Dynamic Pricing in the Presence of Social Learning*

Xi Weng†

November, 2011

Abstract

This paper studies how the interplay of individual and social learning affects price dynamics. I consider a monopolist selling a new experience good over time to many buyers. Buyers learn from their own private experiences (*individual learning*) as well as by observing other buyers' experiences (*social learning*). Individual learning generates *ex post heterogeneity*, which affects the buyers' purchasing decisions and the firm's pricing strategy. When learning is through good news signals, the monopolist's incentive to exploit the known buyers causes experimentation to be terminated too early. After the arrival of a good news signal, the price could instantaneously go down in order to induce the remaining unknown buyer to experiment. When learning is through bad news signals, experimentation is efficient, since only the homogeneous unknown buyers purchase the experience good.

Keywords. Learning. Experimentation. Strategic pricing. Exponential bandit. Good news case. Bad news case.
JEL. D83. C02. C61. C73.

*I am especially grateful to George Mailath for continuous encouragement, guidance and support. I also thank David Dillenberger, Jan Eeckhout, Hanming Fang, Qingmin Liu, Steve Matthews, Guido Menzio, Tymofiy Mylovanov, Mallesh Pai, Andrew Postlewaite, Jeroen Swinkels and seminar participants at Penn and UPF for helpful comments and suggestions.
†Department of Economics, University of Pennsylvania, xweng@sas.upenn.edu.

Contents

1 Introduction

In many markets for new experience goods, the buyers are facing both common and idiosyncratic uncertainty. Take the market for new drugs, for example. The effectiveness of a new drug first depends on the unknown common quality. However, a good quality does not guarantee that the drug is effective for everybody. Each patient's idiosyncratic uncertainty also matters.[1] Patients learn from others' experiences (*social learning*) as well as their own (*individual learning*). The success of the new drug for one patient is good news about product quality, but it does not necessarily mean that the drug would also be effective for other patients.

Consider a monopolist selling a new experience good to many buyers in such a market. 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 success of the product, everyone becomes increasingly pessimistic. In order to keep the buyers purchasing the product, the price has to be reduced. How will the monopolist react when the product is revealed to be effective for one buyer? Will strategic pricing achieve an efficient allocation?

In this paper, dynamic monopoly pricing is modelled as an infinite-horizon, continuous-time process. The monopolist sells a perishable experience good. She cannot price-discriminate across buyers. At each instant of time, the monopolist first posts a 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 (modelled as another good of known characteristics). The experience good generates random lump-sum payoffs according to a Poisson process. The arrival rate of the lump-sum payoffs depends on an unknown product characteristic and an unknown individual attribute, both of which are binary. For tractability, we assume the public arrival of lump-sum payoffs immediately resolves both the common uncertainty and the idiosyncratic uncertainty of the receiver. As a result, there is a simple dichotomy of the learning process: in the *social learning phase*, the uncertainty about the product characteristic has not been resolved; in the *individual learning phase*, there is common knowledge about the product characteristic. A key feature of the model is that buyers become *ex post heterogeneous* in the individual learning phase: some buyers have received lump-sum payoffs, while others have not.

The model setting consists of 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 gives full characterizations of the symmetric Markov

[1] Although the F.D.A. conducts an extensive period of pre-launch testing in the pharmaceutical industry, some drugs enter the market with substantial uncertainty about their product qualities. For example, dietary supplements do not need to be pre-approved by the F.D.A. before entering the market. There is also a "hurry-up mechanism," which allows approval of a drug that has not yet been proved effective in thorough clinical trials but has shown promise that it might benefit patients with life-threatening diseases. A recent example is a cancer drug Avastin, which was approved by the F.D.A. based on one clinical trial (New York Times (2010)).

perfect equilibrium for both cases. In the good news case, because of the ex post heterogeneity, the interplay of individual and social learning leads to implications significantly different from the ones obtained when only social learning exists. In particular, the buyers' purchasing behavior, the equilibrium price path and efficiency all significantly differ from the pure social learning model.

In the benchmark case where there is a single buyer in the market, that buyer's purchasing decision is purely myopic. The key reason is that in this one-buyer case, the equilibrium price is set such that the buyer is indifferent between purchasing the experience good and taking the outside option. The buyer's continuation value is independent of the learning outcomes. Since learning is not valuable, the buyer only compares the instantaneous cost and benefit when making the purchasing decisions.[2] With many buyers, this property also holds when the buyers' payoffs are perfectly correlated, but it no longer applies when the buyers' payoffs are only partially correlated. Consider a situation where two ex ante identical unknown buyers make different purchasing decisions (an "unknown" buyer refers to a buyer whose value of the good has not been fully revealed). One buyer keeps 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 what she is willing to pay. This implies that with multiple buyers and partial payoff correlations, there could be non-trivial intertemporal incentive considerations in making the purchasing decisions.

We first characterize the symmetric Markov perfect equilibrium when there are two buyers. In the social learning phase – when no lump-sum payoff has arrived yet – the critical tradeoff for the monopolist is between selling to both buyers and exiting the market; in the individual learning phase – after lump-sum payoffs have arrived to one buyer – the critical tradeoff is between selling to both buyers and selling only to the known buyer who has received lump-sum payoffs. In both learning phases, the equilibrium purchasing behavior is determined by a cutoff in the posterior belief about the unknown buyer's individual attribute. Each unknown buyer purchases the experience good above this cutoff and takes the outside option below this cutoff.

By comparing cutoffs in different learning phases, we distinguish a mass market from a niche market. The cutoff in the social learning phase is higher than the cutoff in the individual learning phase in a mass market, but lower in a niche market. Along the equilibrium path, in a mass market, the monopolist always sells to both buyers after the arrival of the first lump-sum payoff; in a niche market, if the first lump-sum payoff arrives too late, experimentation by the unknown buyer will

[2] In a dynamic duopoly pricing model (e.g., Bergemann and Välimäki (1996)), learning determines the future competition positions of different sellers. The buyer generally is not making myopic decisions since her continuation value varies with posterior beliefs. But if one seller's price is fixed to a constant, the buyer's optimal decisions become purely myopic in the framework of Bergemann and Välimäki (1996).

be immediately terminated. When experimentation by the unknown buyer occurs in the individual learning phase, the equilibrium price is set the same as in the one-buyer case. Although the unknown buyer is indifferent between the alternatives, the known buyer receives a larger consumer surplus, since she is more optimistic about the experience good than the unknown buyer.

The presence of idiosyncratic uncertainty has two important implications for the equilibrium price. First, in the social learning phase, since there is a future benefit by taking the outside option for a small amount of time, each unknown buyer receives a value higher than the outside option to deter deviation. This *deterrence effect* forces the monopolist to reduce the price in order to provide the extra subsidy. Second, it also affects how price responds to the arrival of lump-sum payoffs. In particular, when the first lump-sum payoff arrives, there might be an instantaneous drop in price. This is driven by two opposing effects on the unknown buyer's reservation value. On the one hand, the arrival of a good news signal makes the unknown buyer more optimistic. This *informational effect* raises the unknown buyer's reservation value. On the other hand, the unknown buyer loses the chance of becoming the first known buyer. The resulting loss of rents lowers the unknown buyer's reservation value. This *continuation value effect* is driven by ex post heterogeneity. If the buyers' payoffs are perfectly correlated, there is no such effect, and the equilibrium price always goes up after the arrival of the first lump-sum payoff.

If the buyers' payoffs are perfectly correlated, efficiency is achieved for any number of buyers since the monopolist is able to fully internalize the social surplus by subsidizing experimentation. However, if the buyers' payoffs are only partially correlated, the equilibrium experimentation level is always lower than the socially efficient one. This is due to the existence of ex post heterogeneity: the known buyers are willing to pay more than the unknown buyers in the individual learning phase. Without price discrimination, the monopolist faces a tradeoff between exploitation of the known buyers and exploration for a higher future value. The exploitation incentive always causes experimentation to be terminated too early. The inefficiency in the individual learning phase reduces the monopolist's incentives to subsidize experimentation in the social learning phase. As a result, the equilibrium experimentation is inefficiently low in the social learning phase as well.

We then characterize the symmetric Markov perfect equilibrium in the bad news case. It is shown that the equilibrium is always efficient as is the case when the buyers' payoffs are perfectly correlated. The key insight is that although buyers become heterogeneous in the individual learning phase, the buyers who have received lump-sum damages will never purchase the experience good. The potential buyers are only the unknown ones, who are ex post homogeneous in a symmetric equilibrium. Another important difference between the good and bade news cases is that no extra subsidy is needed in the bad news case since deviations of an unknown buyer make the deviator more pessimistic. As a result, there is no deterrence effect and no continuation value effect. The instantaneous price reaction to the arrival of the first lump-sum damage is always to go down.

The presence of multi-dimensional beliefs complicates the analysis significantly: the posterior belief about the product characteristic and the posterior beliefs about the individual attributes are all relevant for decision-making. The dimension of the state space is reduced by the fact that given the priors, the posterior about the product characteristic is a function of the posteriors about the individual attributes. When considering the symmetric Markov perfect equilibrium, on the equilibrium path, one posterior is sufficient to represent all the posteriors. But off the equilibrium path, the deviations lead to heterogeneous posterior beliefs about the individual attributes. Even in that case, the problem is transformed in a way such that all value functions can be explicitly derived by solving ordinary differential equations. The benefit of this approach is to ensure that the traditional value matching and smooth pasting conditions can still be applied to characterize the optimal stopping decisions.

Related Literature

Bergemann and Välimäki (1996) and Felli and Harris (1996) are two early papers analyzing the impact of price competition on experimentation. They show that if there is only individual learning, the dynamic duopoly competition with vertically differentiated products can achieve efficiency. However, Bergemann and Välimäki (2000) show that in the presence of social learning, the dynamic duopoly competition cannot achieve efficiency. Bergemann and Välimäki (2002) and Bonatti (2009) allow *ex ante heterogeneity* in the sense that buyers are different in their willingness to pay.[3] Both papers assume a continuum of buyers. At each instant of time, an individual buyer only makes a myopic optimal choice and strategic interactions between the buyers don't exist.

Bergemann and Välimäki (2006) also consider a dynamic monopoly pricing problem, but with a continuum of buyers and independent valuations. The difference in crucial modelling assumptions leads them to investigate different properties of equilibrium price path. The framework of a continuum of buyers makes it impossible to discuss the impact of a single good news signal on price. Instead, Bergemann and Välimäki (2006) are more concerned about whether price would always go down or eventually go up in equilibrium. Bose, Orosel, Ottaviani, and Versterlund (2006) and Bose, Orosel, Ottaviani, and Versterlund (2008) develop another way of modelling dynamic monopoly pricing under social learning. Their model is closer to the herding literature: each short-lived buyer makes a purchasing decision in a pre-determined sequence. In contrast, in our model, all buyers are long-lived and are making purchasing decisions repeatedly.

This paper is also closely connected to the continuous-time strategic experimentation literature. A nonexhaustive list of related papers includes Bolton and Harris (1999), Keller and Rady (1999), Keller and Rady (2010) and Keller, Rady, and Cripps (2005).[4] The analysis of our model setting is

[3] Villas-Boas (2004) also investigates a duopoly model with *ex ante heterogeneity* along a location. He considers a two-period model and is mainly concerned about consumer loyalty, i.e., whether in the second period, buyers return to the seller they bought from in the first period.

[4] The strategic experimentation framework is also used as a building block to investigate broader issues. For example, Strulovici (2010) investigates voting in a strategic experimentation environment; Bergemann and Hege

greatly simplified by the use of exponential bandits, building on Keller, Rady, and Cripps (2005). Most of the papers in the strategic experimentation literature assume a common value environment, where 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 a partial payoff correlation, we introduce multi-dimensional posterior beliefs and show that the dimensionality of the problem can be reduced by expressing one posterior as a function of other posteriors.

In addition to the theoretical body of work, there are a few empirical studies attempting to quantify the importance of learning considerations on consumers' dynamic purchasing behavior. However, most of the existing works have exclusively focused on modelling individual consumer behavior and analyzing the impact of idiosyncratic uncertainty (see, e.g., Ackerberg (2003), Crawford and Shum (2005), Erdem and Keane (1996) and so on). Several recent works, including Ching (2010), Chintagunta, Jiang, and Jin (2009), Kim (2010), use both individual learning and social learning to investigate the diffusion of new drugs. In particular, Ching's paper is based on the passage of the Hatch-Waxman Act in 1984. This act eliminates the clinical trial study requirements for approving generic drugs and encourages more entries of generic drugs that have uncertain product qualities. Ching shows that both individual learning and social learning are needed to explain the slow diffusion of generic drugs into the market.

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

Time $t \in [0, +\infty)$ is continuous. The market consists of $n \geq 2$ buyers indexed by $i = 1, 2, \cdots, n$ and one monopolist, who are all risk-neutral with the common discount rate $r > 0$. The monopolist with a zero cost of production sells a *risky product* with unknown value. 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 $s > 0$.[5] 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 arrival of lump-sum payoffs depends on both an intrinsic characteristic of the product (*common uncertainty*) and the quality of the match between the product and that buyer (*idiosyncratic uncertainty*). The product characteristic is either *high*

(2005), Hörner and Samuelson (2009) and Bonatti and Hörner (2009) consider moral hazard problems when effort affects speed of learning.

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

($\lambda = \lambda_H$) or low ($\lambda = \lambda_L = 0$), and the match between buyer i and the risky product 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 modelled as a Poisson process with intensity $\lambda\kappa_i$. Therefore, a buyer i is able to receive random lump-sum payoffs if and only if both the product characteristic is high and the individual match quality is relevant. Before the game starts, nature chooses randomly and independently the product characteristic and the individual match quality for each buyer. The common priors are such that: $q_0 = \Pr(\lambda = \lambda_H)$, and for each buyer i, $\rho_0 = \Pr(\kappa_i = 1)$. 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, ρ_0 and q_0 are common knowledge.

We consider two cases in 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 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 the common characteristic and 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 and the product characteristic is high. The non-arrival of lump-sum payoffs may be due to either a low characteristic or an irrelevant match. Social learning is important because it provides additional information about the product characteristic even if the buyers' match qualities are drawn independently. Although the assumption $\lambda_L = 0$ seems a little restrictive, the current model is rich enough to include the extreme cases of common value ($\rho_0 = 1, q_0 < 1$) and independent values ($q_0 = 1, \rho_0 < 1$).

