Spring 5-16-2011

Essays on Market Dynamics in the Presence of Learning

Xi Weng
wengxi125@gmail.com

Essays on Market Dynamics in the Presence of Learning

Abstract
I investigate how the presence of learning affects the market dynamics in three different market settings. The first chapter 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. The second chapter is based on the observation that workers learn at different rates about their productivity and therefore expect different wage paths across firms. We show that under strict supermodularity there is always positive assortative matching: differential learning is always dominated by the impact of productivity. Surprisingly, this holds even if learning is faster in the low type firm. The key assumption driving this result is that this is a pure Bayesian learning model.We also derive a new equilibrium condition in this class of continuous time models in addition to the common smooth-pasting and value-matching conditions. This no-deviation condition captures sequential rationality and results in a restriction on the second derivative of the value function. The third chapter develops a continuous-time war of attrition model with learning to investigate whether learning is possible to make it easier to reach an agreement. I show that with exogenous private learning, it may be easier to reach an agreement initially but it becomes more and more difficult over time. The expected delay will always be higher than the expected delay without learning. I also show that when allowing only one player to learn leads to a shorter delay than allowing both to learn.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Economics

First Advisor
George Mailath

Second Advisor
Andrew Postlewaite

Third Advisor
Hanming Fang

Keywords
Microeconomics, Market Dynamics, Learning

Subject Categories
Economic Theory

ESSAYS ON MARKET DYNAMICS IN THE PRESENCE OF LEARNING

XI WENG

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

2011

George J. Mailath, Walter H. Annenberg Professor in the Social Sciences

Supervisor of Dissertation

Dirk Krueger, Professor of Economics

Graduate Group Chairperson

Dissertation Committee

Andrew Postlewaite, Harry P. Kamen Professor of Economics and Professor of Finance

Hanming Fang, Professor of Economics

Qingmin Liu, Assistant Professor of Economics

ESSAYS ON MARKET DYNAMICS

IN THE PRESENCE OF LEARNING

For Yu Zhang.

ACKNOWLEDGMENTS

I would never have been able to finish my dissertation without the guidance of my committee members and support from my family and wife.

I owe my deepest gratitude to my advisor, Professor George Mailath. The supervision from George can be summarized by an old Chinese proverb, “Just as bitter medicine cures sickness, so unpalatable advice benefits conduct.” George was very strict about even small shortcomings in my existing research. After I finished the dissertation, I realized that it is very difficult to count completely how many interesting research insights in my dissertation were inspired by George’s great intelligence. As a father, George was always ready to help me when I was unable to solve any problem. His excellent guidance and continuous support enabled me to overcome the difficulties along the way. I wish that at the end of life, I could have achieved even one tenth of George’s excellence, as a research, an advisor and a person.

I have been so lucky to have Professor Hanming Fang, Professor Qingmin Liu and Professor Andrew Postlewaite on my committee. I am immensely grateful for their inspiring advice and invaluable input in my dissertation. Special thanks goes to Professor Jan Eeckhout, with whom I wrote one chapter of my dissertation.

I would also like to thank my parents. They have always been supporting me and encouraging me with their best wishes and greatest love.

Finally, I would like to thank my wife, Yu Zhang. She was always there cheering me up and stood by me through the good times and bad.

ABSTRACT

ESSAYS ON MARKET DYNAMICS IN THE PRESENCE OF LEARNING

Xi Weng

George Mailath

I investigate how the presence of learning affects the market dynamics in three different market settings. The first chapter 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. The second chapter is based on the observation that workers learn at different rates about their productivity and therefore expect different wage paths across firms. We show that under strict supermodularity there is always positive assortative matching: differential learning is always dominated by the impact of productivity. Surprisingly, this holds even if learning is faster in the low type firm. The key assumption driving this result is that this is a pure Bayesian learning model.We also derive a new equilibrium condition in this class of continuous time models in addition to the common smooth-pasting and value-matching conditions. This *no-deviation condition* captures sequential rationality and results in a restriction on the second derivative of the value function. The third chapter develops a continuous-time war of attrition model with learning to investigate whether learning is possible to make it easier to reach an agreement. I show that with exogenous private learning, it may be easier to

reach an agreement initially but it becomes more and more difficult over time. The expected delay will *always* be higher than the expected delay without learning. I also show that when allowing only one player to learn leads to a shorter delay than allowing both to learn.

Table of Contents

List of Figures

Chapter 1

Dynamic Pricing in the Presence of Social Learning

1.1 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.

[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)).

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 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

[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).

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 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) de-

[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.

velop 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 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

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

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 1.2 introduces the model and defines the solution concept. Section 1.3 and Section 1.4 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 1.5 concludes the paper.

