Vertical Differentiation," *Journal of Economic Theory*, 106(1), 91–125.

[7] ——— (2006): "Dynamic Pricing of New Experience Goods," *Journal of Political Economy*, 114(4), 713–743.

[8] Bergin, J., and W. MacLeod (1993): "Continuous Time Repeated Games," *International Economic Review*, 34(1), 21–37.

[9] Bhaskar, V. (2012): "Dynamic Moral Hazard, Learning and Belief Manipulation," mimeo.

[10] Biais, B., T. Mariotti, J.-C. Rochet, and S. Villeneuve (2010): "Large Risks, Limited Liability, and Dynamic Moral Hazard," *Econometrica*, 78(1), 73–118.

[11] Bolton, P., and C. Harris (1999): "Strategic Experimentation," *Econometrica*, 67(2), 349–374.

[12] Bonatti, A. (2011): "Menu Pricing and Learning," *American Economic Journal: Microeconomics*, 3(3), 124–163.

[13] Bonatti, A., and J. Hörner (2011): "Collaborating," *American Economic Review*, 101(2), 632–63.

[14] Bose, S., G. Orosel, M. Ottaviani, and L. Versterlund (2006): "Dynamic Monopoly Pricing and Herding," *The RAND Journal of Economics*, 37(4), 910–928.

[15] ——— (2008): "Monopoly Pricing in the Binary Herding Model," *Economic Theory*, 37(2), 203–241.

[16] Crawford, G., and M. Shum (2005): "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 73(4), 1137–1173.

[17] Dixit, A. (1993): *The Art of Smooth Pasting*. Routledge, New York.

[18] Erdem, T., and M. Keane (1996): "Decision-making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 15(1), 1–20.

[19] Felli, L., and C. Harris (1996): "Learning, Wage Dynamics and Firm-Specific Human Capital," *Journal of Political Economy*, 104(4), 838–868.

[20] Hörner, J., and L. Samuelson (2013): "Incentives for Experimenting Agents," Yale mimeo.

[21] Keller, G., and S. Rady (1999): "Optimal Experimentation in a Changing Environment," *Review of Economic Studies*, 66(3), 475–507.

[22] ——— (2010): "Strategic Experimentation with Poisson Bandits," *Theoretical Economics*, 5(2), 275–311.

[23] Keller, G., S. Rady, and M. Cripps (2005): "Strategic Experimentation with Exponential Bandits," *Econometrica*, 73(1), 39–68.

[24] Li, J., N. Matouschek, M. Powell, and X. Weng (2013): "Implementing Change in Organizations," mimeo.

[25] Simon, L., and M. Stinchcombe (1989): "Extensive Form Games in Continuous Time: Pure Strategies," *Econometrica*, 57(5), 1171–1214.

[26] Strulovici, B. (2010): "Learning While Voting: Determinants of Collective Experimentation," *Econometrica*, 78(3), 933–971.

[27] Villas-Boas, J. (2004): "Consumer Learning, Brand Loyalty, and Competition," *Marketing Science*, 23(1), 134–145.

[28] Weng, X. (2013): "Can Learning Cause Shorter Delays in Reaching Agreements?," mimeo.