At each instant of time t, the monopolist first announces a price based on the previous history and then each buyer decides which product to purchase conditional on the previous history and the announced price. It is assumed that the monopolist cannot price-discriminate and so charges the same price to all buyers.

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; $a_{it} = 0$ if buyer i purchases the safe product at time t. A 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 beliefs are defined as:

$$q_t \triangleq \Pr[\lambda_H \mid h_t] \quad \text{and} \quad \rho_{it} \triangleq \Pr[\kappa_i = 1 \mid \lambda_H, h_t]$$

such that the posterior belief of receiving lump-sum payoffs is given by

$$\Pr[\lambda\kappa_i = \lambda_H \mid h_t] = \rho_{it} q_t.$$

Given a pair of priors (ρ_0, q_0), the posteriors $(\rho_{1t}, \cdots, \rho_{nt}, q_t)$ evolve according to Bayes' rule. A buyer i who has not received any lump-sum payoff before time t expects an arrival of lump-sum payoffs from the risky product with rate $\lambda_H a_{it}\rho_{it}q_t$. 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:[6]

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

If no buyer has received a lump-sum payoff, then with an expected arrival rate $\lambda_H q_t \sum_{i=1}^n a_{it}\rho_{it}$, some buyer receives a lump-sum payoff and q_t jumps to 1. Otherwise, q_t obeys the following differential equation at those times when a_{it} is right continuous for $\forall i$:

$$\dot{q}_t = -\lambda_H q_t(1-q_t)\sum_{i=1}^n a_{it}\rho_{it}. \tag{2}$$

The posterior belief q can be expressed as a function of ρ_i's. When no buyer has received a lump-sum payoff for a length of time t, let $x_{it} \triangleq \rho_0 e^{-\lambda_H\int_0^t a_{i\tau}d\tau} + 1 - \rho_0$ denote the probability of the event that unknown buyer i has not received lump-sum payoffs for a length of time t conditional on λ_H. By Bayes' rule

$$q_t = \frac{q_0\prod_{i=1}^n x_{it}}{q_0\prod_{i=1}^n x_{it} + 1 - q_0}. \tag{3}$$

[6] If buyer i has not received good news within time t and $t+h$, then the posterior belief $\rho_{i,t+h}$ could be 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 $\tau \le \bar{h}$. Hence by definition,

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

$\dot{q}_t$ is derived similarly.

From equation (1),

$$\rho_{it} = \frac{\rho_0 e^{-\lambda_H \int_0^t a_{i\tau} d\tau}}{x_{it}} \Longrightarrow 1 - \rho_{it} = \frac{1-\rho_0}{x_{it}}. \tag{4}$$

Substituting (4) into (3) yields:

$$q_t = \frac{q_0(1-\rho_0)^n}{q_0(1-\rho_0)^n + (1-q_0)\prod_{i=1}^n (1-\rho_{it})}. \tag{5}$$

Notice that equation (5) also holds when at least one buyer has received lump-sum payoffs. In that situation, at least one of the ρ_{it}'s is one and q_t is also one. After long history of no realization of lump-sum payoffs, the posteriors ρ_{it} would converge to zero while q_t would not. This reflects the fact that ρ_{it} is a conditional probability and q_t is bounded below by $q_0(1-\rho_0)^n$.

A nice property about equation (5) is that it only depends on ρ_{it}'s and does not explicitly depend on previous purchasing decisions or time t. Differential equations (1) and (2) imply: given a particular history of purchasing decisions, both ρ_{it} and q_t can be written as a function of time. In the critical history when nobody has received lump-sum payoffs, ρ_{it} is sufficient to encode time t and the relevant information about previous purchasing decisions, which are needed for the the updating of q_t. Therefore, we are able to express q_t as a function of $\boldsymbol{\rho}_t \triangleq (\rho_{1t}, \cdots, \rho_{nt})$ for a given pair of priors (ρ_0, q_0).

2.2 Strategies and Payoffs

Throughout the paper, we focus on symmetric Markov perfect equilibria. The natural state variables include a posterior about common uncertainty q and posteriors about idiosyncratic uncertainty $\boldsymbol{\rho}$. Given a pair of priors (ρ_0, q_0), it suffices to use posterior beliefs $\boldsymbol{\rho}_t$ as state variables since q can be expressed as a function of $\boldsymbol{\rho}$. This enables us to reduce the dimensionality of the state space by one. The state variable $\boldsymbol{\rho}_t$ is required to be feasible in the sense that

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

Purchasing Decision Given a pair of priors (ρ_0, q_0), buyer i's acceptance policy is a function of states $\boldsymbol{\rho}$ and price P

$$\alpha_i : \Sigma \times \mathbb{R} \to \{0,1\}.$$[7]

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

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

The choice of a_{it} affects not only flow utility but also how beliefs $\boldsymbol{\rho}_t$ and q_t are updated. Given

[7] More accurately, the strategy should be written as $\alpha_i(\boldsymbol{\rho}, P; \rho_0, q_0)$. Throughout the paper, (ρ_0, q_0) will be dropped since no confusion is caused.

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

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

where the expectation is taken over $\{\boldsymbol{\rho}_t : t \in [0, \infty)\}$ with $\boldsymbol{\rho}_0 = \boldsymbol{\rho}$ and $q(\boldsymbol{\rho}_t)$ is given by equation (5).

Pricing Decision Given a pair of priors (ρ_0, q_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 the purchasing decisions and so 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 $\boldsymbol{\rho}_0 = \boldsymbol{\rho}$.

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. Some restrictions on strategies have to be imposed to overcome this issue. In particular, we require the Markovian strategy profile (P, α) to 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 when there is no arrival of lump-sum payoffs.

2.3 Symmetric Markov Perfect Equilibrium

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

Definition 1 *Given a pair of priors (ρ_0, q_0), an admissible Markov strategies 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]

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

[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.

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 favorable to the receiver of this payoff. In this section, we normalize $\xi_f = 0$ and $\xi_l = v > 0$. Assumption 1 implies $g \triangleq \lambda_H v > s > 0$.

Since the arrival of one lump-sum payoff immediately resolves common uncertainty, there are only two situations to consider: a *social learning phase*, where the common uncertainty has not been resolved, and an *individual learning phase*, where the common uncertainty has been resolved. In the individual learning phase, an unknown buyer just needs to learn her individual match quality and for such a buyer i, without the arrival of a lump-sum payoff, posterior belief ρ_i is updated according to equation (1).

In the social learning phase, both individual learning and social learning exist. If unknown buyers behave symmetrically, they share the same posterior belief ρ, and belief q about λ_H is given by equation (5):

$$q = \frac{(1-\rho_0)^n q_0}{(1-\rho_0)^n q_0 + (1-\rho)^n (1-q_0)}. \tag{6}$$

Therefore, in a symmetric Markov perfect equilibrium, it suffices to use the common posterior belief ρ as the unique state variable.

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 where payoffs are given by the aggregate surplus.

Given the priors ρ_0 and q_0, the socially efficient allocation is characterized by a cutoff strategy in posterior belief ρ. There are two cutoffs ρ_I^e and ρ_S^e for the individual learning phase and the social learning phase, respectively. In the individual (social) learning phase, it is optimal for the social planner to keep the unknown buyers experimenting until belief drops to ρ_I^e (ρ_S^e) and no lump-sum payoff has been received before that. A backward procedure is used to solve for the socially efficient allocation. We first characterize the socially efficient allocation in the individual learning phase and then use the optimal social surplus function in the individual learning phase to solve the cooperative problem in the social learning phase.

Socially Efficient Allocation in the Individual Learning Phase In the individual learning phase, suppose k buyers have received good news; then it is socially optimal for them to keep purchasing the risky product by assumption 2 and the social surplus function is

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

where

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

is the optimal value for an unknown buyer with posterior belief ρ.

Since the unknown buyers are facing a standard independent two-armed bandit problem, previous research (see Keller, Rady, and Cripps (2005)) has characterized the optimal cutoff and value function W. It is efficient for the remaining $n-k$ unknown buyers to stop purchasing the risky product once the posterior belief ρ reaches

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

and still no lump-sum payoff has been received. Since in the individual learning phase, the common uncertainty has been resolved ($q=1$), the efficient cutoff ρ_I^e does not depend on the priors ρ_0 and q_0. The value function for a buyer with posterior belief ρ is

$$W(\rho) = \max\left\{ s, g\rho + \frac{\lambda_H s}{r+\lambda_H}\left(\frac{rs}{(r+\lambda_H)(g-s)}\right)^{r/\lambda_H}(1-\rho)\left(\frac{1-\rho}{\rho}\right)^{r/\lambda_H} \right\}. \tag{7}$$

Efficiency in the Social Learning Phase In the social learning phase, the socially efficient allocation solves the symmetric cooperative problem (see claim 2 in the appendix):

$$\Omega_S(\rho) = \sup_{\alpha(\cdot)\in\{0,1\}} \mathbb{E}\left\{ \int_{t=0}^{h} re^{-rt} n[\alpha(\rho_t)\rho_t q(\rho_t) g + (1-\alpha(\rho_t))s]dt + e^{-rh}\Omega(\rho_h \mid \alpha) \right\}$$

where

$$\begin{aligned}\mathbb{E}\Omega(\rho_h \mid \alpha) = q\sum_{k=1}^{n} \binom{n}{k} \rho^k \left(1 - e^{-\lambda_H \int_0^h \alpha_t dt}\right)^k \left(\rho e^{-\lambda_H \int_0^h \alpha_t dt} + 1 - \rho\right)^{n-k} \Omega_k(\rho_h) \\ + \left[q\left(\rho e^{-\lambda_H \int_0^h \alpha_t dt} + 1 - \rho\right)^n + 1 - q\right] \Omega_S(\rho_h)\end{aligned}$$

and

$$\rho_h = \frac{\rho e^{-\lambda_H \int_0^h \alpha_t dt}}{\rho e^{-\lambda_H \int_0^h \alpha_t dt} + 1 - \rho}.$$

In the continuous time framework, the probability that more than two buyers receive lump-sum payoffs at the same time is zero. The Hamilton-Jacobi-Bellman equation (HJB equation hereafter) for the above problem hence is simplified as:

$$r\Omega_S(\rho) = \max\left\{rns, rn\rho q(\rho)g + n\rho q(\rho)\lambda_H(\Omega_1(\rho) - \Omega_S(\rho)) - \lambda_H\rho(1-\rho)\Omega_S'(\rho)\right\}, \tag{8}$$

where $\Omega_1(\rho) = g + (n-1)W(\rho)$ is the social surplus when one buyer receives a lump-sum payoff.

The first part of the maximand corresponds to using the safe product, the second to the risky product. The effect of using the risky product for the social planner can be decomposed into three elements: i) the (normalized) expected payoff rate $rn\rho q(\rho)g$, ii) the jump of the value function to $\Omega_1(\cdot)$ if one buyer receives a lump-sum payoff, which occurs at rate $n\lambda_H$ with probability $pq(\rho)$, and iii) the effect of Bayesian updating on the value function when no lump-sum payoff is received. When no lump-sum payoff is received, both ρ and q are updated. The updating of q is implicitly incorporated as a function of ρ.

The optimal cutoff ρ_S^e is pinned down by solving the following differential equation:

$$r\Omega_S(\rho) = rn\rho q(\rho)g + n\rho q(\rho)\lambda_H(\Omega_1(\rho) - \Omega_S(\rho)) - \lambda_H\rho(1-\rho)\Omega_S'(\rho), \tag{9}$$

with boundary conditions:

$$\Omega_S(\rho_S^e) = ns \quad \text{(value matching condition)} \quad \text{and} \quad \Omega_S'(\rho_S^e) = 0 \quad \text{(smooth pasting condition)}.$$

Substitute the two boundary conditions into differential equation (9) and we immediately show that the cutoff ρ_S^e should satisfy

$$rn\rho q(\rho)g + n\rho q(\rho)\lambda_H\Omega_1(\rho) = (r + n\rho q(\rho)\lambda_H)ns. \tag{10}$$

In the appendix, we show that equation (10) implies a unique solution ρ_S^e for a given pair of priors (ρ_0, q_0). The socially efficient allocation in the social learning phase can be characterized as follows:

Proposition 1 *(Characterize socially efficient allocation) For any posteriors (ρ, q), it is socially efficient to purchase the risky product in the social learning phase if and only if*

$$\rho q > \frac{rs}{(r+\lambda_H)g + (n-1)\lambda_H W(\rho) - n\lambda_H s}.$$

When the common uncertainty is resolved, it is always socially efficient for the unknown buyers to continue experimentation until the posterior reaches ρ_I^e.

Proof. In the appendix. ■

Given the priors, the unique pair of efficient cutoffs $(\rho_S^e(\rho_0, q_0), q_S^e(\rho_0, q_0))$ is determined by equations

$$q_S^e = \frac{(1-\rho_0)^n q_0}{(1-\rho_0)^n q_0 + (1-\rho_S^e)^n(1-q_0)} \tag{11}$$

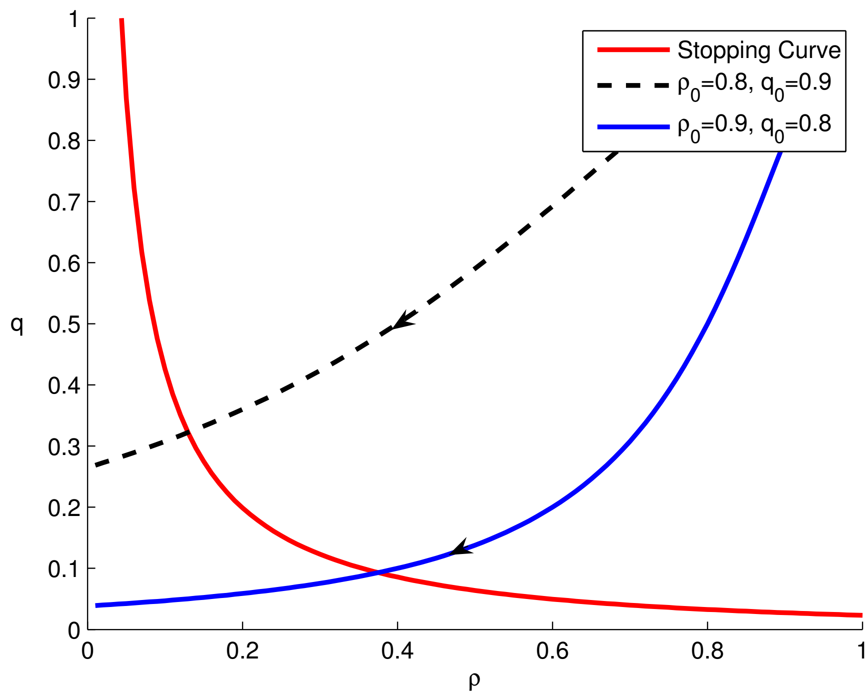

Figure 1: Solutions to the Cooperative Problem with Two Players

and

$$q_S^e = \frac{rs}{\rho_S^e[(r+\lambda_H)g + (n-1)\lambda_H W(\rho_S^e) - n\lambda_H s]}, \tag{12}$$

where $W(\cdot)$ is given by equation (7). Figure 1 is an illustration of how we can use equations (11) and (12) to determine the efficient cutoffs in the social learning phase. Equation (12) describes a stationary stopping curve because it consists of all pairs of stopping cutoffs (ρ_S^e, q_S^e) and this equation is independent of priors (ρ_0, q_0). Equation (11) describes how ρ and q evolve jointly over time starting from ρ_0 and q_0. This equation indeed depends on priors.