1.2 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* ($\lambda = \lambda_H$) 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

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

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 1.1. *(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 1.2. *(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.

1.2.1 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:

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

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]

$$\dot{\rho}_{it} = -\lambda_H a_{it} \rho_{it} (1 - \rho_{it}). \tag{1.1}$$

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

[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 $\tau \leq \bar{h}$. Hence by definition,

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

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

differential equation at those times when a_{it} is right continuous for $\forall i$:

$$\dot{q}_t = -\lambda_H q_t(1-q_t)\sum_{i=1}^{n} a_{it}\rho_{it}. \tag{1.2}$$

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

$$q_t = \frac{q_0 \prod_{i=1}^{n} x_{it}}{q_0 \prod_{i=1}^{n} x_{it} + 1 - q_0}. \tag{1.3}$$

From equation (1.1),

$$\rho_{it} = \frac{\rho_0 e^{-\lambda_H \int_0^t a_{i\tau} d\tau}}{x_{it}} \Longrightarrow 1 - \rho_{it} = \frac{1-\rho_0}{x_{it}}. \tag{1.4}$$

Substituting (1.4) into (1.3) yields:

$$q_t = \frac{q_0(1-\rho_0)^n}{q_0(1-\rho_0)^n + (1-q_0)\prod_{i=1}^{n}(1-\rho_{it})}. \tag{1.5}$$

Notice that equation (1.5) 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 (1.5) is that it only depends on ρ_{it}'s and does not explicitly depend on previous purchasing decisions or time t. Differential equations (1.1) and (1.2) 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).

1.2.2 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

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

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 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 (1.5).

---

[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.

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.

1.2.3 Symmetric Markov Perfect Equilibrium

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

Definition 1.1. *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]

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

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* $\alpha_i(\boldsymbol{\rho}, P) = \alpha_{\pi(i)}(\tilde{\boldsymbol{\rho}}, P)$*.*

1.3 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 1.1 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 (1.1).

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 (1.5):

$$q = \frac{(1-\rho_0)^n q_0}{(1-\rho_0)^n q_0 + (1-\rho)^n (1-q_0)}. \tag{1.6}$$

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

---

[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.

1.3.1 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 1.2 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

$$W(\rho) = \max\left\{s, g\rho + \frac{\lambda_H s}{r+\lambda_H}(\frac{rs}{(r+\lambda_H)(g-s)})^{r/\lambda_H}(1-\rho)(\frac{1-\rho}{\rho})^{r/\lambda_H}\right\}. \tag{1.7}$$

Efficiency in the Social Learning Phase In the social learning phase, the socially efficient allocation solves the symmetric cooperative problem (see claim A.1 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:

$$r\Omega_S(\rho) = \max\left\{rns, rn\rho q(\rho)g + n\rho q(\rho)\lambda_H(\Omega_1(\rho) - \Omega_S(\rho)) - \lambda_H\rho(1-\rho)\Omega_S'(\rho)\right\}, \tag{1.8}$$

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:

$$r\Omega_S(\rho) = rn\rho q(\rho)g + n\rho q(\rho)\lambda_H(\Omega_1(\rho) - \Omega_S(\rho)) - \lambda_H \rho(1-\rho)\Omega_S'(\rho), \tag{1.9}$$

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 (1.9) and we immediately show that the cutoff ρ_S^e should satisfy

$$rn\rho q(\rho)g + n\rho q(\rho)\lambda_H\Omega_1(\rho) = (r + n\rho q(\rho)\lambda_H)ns. \tag{1.10}$$

In the appendix, we show that equation (1.10) 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 1.1. *(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. □