Unlike the individual learning phase, the cutoff ρ_S^e does depend on the priors (ρ_0, q_0). We formulate the problem so that ρ is the unique state variable in order to avoid solving partial differential equations. But the actual optimal stopping decision depends not only on belief ρ but also on q. For a fixed ρ_0, a higher q_0 means that the society can afford to experiment more and thus the efficient cutoff ρ_S^e should be lower. For a fixed pair of priors (ρ_0, q_0), a two-dimensional optimal stopping problem is transformed into a one-dimensional one by expressing q as a function of ρ. As a result, we are able to apply traditional value matching and smooth pasting conditions to solve our optimal stopping problems.

3.2 Characterizing Equilibrium for $n = 2$

In the two-buyer case, there are three situations to consider. When the common uncertainty is not resolved, denote U_S as the value function for each unknown buyer; and J_S as the value function for the monopolist. When one buyer has received lump-sum payoffs, 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 for 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 $P \leq g - s$ and $= 0$ otherwise;

- for $\zeta = S$, the unknown buyers choose acceptance policy α^1_ζ to maximize:

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

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

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

  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:

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

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

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

- beliefs update according to Bayes' rule: ρ_t satisfies the law of motion, i.e., equation (1); $q_\zeta(\rho_t) = 1$ for $\zeta = I$ and $q_\zeta(\rho_t)$ is given by equation (6) for $\zeta = S$;

- when both buyers have received 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 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 turn out to be good, it is optimal for the monopolist charging price $g - s$ to extract all of the surplus.

3.2.1 Niche Market vs. Mass Market

As in the social planner's problem, the equilibrium purchasing behavior can be characterized by two cutoffs $\rho_S^\star$ and $\rho_I^\star$. If no buyer has received lump-sum payoffs, the price is falling over time to keep both unknown buyers experimenting until posterior ρ reaches $\rho_S^\star$. After that, both buyers purchase the safe product. If one buyer has received lump-sum payoffs, the monopolist stops selling to the unknown buyer and only serves the known buyer when posterior belief about the unknown buyer is below $\rho_I^\star$.

The efficient cutoff in the individual learning phase ρ_I^e is always smaller than the efficient cutoff in the social learning phase ρ_S^e for any pair of priors (ρ_0, q_0). Under strategic interactions, it turns out that $\rho_I^\star$ could be either smaller or larger than $\rho_S^\star$. We can distinguish a mass market from a niche market by comparing these two cutoffs.

Definition 2 *(Niche market and mass market)*

1. *The market is niche if the cutoffs determined by (ρ_0, q_0) satisfy: $\rho_S^\star \leq \rho_I^\star$, and*

2. *The market is mass if the cutoffs determined by (ρ_0, q_0) satisfy: $\rho_S^\star > \rho_I^\star$.*

In a mass market, the arrival of good news never terminates experimentation while in a niche market, experimentation is shut down by the arrival of the first lump-sum payoff at $\rho \leq \rho_I^\star$. Obviously, whether a mass or niche market appears in equilibrium depends on the priors, which in turn determines the relative importance of social learning and individual learning. We expect that experimentation would continue after the first arrival of lump-sum payoffs if the individual learning component is quite important and vice versa.

3.2.2 Equilibrium in the Individual Learning Phase

A backward procedure is used to characterize $\rho_I^\star$ and $\rho_S^\star$. In the individual learning phase, the equilibrium cutoff $\rho_I^\star$ and the various value functions are provided by the following proposition.

Proposition 2 *Fix a symmetric Markov perfect equilibrium. In the history such that the common uncertainty is resolved, the unknown buyer purchases the risky product if and only if the posterior belief ρ is larger than*

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

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

$$V_I(\rho) = \max \left\{ s, s + g(1-\rho)(1 - [\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}]^{r/\lambda_H}) \right\}; \tag{13}$$

and the monopolist receives value

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

Proof. In the appendix. ■

It is straightforward to see that the equilibrium cutoff $\rho_I^\star$ is strictly larger than the efficient cutoff ρ_I^e. This is because *ex post* heterogeneity means the known buyer is willing to pay more than the unknown buyer. In the absence of price discrimination, the monopolist faces a tradeoff between exploitation of the known buyers and exploration for a higher future value. The incentive to charge a high price and extract the full surplus from the known buyer causes an early termination of experimentation. Another remark is that the unknown buyer is making a myopic choice in the individual learning phase since there is no learning value attached to the purchasing behavior (the unknown buyer always receives value s regardless of whether she receives the lump-sum payoffs).

3.2.3 Equilibrium in the Social Learning Phase

Now consider the situation where 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. 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., Dixit (1993)) to pin down the equilibrium cutoff.

To keep both unknown buyers experimenting, the unknown buyers' value should be such that i) each buyer has an incentive to participate (i.e., the value is larger than the outside option s); ii) each buyer should not benefit from the following deviations: stopping experimentation for a very small amount of time and then switching back to the specified equilibrium behavior.

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

$$U_S(\rho) \geq \hat{U}(\rho; h) = \int_{t=0}^{h} re^{-rt} s dt + \rho q(1-e^{-\lambda_H h})e^{-rh}U_I(\rho) + [1-\rho q(1-e^{-\lambda_H h})]e^{-rh}U^D(\rho, \rho_h) \quad (14)$$

where $\hat{U}(\rho; h)$ denotes the value for a deviator who deviates for h length of time. The deviator receives a deterministic payoff s within the h length of time. After the deviation, with probability $\rho q(1 - e^{-\lambda_H h})$, 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 lump-sum payoffs 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 (14) 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 result is given by lemma 1 in the appendix. Here we just provide a sketch of the proof.

Sketch of the proof for lemma 1. The main difficulty of the proof is to evaluate the off-equilibrium-path value function $U^D(\rho, \rho_h)$. 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 h sufficiently small, 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:

$$\begin{aligned} U^D(\rho, \rho_h) = \mathbb{E} \int_{t=0}^{h'} & re^{-rt}(\rho_t q_t g - \tilde{P}_t)dt \\ & + \rho \tilde{q}_h (1 - e^{-\lambda_H h'})e^{-rh'} V_I(\rho_{h+h'}) + \rho_h \tilde{q}_h (1 - e^{-\lambda_H h'})e^{-rh'} s \\ & + [1 - \rho \tilde{q}_h (1 - e^{-\lambda_H h'}) - \rho_h \tilde{q}_h (1 - e^{-\lambda_H h'})]e^{-rh'} U(\rho_{h'}, \rho_{h+h'}). \end{aligned} \quad (15)$$

In the above expression, ρ_t is the posterior about the deviator and starts from $\rho_0 = \rho$; $\tilde{q}_h$ is the posterior about the product characteristic after an h-deviation such that: $\tilde{q}_h = \frac{q_0(1-\rho_0)^2}{q_0(1-\rho_0)^2 + (1-q_0)(1-\rho)(1-\rho_h)}$; and $\tilde{P}_t$ is the off-equilibrium-path price set by the monopolist after an h-deviation. Obviously, the value function $U^D(\rho, \rho_h)$ depends on the off-equilibrium-path price and cannot be evaluated directly.

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

$$\begin{aligned} U^{ND}(\rho, \rho_h) = \mathbb{E} \int_{t=0}^{h'} & re^{-rt}(\rho'_t q_t g - \tilde{P}_t)dt \\ & + \rho \tilde{q}_h (1 - e^{-\lambda_H h'})e^{-rh'} s + \rho_h \tilde{q}_h (1 - e^{-\lambda_H h'})e^{-rh'} V_I(\rho_{h'}) \\ & + [1 - \rho \tilde{q}_h (1 - e^{-\lambda_H h'}) - \rho_h \tilde{q}_h (1 - e^{-\lambda_H h'})]e^{-rh'} U(\rho_{h+h'}, \rho_{h'}), \end{aligned} \quad (16)$$

where ρ'_t is the posterior about the non-deviator and starts from $\rho'_0 = \rho_h$.

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-path 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-path 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-path price. If the non-deviator has not received lump-sum payoffs 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 (15) and (16) 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 (16) from equation (15) 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 $\rho_S^\star$.

After $U^D(\rho, \rho_h)$ is evaluated, $\lim_{h\to 0} \frac{U_S(\rho) - \hat{U}(\rho;h)}{h}$ can be computed directly. ■

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 gives an important characterization of the on-equilibrium-path value function U_S:

Proposition 3 *Fix the monopolist's strategy such that $\rho_S^\star$ is the equilibrium cutoff in the social learning phase. In a mass market, given any $\rho > \rho_S^\star$, a necessary and sufficient condition for the* *unknown buyers to keep experimenting is that the value $U_S(\rho)$ satisfies differential equation*

$$0 = 2(r+\lambda_H\rho q)(U_S(\rho)-s)+\lambda_H\rho(1-\rho)U_S'(\rho)+(r+\lambda_H\rho)g(1-\rho)q(\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)})^{r/\lambda_H}-\lambda_H g\rho(1-\rho)q$$
$$-\left[\frac{r+\lambda_H\rho_S^\star}{1-\rho_S^\star}(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H}-\lambda_H(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H}\right]g(1-\rho)^2q(\frac{1-\rho}{\rho})^{r/\lambda_H}. \quad (17)$$

In a niche market, given any $\rho > \rho_S^\star$, a necessary and sufficient condition for the unknown buyers to keep experimenting is that the value $U_S(\rho)$ satisfies differential equation

$$0 = 2(r+\lambda_H\rho q)(U_S(\rho)-s)+\lambda_H\rho(1-\rho)U_S'(\rho)$$
$$+\frac{r\lambda_H g}{r+\lambda_H}\frac{(1-\rho)^2q\rho_S^\star}{1-\rho_S^\star}(\frac{(1-\rho)\rho_S^\star}{\rho(1-\rho_S^\star)})^{r/\lambda_H}-\frac{rg}{r+\lambda_H}\lambda_H\rho(1-\rho)q \quad (18)$$

for $\rho \le \rho_I^\star$; and differential equation

$$0 = 2(r+\lambda_H\rho q)(U_S(\rho)-s)+\lambda_H\rho(1-\rho)U_S'(\rho)+(r+\lambda_H\rho)g(1-\rho)q(\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)})^{r/\lambda_H}-\lambda_H g\rho(1-\rho)q$$
$$-r\left[\frac{r+\lambda_H+\lambda_H\rho_I^\star}{(r+\lambda_H)(1-\rho_I^\star)}(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H}-\frac{\lambda_H}{r+\lambda_H}(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H}\right]g(1-\rho)^2q(\frac{1-\rho}{\rho})^{r/\lambda_H} \quad (19)$$

for $\rho > \rho_I^\star$.

The necessity of proposition 3 just comes from combining lemma 1 and lemma 2. In the appendix, we prove the sufficiency of this result as well: given the on-equilibrium-path value function $U_S(\rho)$ and off-equilibrium-path value function $U^D(\rho,\rho_h)$, it is not optimal for an experimenting buyer to deviate.

The ordinary differential equations in proposition 3 can be solved by using observation 1 in the appendix. In a mass market, for any $\rho > \rho_S^\star$, the value function $U_S(\rho)$ is given by

$$\begin{aligned} U_S(\rho) &= s+\frac{\lambda_H}{2r+\lambda_H}g\rho(1-\rho)q-g(1-\rho)q[\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}]^{r/\lambda_H} \\ &+ \left[\frac{r+\lambda_H\rho_S^\star}{r(1-\rho_S^\star)}(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H}-\frac{\lambda_H}{r}(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H}\right]g(1-\rho)^2q(\frac{1-\rho}{\rho})^{r/\lambda_H} \\ &+ C(1-\rho)^2q(\frac{1-\rho}{\rho})^{2r/\lambda_H}. \end{aligned} \quad (20)$$

In a niche market, for any $\rho_S^\star < \rho \le \rho_I^\star$, the value function $U_S(\rho)$ is given by

$$\begin{aligned} U_S(\rho) &= s+\frac{r\lambda_H}{(2r+\lambda_H)(r+\lambda_H)}g\rho(1-\rho)q-\frac{\lambda_H g}{r+\lambda_H}\frac{\rho_S^\star(1-\rho)^2q}{1-\rho_S^\star}(\frac{(1-\rho)\rho_S^\star}{\rho(1-\rho_S^\star)})^{r/\lambda_H} \\ &+ D(1-\rho)^2q(\frac{1-\rho}{\rho})^{2r/\lambda_H}; \end{aligned} \quad (21)$$

and for $\rho > \rho_I^\star$, the value function $U_S(\rho)$ is given by[9]

$$
\begin{aligned}
U_S(\rho) &= s + \frac{\lambda_H}{2r+\lambda_H} g\rho(1-\rho)q - g(1-\rho)q[\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}]^{r/\lambda_H} \\
&+ \left[\frac{r+\lambda_H+\lambda_H\rho_I^\star}{(r+\lambda_H)(1-\rho_I^\star)}(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H} - \frac{\lambda_H}{r+\lambda_H}(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H}\right] g(1-\rho)^2 q(\frac{1-\rho}{\rho})^{r/\lambda_H} \\
&+ (D - \frac{2\lambda_H g}{2r+\lambda_H}(\frac{\rho_I^\star}{1-\rho_I^\star})^{1+2r/\lambda_H})(1-\rho)^2 q(\frac{1-\rho}{\rho})^{2r/\lambda_H}. \qquad (22)
\end{aligned}
$$

Since there is learning value attached to purchasing behavior, the unknown buyer is not making a myopic choice. The monopolist has to provide extra subsidy to deter deviations because the deviator gains rents by becoming more optimistic: $U_S(\rho) > s$.

Denote the equilibrium price in the social learning phase to be $P_S(\rho)$. Then, the value for a buyer from purchasing the risky product can be characterized by the following HJB equation:

$$rU_S(\rho) = r(\rho q(\rho)g - P_S(\rho)) + \lambda_H\rho q(\rho)(U_I(\rho) - U_S(\rho)) + \lambda_H\rho q(\rho)(V_I(\rho) - U_S(\rho)) - \lambda_H\rho(1-\rho)U_S'(\rho) \qquad (23)$$

where $q(\rho) = \frac{q_0(1-\rho_0)^2}{q_0(1-\rho_0)^2 + (1-q_0)(1-\rho)^2}$, $U_I(\rho) = s$, and $V_I(\rho)$ is given by equation (13).

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

$$rJ_S(\rho) = 2rP_S(\rho) + 2\lambda_H\rho q(\rho)(J_I(\rho) - J_S(\rho)) - \lambda_H\rho(1-\rho)J_S'(\rho). \qquad (24)$$

where $J_I(\rho)$ is given by proposition 2.

Equations (23) and (24) are value functions if both unknown buyers purchase the risky product. The RHS of equation (23) can be decomposed into four elements: i) the expected payoff rate from purchasing the risky product $r(\rho q(\rho)g - P_S(\rho))$; 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 (24) could be interpreted similarly.

The on-equilibrium-path price $P_S(\rho)$ can be derived from the on-equilibrium-path value function $U_S(\rho)$. It is straightforward to show: in a mass market,

$$P_S(\rho) = \rho q(\rho)g - s + \frac{\lambda_H}{2r+\lambda_H} g\rho(1-\rho)q(\rho) + Cq(\rho)(1-\rho)^2(\frac{1-\rho}{\rho})^{2r/\lambda_H} \qquad (25)$$

for $\rho > \rho_S^\star$; while in a niche market,

$$P_S(\rho) = \rho q(\rho)g - s - \frac{\lambda_H}{2r+\lambda_H} g\rho(1-\rho)q(\rho) + Dq(\rho)(1-\rho)^2(\frac{1-\rho}{\rho})^{2r/\lambda_H} \qquad (26)$$

[9]The undetermined coefficient in the differential equation is chosen such that $U_S(\rho)$ is continuous at $\rho_I^\star$.

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

$$P_S(\rho) = \rho q(\rho) g - s + \frac{\lambda_H}{2r+\lambda_H} g\rho(1-\rho)q(\rho) + (D - \frac{2\lambda_H g}{2r+\lambda_H}(\frac{\rho_I^\star}{1-\rho_I^\star})^{1+2r/\lambda_H}) q(\rho)(1-\rho)^2(\frac{1-\rho}{\rho})^{2r/\lambda_H} \tag{27}$$

for $\rho > \rho_I^\star$. In the above equations, C and D are constants in equations (20) to (22). Notice in equations (26) and (27), the signs in front of term $\frac{\lambda_H}{2r+\lambda_H} g\rho(1-\rho)q(\rho)$ are different. This reflects the change in continuation value when ρ drops below $\rho_I^\star$. By proposition 2, for $\rho \leq \rho_I^\star$, 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, it is easy to check that in a niche market, the price $P_S(\rho)$ is still continuous at $\rho_I^\star$.

We substitute the price expression $P_S(\rho)$ into equation (24) 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 4 *(Characterize the symmetric Markov perfect equilibrium) In the social learning phase, the unknown buyers purchase the risky product under posterior beliefs (ρ, q) if and only if*

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

A mass market appears if and only if

$$\frac{1-q_0}{q_0(1-\rho_0)^2} > \frac{g}{(1-\rho_I^\star)s}. \tag{28}$$

Moreover, for all $\rho_0 < 1$ and $q_0 < 1$, the symmetric Markov perfect equilibrium is inefficient so that experimentation is terminated too early.

Proof. In the appendix. ■

The unique equilibrium cutoff $\rho_S^\star$ is characterized by equation

$$\rho q(\rho) = \frac{rs}{rg + \lambda_H(V_I(\rho) + J_I(\rho)) - \lambda_H s}. \tag{29}$$

It is straightforward to show the equilibrium is inefficient by comparing the efficient stopping curve with the equilibrium stopping curve. The inefficiency in the individual learning phase causes a leakage of the social surplus for the monopolist, which reduces the monopolist's incentives to subsidize experimentation in the social learning phase. Therefore, the equilibrium experimentation is terminated too early in the social learning phase as well.

There are two remarks about proposition 4. First, 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

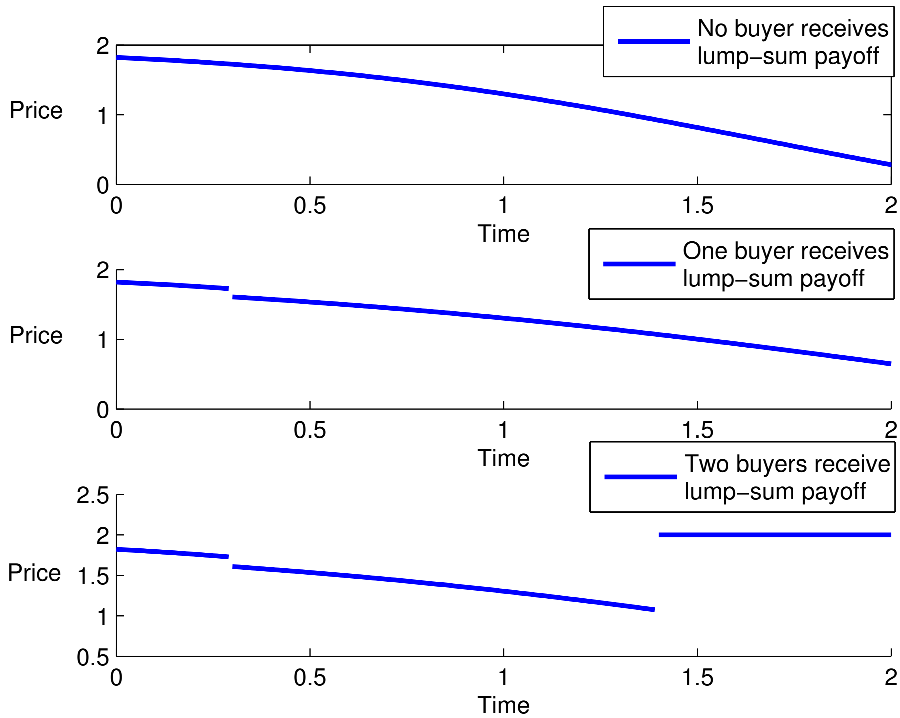

Figure 2: Equilibrium Price Dynamics

solving an optimal stopping problem given the price 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. Second, the appearance of a mass market depends on the relative importance of social learning and individual learning. Given q_0, when ρ_0 goes up, the monopolist has higher incentives to keep the remaining unknown buyer experimenting. A mass market is more likely to appear as a result.

3.2.4 Equilibrium Price Path

After solving for the equilibrium cutoff $\rho_S^\star$, the constants C and D in equations (20) and (21) can be pinned down from the value matching condition and then the expression for the equilibrium prices can be derived. Figure 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.

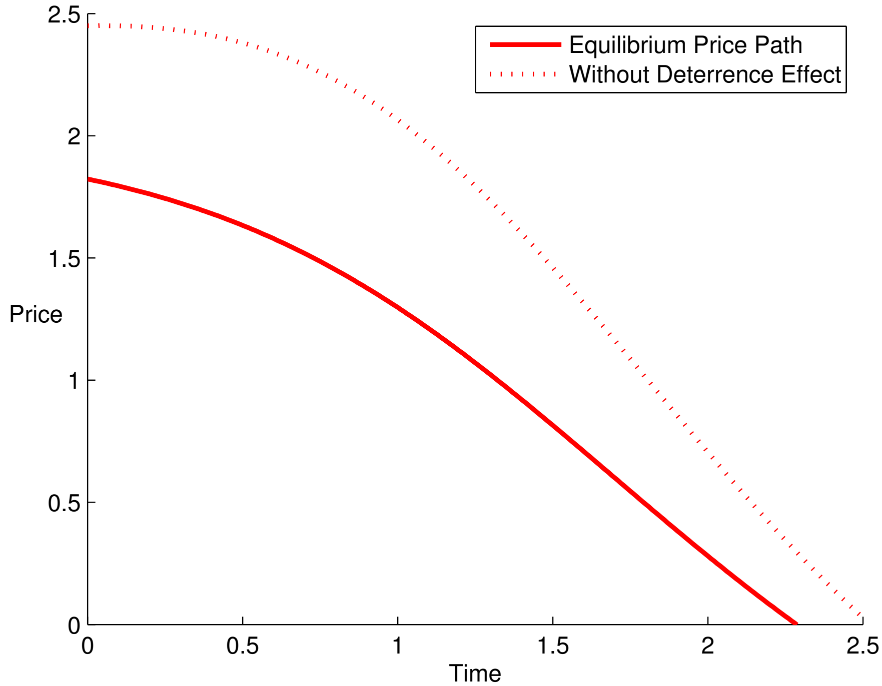

Figure 3: Deterrence Effect

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

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

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 price $P_S(\rho)$ is strictly less than $\tilde{P}_S(\rho)$ because of this deterrence effect. Figure 3 compares the equilibrium price path with and without the deterrence effect. It shows that the price reduction caused by the deterrence effect is quite significant.

Second, the instantaneous price reaction to the arrival of the first lump-sum payoff might be ambiguous. In particular, when the first lump-sum payoff arrives, there could be an instantaneous drop in price in order to encourage the buyer who remains unsure to experiment as shown by figure 2. To understand the negative response of the price to the arrival of a good news signal, we first compare the equilibrium price in the individual learning phase $P_I(\rho)$ and the price without the deterrence effect $\tilde{P}_S(\rho)$. Equation

$$P_I(\rho) - \tilde{P}_S(\rho) = \rho(1 - q(\rho))g - \frac{\lambda_H}{r} \rho q(\rho)(V_I(\rho) - s)$$

shows that the arrival of good news brings two opposite effects on the reservation value of the buyer who remains unsure. There is a positive informational effect captured by $\rho(1-q(\rho))g$: the arrival of good news reveals that the product characteristic is high and hence makes the unknown buyer more optimistic about the unconditional probability of receiving lump-sum payoffs. However, there is another negative continuation value effect: the buyer who remains unsure loses the chance of becoming the first known buyer to extract rents. The price has to be lower to compensate for the loss of rents if the monopolist wishes to make a sale to the unknown buyer.

The comparison of the informational effect and the continuation value effect depends on the comparison of $1-q(\rho)$ and $q(\rho)(V_I(\rho)-s)$.

Corollary 1 *For $\rho_0 < 1$ and $q_0 < 1$, $\frac{q(\rho)(V_I(\rho)-s)}{1-q(\rho)}$ is strictly increasing in ρ.*

Proof. Plug the formula of $q(\rho)$ and $V_I(\rho)$ into $\frac{q(\rho)(V_I(\rho)-s)}{1-q(\rho)}$ and we can get $\frac{q(\rho)(V_I(\rho)-s)}{1-q(\rho)}$ is proportional to

$$\frac{1-[\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}]^{r/\lambda_H}}{1-\rho},$$

which is strictly increasing in ρ. ■

The above corollary implies: in the early days of the market, ρ is higher and it is more likely to have $\tilde{P}_S(\rho) > P_I(\rho)$; in the late days of the market, ρ is lower and it is more likely to have $\tilde{P}_S(\rho) < P_I(\rho)$. Since the equilibrium price $P_S(\rho)$ is strictly below $\tilde{P}_S(\rho)$ due to the deterrence effect, the above statement also holds if we replace $\tilde{P}_S(\rho)$ with $P_S(\rho)$. Figure 4 describes a situation where with the same priors, the price might either drop or jump depending on the arrival time of the first lump-sum payoff.

3.3 Efficiency

This section discusses the efficiency property of the symmetric Markov perfect equilibrium for an arbitrary number of buyers. We first investigate the extreme case of the perfect payoff correlation ($\rho = 1$) and then compare that result to the one in the partial payoff correlation case.

Perfect Payoff Correlation Under this special case, buyers are *ex post homogeneous*. In other words, immediately after one buyer receives a lump-sum payoff, it becomes common knowledge that all buyers are able to receive lump-sum payoffs, and the monopolist should immediately raise the price to $g-s$ to extract all of the surplus.

In the social learning phase, similarly the monopolist should set a price such that i) each experimenting buyer has an incentive to participate (i.e., each buyer's value is larger than the outside option); ii) it is not optimal for each experimenting buyer to have "one-shot" deviations. The common value assumption simplifies the analysis of the "one-shot deviation" problem since the deviator always has the same posterior belief as the buyers who have not deviated. It turns out

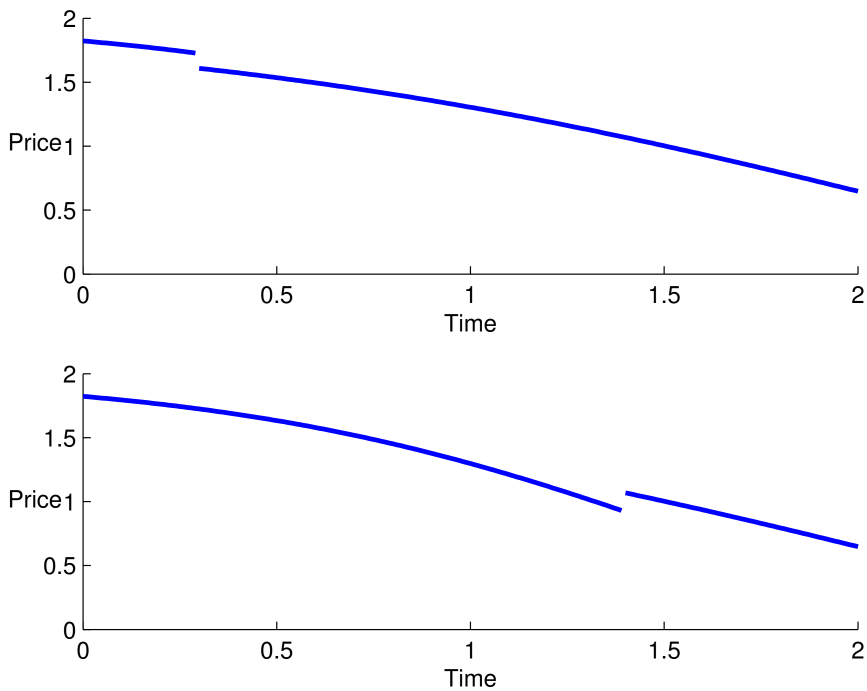

Figure 4: Instantaneous Price Response to the First Arrival of Good News

that under the common value case, restrictions i) and ii) coincide and the strategic equilibrium is always efficient.

Proposition 5 *When the buyers' payoffs are perfectly correlated ($\rho = 1$), the unknown buyers will always receive value s in equilibrium and the symmetric Markov perfect equilibrium is efficient.*

Proof. In the appendix. ■

The intuitive explanation for the above efficiency result is that the ex post homogeneity means the monopolist does not need to face the tradeoff between exploitation and exploration. This enables the monopolis to completely internalize the social surplus and overcome the free riding problem by subsidizing experimentation.

Partial Payoff Correlation Since ex post heterogeneity exists in the partial payoff correlation case, it is natural to conjecture that the inefficiency result in proposition 4 can be extended to a general n case. The induction argument is used to avoid solving for every value function explicitly.

Theorem 1 *Consider a market with any $n \geq 2$ buyers. The symmetric Markov perfect equilibrium is inefficient in both the social learning and individual learning phases if $\rho_0 < 1$ and $q_0 < 1$. Moreover, the equilibrium experimentation is always terminated too early.*

Proof. In the appendix. ■

We are in a position to summarize the roles played by ex post heterogeneity. First, in the social learning phase, ex post heterogeneity means there is a future benefit for the deviator by becoming more optimistic than the non-deviators. The monopolist has to provide extra subsidy to deter deviations. In the common value case, such a future benefit does not exist and there is no need to provide extra subsidy. 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 generates a loss of rents for the buyers who stay unsure upon the arrival of the first lump-sum payoff. The reduction in continuation values leads to an ambiguous instantaneous price reaction to the arrival of the first lump-sum payoff. On the contrary, in the common value case, the equilibrium value for the buyers is always the same as the outside option and there is no continuation value effect. Hence, upon the arrival of the first lump-sum payoff, the instantaneous reaction of the equilibrium price is always to go up. 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. On the other hand, in the common value case, there is no ex post heterogeneity and the monopolist is able to fully internalize the social surplus.

4 Equilibrium in the Bad News Case

In the bad news case, the arrival of lump-sum payoffs (we call them 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 the safe one only when the buyers cannot receive lump-sum damages.

4.1 Socially Efficient Allocation

Different from the good news case, large priors (ρ_0, q_0) mean that the probability of receiving lump-sum damages is high and this discourages the social planner from taking 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 separately in the individual learning and social learning phases.

Socially Efficient Allocation in the Individual Learning Phase In the individual learning phase, suppose k buyers have received lump-sum damages. The social surplus function could be written as (the known buyers will take the safe product and receive s for sure)

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

where

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

defines the optimal control problem for the unknown buyer. The corresponding HJB equation is

$$W(\rho) = \max\left\{s, A-\rho B + \frac{1}{r}\left[\lambda_H\rho(s-W(\rho)) - \lambda_H\rho(1-\rho)W'(\rho)\right]\right\}. \tag{30}$$

Solve the optimal starting problem defined by equation (30) and we get the following result:

Proposition 6 *In the individual learning phase, if $k \geq 1$ buyers are known to receive lump-sum damages, it is socially efficient for those k 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*

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

The value functions for a typical buyer with posterior belief ρ is given by:

$$W(\rho) = \max\left\{s, A - \frac{\lambda_H A + rB - \lambda_H s}{r+\lambda_H}\rho\right\}.$$

Socially Efficient Allocation in the Social Learning Phase In the social learning phase, we similarly write down the HJB equation as:

$$\Omega_S(\rho) = \max\left\{ns, n(A-\rho q(\rho)B) + \frac{1}{r}[\lambda_H n\rho q(\rho)(\Omega_1(\rho) - \Omega_S(\rho)) - \lambda_H\rho(1-\rho)\Omega_S'(\rho)]\right\}. \tag{31}$$

The optimal starting problem (31) is solved by solving differential equation

$$(r+\lambda_H n\rho q)\Omega_S(\rho) = rn(A-\rho qB) + \lambda_H n\rho q[(n-1)W(\rho)+s] - \lambda_H\rho(1-\rho)\Omega_S'(\rho), \tag{32}$$

with boundary condition $\Omega_S(\rho_S^e) = ns$.[10]

The socially efficient allocation in the social learning phase is characterized by the following proposition:

[10] Notice that $W(\rho)$ is not continuously differentiable at ρ_I^e (smoothing pasting condition is no longer satisfied). But it is Lipschitz continuous and hence the solution to the above boundary value problem is still unique.

Proposition 7 *Given any $q_0 < 1$, there exists a unique $\rho_S^e(q_0) > \rho_I^e$ ($\rho_S^e(q_0)$ could be one) such that it is socially efficient to start experimentation in the social learning phase if and only if $\rho \leq \rho_S^e(q_0)$.*

Proof. In the appendix. ■

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 where 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:

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

  where τ is the first (possibly infinite) time at which a new unknown buyer receives good news;
- given (α_k^0, α_k^1), the monopolist chooses price $P_k(\rho_t)$ to maximize

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

- beliefs update according to Bayes' rule: ρ_t satisfies the law of motion, i.e., equation (1); $q_k(\rho_t) = 1$ for $k \geq 1$ and $q_k(\rho_t)$ is given by equation (6) for $k = 0$;
- for $k = n$, the monopolist will not serve any buyer such that $J_n = 0$ and $V_n = s$.

First, it is straightforward to see that the known buyers will buy the risky product if the price is lower than $A - B - s$ and not buy otherwise. Second, the assumption $A - B - s < 0$ implies that selling to the known buyers is purely losing money. Hence, a profit-maximizing monopolist should never set the price lower than $A - B - s$ in order to sell to the known buyers. This also implies that V_k is always s. Third, 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, the cutoff strategy for the monopolist means that she will *start* selling to the unknown buyers if the belief ρ is lower than a certain cutoff. 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 monopolist also has to charge a price such that both the participation constraint and the no profitable one-shot deviation constraint are satisfied. In the bad news case, it turns out that the "one-shot" deviations don't impose more restrictions than the participation constraint.

Claim 1 *In equilibrium, the most pessimistic unknown buyer's value is always s.*

Claim 1 implies that the on-equilibrium-path value for each unknown buyer is always s since they are equally pessimistic. This is different from proposition 3 in the good news case. In the good news case, a one-shot deviation makes the non-deviators more pessimistic if they haven't received any lump-sum payoffs during the deviation period. In that situation, the price charged by the monopolist is lower than what the deviator is willing to pay. The deviator can benefit from a deviation and thus the equilibrium value for the experimenting buyers has to be larger than s to deter deviations. However, in the bad news case, a one-shot deviation makes the deviator more pessimistic. After the deviation, if the monopolist wishes to serve all unknown buyers, the optimal price is determined by what 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 on-equilibrium-path value to be s is enough to deter deviations.

The equilibrium price path could be derived from claim 1: in the individual learning phase, the monopolist would charge $P_I(\rho) = A - \rho B - s$ and in the social learning phase, the monopolist would charge $P_S(\rho) = A - \rho q(\rho) B - s$. The arrival of the first lump-sum damage will unanimously lead to a drop in price if $q_0 < 1$ but the subsequent arrival of lump-sum damages will not have any impact on price. The negative response in price to the arrival of the first lump-sum damage reflects the fact that there is no continuation value effect from claim 1. The informational effect always discourages the unknown buyers from experimenting and reduces the price. But the subsequent arrival of bad news reveals no more information to the remaining unknown buyers and hence has no effect on the price at all. Solve the monopolist's optimal starting problem and we get the following theorem:

Theorem 2 *Consider a market with $n \geq 2$ buyers. The symmetric Markov perfect equilibrium is efficient in both the social learning and the individual learning phases.*

Proof. In the appendix. ■

The above theorem is very intuitive: different from the good news model, there is no tradeoff between exploitation and exploration in the individual learning phase because the buyers who have received lump-sum damages will never purchase the risky product. 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 in the individual learning phase. The efficiency in the social learning phase is a little surprising. It seems that the monopolist cannot fully internalize social surplus since the unknown buyers can benefit from social learning by switching to the safe product. The intuition turns out to be incorrect. In the good news case, society benefits from the arrival of good news but the receivers of the lump-sum payoffs pay less than what they are willing to pay. In other words, the known buyers "steal" some of the social surplus from the monopolist and this causes inefficiency. On the contrary, in the bad news case, society benefits from the *non-arrival* of the bad news. The unknown buyers cannot "steal" social surplus from the monopolist when no lump-sum damages have been received.

5 Conclusion

By combining common and idiosyncratic uncertainty, this paper relaxes the usual common value assumption made in the social learning literature (see, e.g., Banerjee (1992), Bikhchandani, Hirshleifer, and Welch (1992) and Rosenberg, Solan, and Vieille (2007)).[11] We consider a dynamic monopoly pricing environment where the monopolist cannot price-discriminate among the buyers. The partial payoff correlation among the buyers generates ex post heterogeneity. If the monopolist wishes to make a sale 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 generates a reduction in the continuation value for the buyers who stay unsure, 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 level of experimentation. On the contrary, if there is a perfect payoff correlation among the buyers, the arrival of the first good news signal always leads to a jump in price and the equilibrium is efficient.

There is another subtle off-equilibrium-path implication. By taking the outside option, each buyer can extract rents 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

[11] An exception is Murto and Välimäki (2009), who consider partial payoff correlation in an observational learning setting.

unknown buyers, each unknown buyer receives a value higher than the outside option to deter deviations. Such a deterrence effect leads to a significant reduction in the equilibrium price. If there is perfect payoff correlation among the buyers, there is no need to provide such an extra subsidy.

However, in the bad news case, the above implications do not exist for two reasons. 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. Off the equilibrium path, a buyer cannot benefit from deviations because the deviator becomes more pessimistic after a deviation.

There are several extensions to consider in the future. For tractability, we have assumed that the arrival of lump-sum payoffs immediately resolves the common uncertainty and the idiosyncratic uncertainty of the receiver. It is possible to consider a model where the arrival of lump-sum payoffs cannot immediately resolve the common uncertainty or the idiosyncratic uncertainty of the receiver. For example, we may assume lump-sum payoffs arrive at another Poisson rate when the product characteristic is low. 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 Bergemann and Välimäki (2002), who consider a model with a continuum of buyers such that buyers are choosing according to their myopic preferences at each instant in time. 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

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

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:

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

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 which are independent of previous prices since allowing pricing as a function of previous prices may generate more complicated problems.[12] 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 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 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 have to be imposed to guarantee the induced outcome is unique. Before doing that, we define an outcome h to be

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

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 Bergin and McLeod (1993) shows the following:

Proposition 8 *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 8. In proposition 8, condition 2 is the key condition. This condition is slightly different from the inertia condition proposed in Bergin and McLeod (1993). The modification is needed to handle the possible situation when the arrival of a lump-sum payoff at time t results in the purchasing decisions a_t to be not right continuous in time.

Any Markovian strategy profile (P, α) which induces a weakly admissible strategy profile generates a unique distribution over compatible outcomes. But the notion of weak admissibility does not guarantee that the induced outcome allows us to use equations (1) and (2) to update beliefs.

Definition 3 *A Markovian strategy profile (P, α) is* strongly admissible *in the good news case if it satisfies:*[13]

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 $||\boldsymbol{\rho}' - \boldsymbol{\rho}|| \leq \epsilon$;*[14]

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 3 are sufficient to guarantee that (P, α) induces a weakly admissible strategy profile. More than that, these two conditions imply 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 equations (1) and (2) to update beliefs. In the good news case, condition 1 implies P_t is decreasing when it is continuous but it also allows 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

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

[14]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.

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

Definition 3 is too strong in the sense that even cutoff strategies may not be strongly admissible.[15] We use the completion argument in Bergin and McLeod (1993) 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. But the proof of theorem 2 in Bergin and McLeod (1993) 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 to the consistent outcomes.

In the 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 is well behaved even off the equilibrium path. The conditions imposed are counterparts of conditions 1-3 in definition 3.

Definition 4 *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 $P(h, t+\epsilon) \le P(h, t)$ for all $\epsilon \le \bar{\epsilon}_1$;

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 to the consistent outcomes which can be similarly defined.

B Proofs of Results from Section 3

B.0 General Solution to Linear First Order Ordinary Differential Equations

The following observation is widely used throughout the paper to solve linear first order ordinary differential equations.

[15]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.

Observation 1 *Given that f and g are continuous functions on an interval I, the ordinary differential equation $y' + f(x)y = g(x)$ has a general solution*

$$y(x) = \frac{H(x)}{h(x)}$$

where $h(x) = e^{R(x)}$, $R(x)$ is an antiderivative of $f(x)$ on I and $H(x)$ is an antiderivative of $h(x)g(x)$ on I.[16]

Proof. Multiply both sides of differential equation $y' + f(x)y = g(x)$ by $h(x)$. Then the original differential equation becomes

$$\frac{d}{dx}(h(x)y(x)) = h(x)g(x).$$

After integration, it is straightforward to see that the general solution is $y(x) = \frac{H(x)}{h(x)}$. ■

B.1 Proof of Proposition 1

Proof. Before proving the proposition, we first show the socially optimal allocation is indeed symmetric.

Claim 2 *The socially optimal allocation is symmetric when buyers are homogeneous.*

Proof. For any posteriors $\boldsymbol{\rho}$, denote the social surplus to be $\Omega(\boldsymbol{\rho})$. The social planner's problem can be written as:

$$\Omega(\boldsymbol{\rho}) = \sup_{\boldsymbol{\alpha}(\cdot)\in\{0,1\}^n} \mathbb{E}\left\{ \int_{t=0}^{h} re^{-rt} \sum_{i=1}^{n} [\alpha_i(\boldsymbol{\rho}_t)\rho_{it}q(\boldsymbol{\rho}_t)g + (1-\alpha_i(\rho_t))s]dt + e^{-rh}\Omega(\boldsymbol{\rho}_h \mid \alpha) \right\}.$$

Consider any $\tilde{\boldsymbol{\rho}}$ which is a permutation of $\boldsymbol{\rho}$. Naturally, the social surplus should be the same: $\Omega(\boldsymbol{\rho}) = \Omega(\tilde{\boldsymbol{\rho}})$ since the strategies $\boldsymbol{\alpha}$ can be permuted as well. Suppose buyers are homogeneous with the same prior ρ_0 and denote $\boldsymbol{\rho_0} = (\rho_0, \cdots, \rho_0)$. From the HJB equation, it is socially optimal for buyer i to purchase the risky product if and only if:

$$r\rho_0 q_0 g + \rho_0 q_0 \lambda_H(\Omega_1(\rho_0) - \Omega(\boldsymbol{\rho}_0)) - \lambda_H \rho(1-\rho)\frac{\partial\Omega(\boldsymbol{\rho}_0)}{\partial\rho_i} > rs.$$

Since $\Omega(\boldsymbol{\rho}) = \Omega(\tilde{\boldsymbol{\rho}})$, for any $j \neq i$, we can switch i and j without affecting the partial derivatives. In other words, the partial derivatives are identical when buyers are homogeneous: $\frac{\partial\Omega(\boldsymbol{\rho}_0)}{\partial\rho_i} = \frac{\partial\Omega(\boldsymbol{\rho}_0)}{\partial\rho_j}$. Therefore, it is socially optimal for buyer i to purchase the risky product if and only if it is also optimal for buyer j to purchase. This implies the socially optimal allocation is symmetric. ■

Notice in equation

$$rn\rho q(\rho)g + n\rho q(\rho)\lambda_H\Omega_1(\rho) = (r + n\rho q(\rho)\lambda_H)ns, \tag{33}$$

$\Omega_1(\cdot)$ is a piece-wise function since $W(\cdot)$ is a piece-wise function. The next result claims that ρ_S^e is always larger than ρ_I^e such that $\Omega_1(\rho_S^e) > (n-1)s + g$.

Claim 3 *Beginning from any combination of $\rho_0 < 1$ and $q_0 < 1$, the efficient cutoff in the social learning phase will always be larger than the efficient cutoff in the individual learning phase: $\rho_S^e > \rho_I^e$.*

[16] An antiderivative of a function $f(x)$ is defined as any function $F(x)$ whose derivative is $f(x)$: $F'(x) = f(x)$.

Proof. We first substitute the expression $\Omega_1(\rho) = g + (n-1)W(\rho)$ into equation (33) and get

$$rn\rho q(\rho)g + n\rho q(\rho)\lambda_H[g + (n-1)W(\rho)] = (r + n\rho q(\rho)\lambda_H)ns. \tag{34}$$

By contradiction, assume $\rho_S^e \le \rho_I^e$ and $W(\rho_S^e) = s$ by definition. Equation (34) then gives us a cutoff $\tilde{\rho}_S^e$ satisfying

$$\tilde{\rho}_S^e q(\tilde{\rho}_S^e) = \rho_I^e = \frac{rs}{(r+\lambda_H)g - \lambda_H s}.$$

As $q(\tilde{\rho}_S^e) < 1$, the above equation implies that: $\tilde{\rho}_S^e > \rho_I^e$, which contradicts the assumption $\rho_S^e \le \rho_I^e$. Therefore, it must be true that $\rho_S^e > \rho_I^e$ and thus $W(\rho_S^e) > s$. ■

From claim 3, ρ_S^e should satisfy equation (34) where $q(\rho_S^e)$ is given by equation (6). Given the priors, the efficient cutoffs $(\rho_S^e(\rho_0, q_0), q_S^e(\rho_0, q_0))$ can be solved jointly:

$$q_S^e = \frac{rs}{\rho_S^e[(r+\lambda_H)g + (n-1)\lambda_H W(\rho_S^e) - n\lambda_H s]}. \tag{35}$$

$$q_S^e = \frac{(1-\rho_0)^n q_0}{(1-\rho_0)^n q_0 + (1-\rho_S^e)^n (1-q_0)}. \tag{36}$$

Clearly, $W(\rho_S^e)$ is increasing in ρ_S^e and thus q_S^e is decreasing in ρ_S^e from equation (35). Equation (36) describes how ρ and q evolve jointly over time: since both ρ and q decrease over time, q_S^e is increasing in ρ_S^e. Hence the intersection of equations (36) and (35) is unique. Equation (35) describes the stopping curve such that it is socially efficient to keep experimenting if

$$\rho q > \frac{rs}{(r+\lambda_H)g + (n-1)\lambda_H W(\rho_S^e) - n\lambda_H s}.$$

Finally, we still have to check that it is indeed the case that $\rho_S^e > \rho_I^e$. Notice that ρ_S^e is decreasing in q_S^e on the stopping curve. If $q = 1$, it is easy to check the unique cutoff ρ_S^e is the same as $\rho_I^e = \frac{rs}{(r+\lambda_H)g - \lambda_H s}$. And for $q_S^e < 1$, we should have $\rho_S^e > \rho_I^e$. ■

B.2 Proof of Proposition 2

Proof. In the individual learning phase, 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 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 $g\rho - s$ 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

$$rJ_I(\rho) = \max\left\{r(g-s), 2r(g\rho - s) + \rho\lambda_H(2(g-s) - J_I(\rho)) - \lambda_H\rho(1-\rho)J_I'(\rho)\right\}. \tag{37}$$

On the RHS of equation (37), $g - s$ 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 also may 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

$$rV_I(\rho) = r(g(1-\rho) + s) + \rho\lambda_H(s - V_I(\rho)) - \lambda_H\rho(1-\rho)V_I'(\rho) \tag{38}$$

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 (38) 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.3 Characterize $\lim_{h\to 0} \frac{U_S(\rho) - \hat{U}(\rho;h)}{h}$

Lemma 1 *Fix a pair of priors (ρ_0, q_0) such that $\rho_S^\star$ is the equilibrium cutoff in the social learning phase. In a mass market, for any $\rho > \rho_S^\star$,*

$$\begin{aligned} \lim_{h\to 0} \frac{U_S(\rho) - \hat{U}(\rho;h)}{h} &= 2(r + \lambda_H\rho q)(U_S(\rho) - s) + \lambda_H\rho(1-\rho)U_S'(\rho) \\ &\quad + (r + \lambda_H\rho)g(1-\rho)q(\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)})^{r/\lambda_H} - \lambda_H g\rho(1-\rho)q \\ &\quad - \left[\frac{r + \lambda_H\rho_S^\star}{1-\rho_S^\star}(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H} - \lambda_H(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H}\right] g(1-\rho)^2 q(\frac{1-\rho}{\rho})^{r/\lambda_H}. \end{aligned} \tag{39}$$

In a niche market, for $\rho_S^\star < \rho \le \rho_I^\star$,

$$\begin{aligned} \lim_{h\to 0} \frac{U_S(\rho) - \hat{U}(\rho;h)}{h} &= 2(r + \lambda_H\rho q)(U_S(\rho) - s) + \lambda_H\rho(1-\rho)U_S'(\rho) \\ &\quad - \frac{rg}{r+\lambda_H}\lambda_H\rho(1-\rho)q + \frac{r\lambda_H g}{r+\lambda_H}\frac{\rho_S^\star(1-\rho)^2 q}{1-\rho_S^\star}(\frac{(1-\rho)\rho_S^\star}{\rho(1-\rho_S^\star)})^{r/\lambda_H}; \end{aligned} \tag{40}$$

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

$$\begin{aligned} \lim_{h\to 0} \frac{U_S(\rho) - \hat{U}(\rho;h)}{h} &= 2(r + \lambda_H\rho q)(U_S(\rho) - s) + \lambda_H\rho(1-\rho)U_S'(\rho) \\ &\quad + (r + \lambda_H\rho)g(1-\rho)q(\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)})^{r/\lambda_H} - \lambda_H g\rho(1-\rho)q \\ &\quad - r\left[\frac{r + \lambda_H + \lambda_H\rho_I^\star}{(r+\lambda_H)(1-\rho_I^\star)}(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H} - \frac{\lambda_H}{r+\lambda_H}(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H}\right] g(1-\rho)^2 q(\frac{1-\rho}{\rho})^{r/\lambda_H}. \end{aligned} \tag{41}$$

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:

$$\lim_{h\to 0}\frac{U_S(\rho)-\hat{U}(\rho;h)}{h}=(r+\lambda_H\rho q(\rho))(U_S(\rho)-s)+\lim_{h\to 0}\frac{U_S(\rho)-U^D(\rho,\rho_h)}{h}. \tag{42}$$

The main issue is to evaluate $U^D(\rho,\rho_h)$ for $\rho>\rho_h$. We proceed in the following steps:

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

   Fix $h>0$ to be sufficiently small and the monopolist will still sell to both buyers after an h-deviation.[17] Therefore, there exists $\bar{h}'$ such that for all $h'\le\bar{h}'$, we have:

$$\begin{aligned}U^D(\rho,\rho_h)=\mathbb{E}\int_{t=0}^{h'}&re^{-rt}(\rho_t q_t g-\tilde{P}_t)dt\\&+\rho\tilde{q}_h(1-e^{-\lambda_H h'})e^{-rh'}V_I(\rho_{h+h'})+\rho_h\tilde{q}_h(1-e^{-\lambda_H h'})e^{-rh'}s\\&\qquad+[1-\rho\tilde{q}_h(1-e^{-\lambda_H h'})-\rho_h\tilde{q}_h(1-e^{-\lambda_H h'})]e^{-rh'}U(\rho_{h'},\rho_{h+h'}).\end{aligned} \tag{43}$$

   In the above expression, ρ_t is the posterior about the deviator and starts from $\rho_0=\rho$; $\tilde{q}_h$ is the posterior about the product characteristic after an h-deviation: $\tilde{q}_h=\frac{q_0(1-\rho_0)^2}{q_0(1-\rho_0)^2+(1-q_0)(1-\rho)(1-\rho_h)}$; and $\tilde{P}_t$ is the off-equilibrium-path price set by the monopolist after an h-deviation.

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

$$\begin{aligned}U^{ND}(\rho,\rho_h)=\mathbb{E}\int_{t=0}^{h'}&re^{-rt}(\rho'_t q_t g-\tilde{P}_t)dt\\&+\rho\tilde{q}_h(1-e^{-\lambda_H h'})e^{-rh'}s+\rho_h\tilde{q}_h(1-e^{-\lambda_H h'})e^{-rh'}V_I(\rho_{h'})\\&\qquad+[1-\rho\tilde{q}_h(1-e^{-\lambda_H h'})-\rho_h\tilde{q}_h(1-e^{-\lambda_H h'})]e^{-rh'}U(\rho_{h+h'},\rho_{h'}),\end{aligned} \tag{44}$$

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

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

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

   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:

$$\begin{aligned}U^{ND}(\rho,\rho_h)\ge\tilde{U}(h')=\int_{t=0}^{h'}&re^{-rt}sdt+\rho\tilde{q}_h(1-e^{-\lambda_H h'})e^{-rh'}s\\&+[1-\rho\tilde{q}_h(1-e^{-\lambda_H h'})]e^{-rh'}U(\rho_h,\rho_{h'}).\end{aligned} \tag{45}$$

   Since the ρ_h buyer is more pessimistic about the probability of receiving lump-sum payoffs, the optimal off-equilibrium-path price $\tilde{P}$ is set such that the ρ_h buyer has incentives to experiment.

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

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

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

is left-continuous in ρ and ρ_h implies that in equilibrium, the following equation is satisfied:[18]

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

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

$$(r+\lambda_H\rho\tilde{q}_h)\tilde{U}(\rho;\rho_h)=(r+\lambda_H\rho\tilde{q}_h)s-\lambda_H\rho(1-\rho)\tilde{U}'(\rho;\rho_h) \tag{46}$$

where the expression for $\tilde{q}_h$ is provided by equation (5)

$$\tilde{q}_h(\rho)=\frac{q_0(1-\rho_0)^2}{q_0(1-\rho_0)^2+(1-q_0)(1-\rho)(1-\rho_h)}.$$

The off-equilibrium-path 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-path value function $\tilde{U}(\rho;\rho_h)$.**

   Equation (46) is an ordinary differential equation with general solution:

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

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

$$C_h=\frac{U_S(\rho_h)-s}{(1-\rho_h)q_h(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H}}; \tag{47}$$

   where q_h satisfies: $q_h=\frac{q_0(1-\rho_0)^2}{q_0(1-\rho_0)^2+(1-q_0)(1-\rho_h)^2}$.

   Since on the equilibrium path, experimentation stops at $\rho_S^\star$, the unknown buyer receives a value less than the outside ($U_S(\rho)<s$) for $\rho<\rho_S^\star$. Equation (47) implies that the non-deviator's posterior will never be lower than ρ_S^* 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

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

[18] 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.

corresponds to the deviator's posterior when the non-deviator's posterior drops to $\rho_S^\star$.

3. **Solve for the off-equilibrium-path 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 $\rho(t)$ and $\rho_h(t)$ are posterior beliefs after t length of time beginning from ρ and ρ_h (given that no lump-sum payoff is received during this period). The posteriors can be expressed as:

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

and

$$\tilde{q}_h(t) = \frac{q_0(1-\rho_0)^2}{q_0(1-\rho_0)^2 + (1-q_0)(1-\rho(t))(1-\rho_h(t))}.$$

Given any $t < h'$, the monopolist would also make a sale to both buyers $\rho(t)$ and $\rho_h(t)$. Subtract equation (44) from (43) yields:

$$\begin{aligned} Z(t) = \mathbb{E} \int_0^{h''} & re^{-r\tau}(\rho_\tau q_\tau g - \rho'_\tau q_\tau g) d\tau \\ & + e^{-rh''}(1 - e^{-\lambda_H h''}) \left\{ \rho(t)\tilde{q}_h(t)[V_I(\rho_h(t+h'')) - s] + \rho_h(t)\tilde{q}_h(t)[s - V_I(\rho(t+h''))] \right\} \\ & \quad + e^{-rh''} \left[1 - \rho(t)\tilde{q}_h(t)(1 - e^{-\lambda_H h''}) - \rho_h(t)\tilde{q}_h(t)(1 - e^{-\lambda_H h''}) \right] Z(t+h''). \end{aligned} \tag{48}$$

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

$$(r + \lambda_H \rho(t)\tilde{q}_h(t) + \lambda_H \rho_h(t)\tilde{q}_h(t))Z(t) - \dot{Z}(t) = H(t) \tag{49}$$

where

$$H(t) = r(\rho(t) - \rho_h(t))\tilde{q}_h(t)g + \lambda_H \rho(t)\tilde{q}_h(t)(V_I(\rho_h(t)) - s) - \lambda_H \rho_h(t)\tilde{q}_h(t)(V_I(\rho(t)) - s).$$

Next, the explicit expression for Z can be derived for mass and niche markets, respectively. In a *mass market*, both $\rho(t)$ and $\rho_h(t)$ are larger than $\rho_I^\star$. In that case,

$$V_I(\rho) = s + g(1-\rho)(1 - [\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}]^{r/\lambda_H})$$

and

$$\begin{aligned} H(t) = r(\rho(t) - \rho_h(t))\tilde{q}_h(t)g + \lambda_H \rho(t)\tilde{q}_h(t)g(1-\rho_h(t))(1 - [\frac{(1-\rho_h(t))\rho_I^\star}{\rho_h(t)(1-\rho_I^\star)}]^{r/\lambda_H}) \\ - \lambda_H \rho_h(t)\tilde{q}_h(t)g(1-\rho(t))(1 - [\frac{(1-\rho(t))\rho_I^\star}{\rho(t)(1-\rho_I^\star)}]^{r/\lambda_H}). \end{aligned}$$

The solution to differential equation (49) is

$$Z(t) = (\rho(t) - \rho_h(t))\tilde{q}_h(t)g - [(1-\rho_h(t))(\frac{1-\rho_h(t)}{\rho_h(t)})^{r/\lambda_H} - (1-\rho(t))(\frac{1-\rho(t)}{\rho(t)})^{r/\lambda_H}]\tilde{q}_h(t)g(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H} + Ce^{rt}(1-\rho(t))(1-\rho_h(t))\tilde{q}_h(t). \quad (50)$$

From the expressions of $\rho(t)$ and $\rho_h(t)$, time t can be inversely expressed as either

$$-\frac{1}{\lambda_H}\log[\frac{(1-\rho)\rho(t)}{\rho(1-\rho(t))}] \quad \text{or} \quad -\frac{1}{\lambda_H}\log[\frac{(1-\rho_h)\rho_h(t)}{\rho_h(1-\rho_h(t))}].$$

As a result, $Ce^{rt}(1-\rho(t))(1-\rho_h(t))\tilde{q}_h(t)$ can be written as:

$$\tilde{D}_1(1-\rho(t))(1-\rho_h(t))\tilde{q}_h(t)(\frac{1-\rho_h(t)}{\rho_h(t)})^{r/\lambda_H} + \tilde{D}_2(1-\rho(t))(1-\rho_h(t))\tilde{q}_h(t)(\frac{1-\rho(t)}{\rho(t)})^{r/\lambda_H}.$$

When the two buyers are identical, there should be no difference in the values: $Z(\rho(t), \rho_h(t)) = 0$ for $\rho(t) = \rho_h(t)$. This implies $\tilde{D}_1 = -\tilde{D}_2 = D_h$. Drop the time index t to transform $Z(t)$ back into $Z(\rho, \rho_h)$:

$$Z(\rho, \rho_h) = (\rho - \rho_h)\tilde{q}_h g - [(1-\rho_h)(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (1-\rho)(\frac{1-\rho}{\rho})^{r/\lambda_H}]\tilde{q}_h g(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H} + D_h(1-\rho)(1-\rho_h)\tilde{q}_h[(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (\frac{1-\rho}{\rho})^{r/\lambda_H}]. \quad (51)$$

Observe that: after the non-deviator stops purchasing the risky product, the deviator always receives the outside option. This implies a boundary condition for $Z(\rho, \rho_h)$: $Z(f(\rho_S^\star; h), \rho_S^\star) = 0$. The constant D_h can be pinned down by the boundary condition:

$$D_h = -\frac{(e^{\lambda_H h} - 1)g}{1 - e^{-rh}}(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H} + \frac{\left[1 + (e^{\lambda_H h} - 1)\rho_S^\star - e^{-rh}\right] g}{(1-\rho_S^\star)(1-e^{-rh})}(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H}. \quad (52)$$

Summing up U^{ND} and Z yields an expression for $U^D(\rho, \rho_h)$:

$$U^D(\rho, \rho_h) = s + (\rho - \rho_h)\tilde{q}_h g + \frac{(1-\rho)\tilde{q}_h(\frac{1-\rho}{\rho})^{r/\lambda_H}}{(1-\rho_h)q_h(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H}}(U_S(\rho_h) - s) - [(1-\rho_h)(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (1-\rho)(\frac{1-\rho}{\rho})^{r/\lambda_H}]\tilde{q}_h g(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H} + D_h(1-\rho)(1-\rho_h)\tilde{q}_h[(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (\frac{1-\rho}{\rho})^{r/\lambda_H}], \quad (53)$$

where D_h is given by equation (52).

In a *niche market*, 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 $H(t) = r(\rho(t) - \rho_h(t))\tilde{q}_h(t)g$. It is straightforward to solve differential equation (49):

$$Z(t) = \frac{rg}{r + \lambda_H}(\rho(t) - \rho_h(t))\tilde{q}_h(t) + Ce^{rt}(1-\rho(t))(1-\rho_h(t))\tilde{q}_h(t). \quad (54)$$

Repeating the above procedure yields

$$Z_3(\rho,\rho_h) = \frac{rg}{r+\lambda_H}(\rho-\rho_h)\tilde{q}_h + D_{h3}(1-\rho)(1-\rho_h)\tilde{q}_h[(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (\frac{1-\rho}{\rho})^{r/\lambda_H}], \quad (55)$$

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

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

Similarly, we solve Z as:

$$Z_2(\rho,\rho_h) = \frac{rg}{r+\lambda_H}(\rho-\rho_h)\tilde{q}_h - \frac{\lambda_H g}{r+\lambda_H}\rho_h(1-\rho)\tilde{q}_h + \rho_h(1-\rho)\tilde{q}_h g[\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}]^{r/\lambda_H}$$
$$+ D_{h2}(1-\rho)(1-\rho_h)\tilde{q}_h(\frac{1-\rho}{\rho})^{r/\lambda_H}. \quad (56)$$

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

$$Z_1(\rho,\rho_h) = (\rho-\rho_h)\tilde{q}_h g - [(1-\rho_h)(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (1-\rho)(\frac{1-\rho}{\rho})^{r/\lambda_H}]\tilde{q}_h g(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H}$$
$$+ D_{h1}(1-\rho)(1-\rho_h)\tilde{q}_h[(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (\frac{1-\rho}{\rho})^{r/\lambda_H}]. \quad (57)$$

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 (42) and we get the equations stated in lemma 1. ■

B.4 "Binding" Incentive Constraint

Lemma 2 *Fix a pair of priors (ρ_0, q_0) such that $\rho_S^\star$ is the equilibrium cutoff in the social learning phase. For $\rho > \rho_S^\star$, we must have:*

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

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

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

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

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

But 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.5 Proof of Proposition 3

Proof. The necessity part directly comes from lemma 1 and lemma 2. To prove the sufficiency part, the first step is to show there does not exist profitable one-shot deviations.

Lemma 3 *The value functions derived are sufficient to deter one-shot deviations: it is not profitable for an experimenting buyer to deviate for any $h \geq 0$ length of time.*

Proof. After a buyer deviates h length of time, the monopolist can either make a sell 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 immediate to see that it is not profitable to deviate. Therefore, the interesting case happens when the monopolist makes a sell to both buyers after an h-deviation.

In a mass market, the value associated with an $h > 0$ deviation is given by:

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

where $U^D(\rho, \rho_h)$ satisfies equation (53).

Rearranging terms yields

$$\hat{U}(\rho; h) - s = e^{-rh}[1 - \rho q(1 - e^{-\lambda_H h})](U^D(\rho, \rho_h) - s). \tag{58}$$

Using the expressions that

$$\rho_h = \frac{\rho e^{-\lambda_H h}}{1 - \rho(1 - e^{-\lambda_H h})} \quad \text{and} \quad \tilde{q}_h = \frac{q[1 - \rho(1 - e^{-\lambda_H h})]}{1 - \rho q(1 - e^{-\lambda_H h})},$$

we can directly evaluate $U_S(\rho) - \hat{U}(\rho; h)$ and get

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

A sufficient condition for $U_S(\rho) - \hat{U}(\rho; h) \geq 0$ is that 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, for any $h > 0$, there is no profitable one-shot deviation.

In a niche market, we have to consider the following two cases.

Case 1. $\rho \leq \rho_I^\star$. In this case, it is straightforward to show

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

and

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

In order to show $\hat{U}(\rho;h) \leq U(\rho)$, it suffices to prove for all $h \geq 0$, $S(h) \geq 0$ and $T(h) \geq 0$, which have been shown already.
Case 2. $\rho > \rho_I^\star$. In this case, $\rho_h > \rho_I^\star$ for h sufficiently small and we have:

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

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

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

We have shown that $T(h) \geq 0$. It is straightforward to check that

$$X(h) \triangleq e^{(r+\lambda_H)h} S(h) - \frac{rT(h)}{r+\lambda_H} - \frac{(r+\lambda_H)e^{-rh} - \lambda_H - re^{-(r+\lambda_H)h}}{r+\lambda_H} \geq 0.$$

This implies that it is not profitable to deviate in a niche market as well. ■

The next step is to show after some deviations, both the deviator and the non-deviator do not want to have another deviation.

Lemma 4 *Given the deviator has deviated h length of time in total such that the posterior beliefs are ρ and ρ_h, respectively, it is not profitable for both buyers to have another deviation.*

Proof. First, assume after the deviation, the monopolist is selling only to the deviator. Then setting $U^D(\rho, \rho_h) = s$ is sufficient to deter deviations. If the monopolist is making a sell to both buyers, then given the expressions of off the equilibrium path value function $U^D(\rho, \rho_h)$, we are also able to show it is not profitable to deviate for h' length of time. The proof is similar to the tedious proof of lemma 3 and is omitted.

Second, for the non-deviator, if the monopolist is only selling to the deviator, it is not profitable for the non-deviator to purchase the risky product since she is more pessimistic. We only need to show, if the monopolist is selling to both buyers, 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 any longer once h' exceeds h. The value associated with an h'-deviation is provided by:

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

Given

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

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 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 T sufficiently large 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 T sufficiently large, this new strategy also generates a value higher than $U_S(\rho)$.[19] Similarly define $\hat{\alpha}_1^k$ 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 $[0, T)$. Lemma 3 and lemma 4 together imply that any finite deviation is not profitable: $U_1(\hat{\alpha}_1^k, P^*, \alpha_2^*; \rho) - U_S(\rho) \leq 0$ for all k. But by the construction of admissible strategies,

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

which leads to a contradiction. ■

B.6 Proof of Proposition 4

Proof. In a niche market, $U_S(\rho_S^\star) = s$ and equation (21) implies

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

Substituting this expression into equation (26) yields

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

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

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

In a mass market, similarly we get $\rho_S^\star$ should also satisfy

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

Thus, the equilibrium cutoff $\rho_S^\star$ is characterized by equation (29) regardless of whether it is a mass or niche market. Since $\rho q(\rho)$, $V_I(\rho)$ and $J_I(\rho)$ are all increasing in ρ, the solution to the above equation is unique given a pair of priors (ρ_0, q_0).

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

Furthermore, a mass market appears ($\rho_S^\star > \rho_I^\star$) if and only if

$$\rho_I^\star q(\rho_I^\star) < \frac{rs}{rg + \lambda_H(V_I(\rho_I^\star) + J_I(\rho_I^\star)) - \lambda_H s}$$

or equivalently,

$$\frac{q_0(1-\rho_0)^2}{q_0(1-\rho_0)^2 + (1-q_0)(1-\rho_I^\star)^2} < \frac{\rho_I^e}{\rho_I^\star}.$$

Rearrange terms and we get the condition stated in the proposition.

From proposition 1, the efficient cutoff ρ_S^e is characterized by equation

$$\rho q(\rho) = \frac{rs}{(r+\lambda_H)g + \lambda_H W(\rho) - 2\lambda_H s}.$$

First, $J_I(\rho) + V_I(\rho) + s$ represents the total equilibrium surplus in the individual learning phase, and hence must be strictly less than the socially optimal surplus $\Omega_1(\rho) = g + W(\rho)$ for any $\rho > \rho_I^e$ since equilibrium is inefficient in the individual learning phase. Therefore,

$$rg + \lambda_H(V_I(\rho) + J_I(\rho)) - \lambda_H s < (r+\lambda_H)g + \lambda_H W(\rho) - 2\lambda_H s. \tag{59}$$

Second, it cannot be the case that $\rho_S^\star \le \rho_I^e$ for $q_0 < 1$. Otherwise, $V_I(\rho_S^\star) = s, J_I(\rho_S^\star) = g - s$ and $V_I(\rho_S^\star) + J_I(\rho_S^\star) = g$ imply

$$\rho_S^\star \times q(\rho_S^\star) = \rho_I^e = \frac{rs}{rg + \lambda_H(g-s)}. \tag{60}$$

The above equation contradicts the assumption that $\rho_S^\star \le \rho_I^e$.

Since $W(\cdot)$ is a strictly increasing function for $\rho > \rho_I^e$, inequality (59) implies that $\rho_S^\star > \rho_S^e$. ■

B.7 Proof of Proposition 5

Proof. Given the monopoly price $P_S(q)$ (notice $\rho = 1$ and we should switch to use q as the state variable), the value function for a representative unknown buyer can be written as

$$rU_S(q) = r(gq - P_S(q)) + nq\lambda_H(s - U_S(q)) - n\lambda_H q(1-q)U_S'(q). \tag{61}$$

Participation constraint implies that $U_S(q) \ge s$ and there is also an incentive compatibility constraint which means "one-shot deviations" are not profitable:

$$U_S(q) \ge \hat{U}(q;h) = \int_{t=0}^{h} re^{-rt}sdt + e^{-rh}q(1-e^{-(n-1)\lambda_H h})s + e^{-rh}(1-q+qe^{-(n-1)\lambda_H h})U_S(q_h)$$

for any $h > 0$ where $q_h = \frac{qe^{-(n-1)\lambda_H h}}{1-q+qe^{-(n-1)\lambda_H h}}$. Let h go to zero and the incentive constraint is binding such that the following differential equation is satisfied:

$$U_S(q) = s + \frac{n-1}{r}\left[q\lambda_H(s - U_S(q)) - \lambda_H q(1-q)U_S'(q)\right]$$

for $q \ge q_S^\star$. The general solution is

$$U_S(q) = s + D_S(1-q)(\frac{1-q}{q})^{r/((n-1)\lambda_H)}.$$

On the other hand, given price $P_S(\rho)$, the monopolist's value function is given by:

$$rJ_S(q) = nrP_S(q)dt + nq\lambda_H(n(g-s) - J_S(q)) - n\lambda_H q(1-q)J_S'(q). \tag{62}$$

At the optimal stopping cutoff $q_S^\star$, value matching and smooth pasting conditions are satisfied:

$$U_S(q_S^\star) = s, \quad J_S(q_S^\star) = 0 \quad \text{and} \quad J_S'(q_S^\star) = 0. \tag{63}$$

Boundary conditions (63) imply that $U_S(q_S^\star) = s$ for some $q_S^\star < 1$. As a consequence, it must be the case that $D_S = 0$ and $U_S(q)$ is always s. From equation (61), the equilibrium price is $P_S(q) = gq - s$. Substituting the price expression into equation (62) yields

$$rJ_S(q) = nr(gq-s) + nq\lambda_H(n(g-s) - J_S(q)) - n\lambda_H q(1-q)J_S'(q).$$

This is an ordinary differential equation with boundary conditions

$$J_S(q_S^\star) = 0 \quad \text{and} \quad J_S'(q_S^\star) = 0.$$

It is easy to solve $q_S^\star$ as:

$$q_S^\star = q_S^e = \frac{rs}{n\lambda_H(g-s) + rg}.$$

Therefore, the Markov perfect equilibrium is efficient. ■

B.8 Proof of Theorem 1

Proof. In the individual learning phase, 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

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

for all $1 \leq k \leq n-2$.

Proof. 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 buyers have already received lump-sum payoffs, the monopolist faces the same tradeoff of exploitation and exploration as in the two-buyer case. The monopolist has to charge $g\rho - s$ to keep the unknown buyer experimenting and her value function from selling to the unknown buyer is written as:

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

with boundary conditions

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

It is straightforward to see that:

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

and

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

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

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

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

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

This contradicts the assumption that $\rho^\star_{n-2} > \rho^\star_{n-1}$. Therefore, it must be the case that $\rho^\star_{n-2} \le \rho^\star_{n-1}$ 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 $1 \le j \le n-1$, assume

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

for $k \ge j+1$. 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

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

Standard induction argument then implies that for all $1 \le k \le n-1$, 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

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

for all $1 \le k \le n-2$. ■

Lemma 5 means the equilibrium is inefficient in the individual learning phase. From the boundary conditions, the equilibrium cutoff $\rho^\star_S$ in the social learning phase should satisfy

$$\rho^\star_S q(\rho^\star_S) = \frac{rs}{rg + \lambda_H\left[V_1(\rho^\star_S) + J_1(\rho^\star_S) + (n-1)U_1(\rho^\star_S)\right] - n\lambda_H s}.$$

The inefficiency in the individual learning phase means

$$V_1(\rho) + J_1(\rho) + (n-1)U_1(\rho) < g + (n-1)W(\rho) = \Omega_1(\rho)$$

for $\rho > \rho^e_I$ and hence

$$rg + \lambda_H[V_1(\rho) + J_1(\rho) + (n-1)U_1(\rho)] - n\lambda_H s < (r+\lambda_H)g + \lambda_H(n-1)W(\rho) - n\lambda_H s.$$

This implies that the equilibrium is inefficient in the social learning phase as well: $\rho^\star_S > \rho^e_S$. ■

C Proofs of Results from Section 4

C.1 Proof of Proposition 7

Proof. Notice the derivative of

$$\frac{r}{\lambda_H}\log(\frac{\rho}{1-\rho})+\log(\frac{q_0(1-\rho_0)^n+(1-q_0)(1-\rho)^n}{(1-\rho)^n})$$

is $\frac{r+\lambda_H n\rho q}{\lambda_H\rho(1-\rho)}$. From observation 1, a general solution to differential equation (32) is

$$\Omega_S(\rho)=\frac{\int h(x)\frac{rn[A-xq(x)B]+\lambda_H nxq(x)[(n-1)W(x)+s]}{\lambda_H x(1-x)}dx}{h(\rho)}$$

where

$$h(\rho)=(\frac{\rho}{1-\rho})^{r/\lambda_H}\frac{q_0(1-\rho_0)^n+(1-q_0)(1-\rho)^n}{(1-\rho)^n}.$$

First, we show ρ_I^e is always smaller than ρ_S^e.

Lemma 6 *Given any $q_0<1$, the efficient cutoff for starting experimentation in the social learning phase is larger than the efficient cutoff in the individual learning phase: $\rho_S^e>\rho_I^e$.*

Proof. For $\rho\le\rho_I^e$,

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

We solve for $\Omega_S(\rho)$ using integration by parts:

$$\Omega_S(\rho)=\frac{\int h(x)\frac{rn[A-xq(x)B]+\lambda_H nxq(x)[(n-1)W(x)+s]}{\lambda_H x(1-x)}dx}{h(\rho)}=n\left[A-\frac{\lambda_H}{r+\lambda_H}\rho q(\frac{rB}{\lambda_H}+A-s)\right]+\frac{C}{h(\rho)}.$$

Since 0 is included in the domain of $\Omega_S(\cdot)$, the constant term C must be 0 to guarantee $\Omega_S(\cdot)$ is bounded away from infinity. Therefore,

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

Suppose on the contrary, we have $\rho_S^e\le\rho_I^e$, then ρ_S^e should satisfy

$$n\left[A-\frac{\lambda_H}{r+\lambda_H}\rho_S^e q(\rho_S^e)(\frac{rB}{\lambda_H}+A-s)\right]=ns\Longrightarrow\rho_S^e q(\rho_S^e)=\rho_I^e.$$

This leads to a contradiction since $q<1$. ■

For $\rho>\rho_I^e$, $W(\rho)=s$ and by observation 1,

$$\Omega_S(\rho)=\frac{\int_{\rho_I^e}^{\rho}h(x)\frac{rn[A-xq(x)B]+\lambda_H n^2xq(x)s}{\lambda_H x(1-x)}dx+C}{h(\rho)}.$$

The constant C is chosen such that $\Omega_S(\rho)$ is continuous at ρ_I^e:

$$C=h(\rho_I^e)\Omega_S(\rho_I^e)=h(\rho_I^e)n\left[A-\frac{\lambda_H}{r+\lambda_H}\rho_I^e q(\rho_I^e)(\frac{rB}{\lambda_H}+A-s)\right]>0.$$

At the efficient starting cutoff $\rho_S^e(q_0)$, $\Omega_S(\rho_S^e; q_0) = ns$. Substituting the expression of $\Omega_S(\rho)$ into the above equation yields:

$$C - h(\rho_I^e)ns + \int_{\rho_I^e}^{\rho_S^e} h(x)\frac{rn[A - xq(x)B - s]}{\lambda_H x(1-x)}dx = 0.$$

Notice

$$C - h(\rho_I^e)ns = h(\rho_I^e)n\left[A - s - \frac{\lambda_H}{r+\lambda_H}\rho_I^e q(\rho_I^e)(\frac{rB}{\lambda_H} + A - s)\right] > 0$$

doesn't depend on ρ_S^e. This implies: if an interior solution $\rho_S^e(q_0)$ exists, it must be the case that

$$\int_{\rho_I^e}^{\rho_S^e} h(x)\frac{rn[A - xq(x)B - s]}{\lambda_H x(1-x)}dx < 0$$

and hence $A - \lambda_H \rho_S^e q_0 B - s < 0$. Suppose for a given q_0, there exist two efficient cutoffs ρ_1 and $\rho_2 > \rho_1$. Then we have

$$\int_{\rho_I^e}^{\rho_1} h(x)\frac{rn[A - xq(x)B - s]}{\lambda_H x(1-x)}dx = \int_{\rho_I^e}^{\rho_2} h(x)\frac{rn[A - xq(x)B - s]}{\lambda_H x(1-x)}dx,$$

which is impossible since

$$h(x)\frac{rn[A - xq(x)B - s]}{\lambda_H x(1-x)} < 0$$

for $x \in (\rho_1, \rho_2)$. Therefore, if there exists some ρ_S^e satisfying $\Omega_S(\rho_S^e; q_0) = ns$, such ρ_S^e must be unique. When there does not exist ρ_S^e satisfying

$$C - h(\rho_I^e)ns + \int_{\rho_I^e}^{\rho_S^e} h(x)\frac{rn[A - xq(x)B - s]}{\lambda_H x(1-x)}dx = 0,$$

just set $\rho_S^e = 1$ since it is always beneficial to take the risky product. To summarize, for any q_0, there is a unique $\rho_S^e(q_0)$ such that it is socially efficient to start experimentation if and only if $\rho \leq \rho_S^e(q_0)$. ■

C.2 Proof of Theorem 2

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

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

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 $\rho \leq \rho^\star_{n-1} = \rho^e_I$. 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 $\rho \geq \rho^e_I$ for all $k \geq 1$. Therefore, the symmetric Markov perfect equilibrium is efficient in the individual learning phase. In the social learning phase, for $\rho \leq \rho^e_I$, the monopolist's value function is

$$J_S(\rho) = n\left(A - \rho q B - s\right) + \frac{1}{r}\left[n\lambda_H\rho q(J_1(\rho) - J_S(\rho)) - \lambda_H\rho(1-\rho)J'_S(\rho)\right].$$

The solution to the above differential equation is given by:

$$J_S(\rho) = n(A-s) - n\rho q(\rho)\frac{\lambda_H}{r+\lambda_H}(A - s + \frac{rB}{\lambda_H}).$$

It is easy to check that for any $q < 1$, $J_S(\rho) > 0$ for all $\rho \leq \rho^e_I$ and hence the equilibrium cutoff in the social learning phase must be larger than ρ^e_I. For $\rho > \rho^e_I$,

$$J_S(\rho) = n\left[A - \rho q B - s\right] - \frac{1}{r}\left[n\lambda_H\rho q J_S(\rho) + \lambda_H\rho(1-\rho)J'_S(\rho)\right].$$

Solving the above differential equation yields

$$J_S(\rho) = \frac{\int_{\rho^e_I}^{\rho} h(x)\frac{rn(A - xq(x)B - s)}{\lambda_H x(1-x)}dx + D}{h(\rho)}$$

where

$$h(\rho) = (\frac{\rho}{1-\rho})^{r/\lambda_H}\frac{q_0(1-\rho_0)^n + (1-q_0)(1-\rho)^n}{(1-\rho)^n}.$$

The constant D is chosen such that $J_S(\cdot)$ is continuous at ρ^e_I. This implies: $D = C - h(\rho^e_I)ns$, where C is the constant given in the proof of proposition 7. From integration by parts,

$$\int_{\rho^e_I}^{\rho} h(x)\frac{rn(A - xq(x)B - s)}{\lambda_H x(1-x)}dx = \int_{\rho^e_I}^{\rho} h(x)\frac{rn(A - xq(x)B) + \lambda_H n^2 xq(x)s}{\lambda_H x(1-x)}dx - ns(h(\rho) - h(\rho^e_I)).$$

As a consequence, $J_S(\rho) = \Omega_S(\rho) - ns$.

For a fixed q_0, the monopolist starts selling her product as long as $J_S(\rho_0; q_0) \geq 0$, which implies that the equilibrium cutoff $\rho_S^\star(q_0)$ must be the same as $\rho_S^e(q_0)$. Therefore, the symmetric Markov perfect equilibrium is efficient in the social learning phase as well. ■

References

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

Banerjee, B. (1992): "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107(3), 797–817.

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

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

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

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

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

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

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100(5), 992–1026.

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

Bonatti, A. (2009): "Menu Pricing and Learning," mimeo.

Bonatti, A., and J. Hörner (2009): "Collaborating," forthcoming in *American Economic Review.*

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

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

Ching, A. (2010): "A Dynamic Oligopoly Structural Model for the Prescription Drug Market After Patent Expiration," *International Economic Review*, 51(4), 1175–1207.

Chintagunta, P., R. Jiang, and G. Jin (2009): "Information, Learning, and Drug Diffusion: the Case of Cox-2 Inhibitors," *Quantitative Marketing and Economics*, 7(4), 399–433.

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

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

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.

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

Hörner, J., and L. Samuelson (2009): "Incentives for Experimenting Agents," Yale mimeo.

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

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

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

Kim, M. (2010): "Physician Learning and New Drug Diffusion," mimeo.

Murto, P., and J. Välimäki (2009): "Learning and Information Aggregation in an Exit Game," mimeo.

New York Times (2010): "Panel Urges F.D.A. to Revoke Approval of Drug for Breast Cancer Treatment," `http://prescriptions.blogs.nytimes.com/2010/07/20/panel-urges-f-d-a-to-revoke-approval-of-drug-for-breast-cancer-treatment`.

Rosenberg, D., E. Solan, and N. Vieille (2007): "Social Learning in One-Arm Bandit Problems," *Econometrica*, 75(6), 1591–1611.

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

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