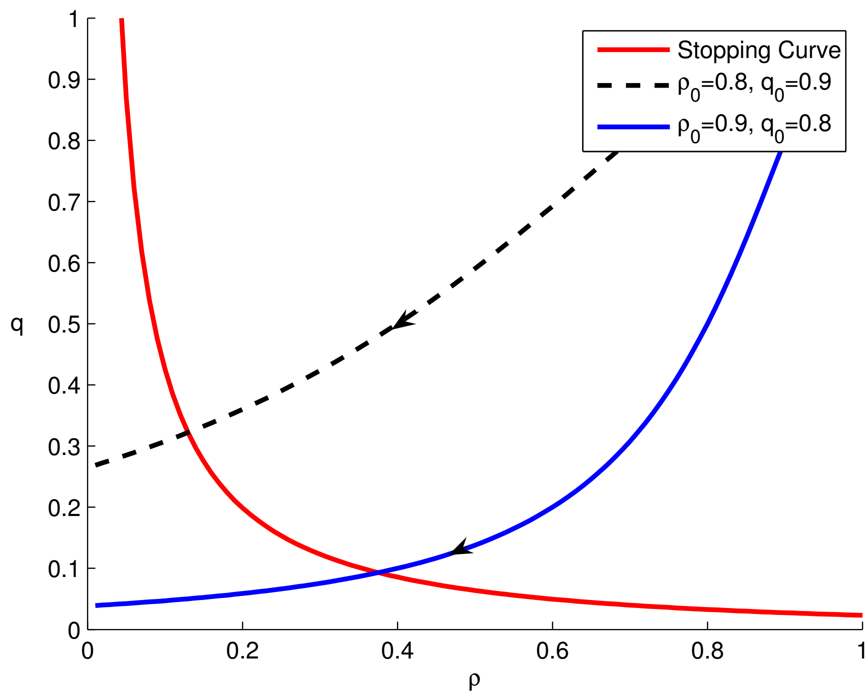

Figure 1.1: Solutions to the Cooperative Problem with Two Players

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

$$q_S^e = \frac{(1-\rho_0)^n q_0}{(1-\rho_0)^n q_0 + (1-\rho_S^e)^n (1-q_0)} \tag{1.11}$$

and

$$q_S^e = \frac{rs}{\rho_S^e[(r+\lambda_H)g + (n-1)\lambda_H W(\rho_S^e) - n\lambda_H s]}, \tag{1.12}$$

where $W(\cdot)$ is given by equation (1.7). Figure 1.1 is an illustration of how we can use equations (1.11) and (1.12) to determine the efficient cutoffs in the social learning phase. Equation (1.12) 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 (1.11) 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^e_S 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.

1.3.2 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 1.1 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:

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

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

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

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 (1.1); $q_\zeta(\rho_t) = 1$ for $\zeta = I$ and $q_\zeta(\rho_t)$ is given by equation (1.6) 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.

Niche Market vs. Mass Market

As in the social planner's problem, the equilibrium purchasing behavior can be characterized by two cutoffs $\rho^\star_S$ and $\rho^\star_I$. If no buyer has received lump-sum payoffs, the price is falling over time to keep both unknown buyers experimenting until posterior ρ reaches $\rho^\star_S$. 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 $\rho^\star_I$.

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 1.2. *(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.

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 1.2. *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

$$V_I(\rho) = \max\left\{s, s + g(1-\rho)(1 - [\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}]^{r/\lambda_H})\right\}; \tag{1.13}$$

and the monopolist receives value

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

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).

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 $h \leq \bar{h}$,

$$U_S(\rho) \geq \hat{U}(\rho; h) = \int_{t=0}^{h} re^{-rt} s dt + \rho q(1-e^{-\lambda_H h})e^{-rh}U_I(\rho) + [1-\rho q(1-e^{-\lambda_H h})]e^{-rh}U^D(\rho, \rho_h) \tag{1.14}$$

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 (1.14) 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 A.1 in the appendix. Here we just provide a sketch of the proof.

Sketch of the proof for lemma A.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:

$$
\begin{aligned}
U^D(\rho, \rho_h) = \mathbb{E} \int_{t=0}^{h'} & re^{-rt}(\rho_t q_t g - \tilde{P}_t)dt \\
& + \rho \tilde{q}_h (1 - e^{-\lambda_H h'}) e^{-rh'} V_I(\rho_{h+h'}) + \rho_h \tilde{q}_h (1 - e^{-\lambda_H h'}) e^{-rh'} s \\
& + [1 - \rho \tilde{q}_h (1 - e^{-\lambda_H h'}) - \rho_h \tilde{q}_h (1 - e^{-\lambda_H h'})] e^{-rh'} U(\rho_{h'}, \rho_{h+h'}). \quad (1.15)
\end{aligned}
$$

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:

$$
\begin{aligned}
U^{ND}(\rho, \rho_h) = \mathbb{E} \int_{t=0}^{h'} & re^{-rt}(\rho'_t q_t g - \tilde{P}_t)dt \\
& + \rho \tilde{q}_h (1 - e^{-\lambda_H h'}) e^{-rh'} s + \rho_h \tilde{q}_h (1 - e^{-\lambda_H h'}) e^{-rh'} V_I(\rho_{h'}) \\
& + [1 - \rho \tilde{q}_h (1 - e^{-\lambda_H h'}) - \rho_h \tilde{q}_h (1 - e^{-\lambda_H h'})] e^{-rh'} U(\rho_{h+h'}, \rho_{h'}), \quad (1.16)
\end{aligned}
$$

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 (1.15) and (1.16) 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 (1.16) from equation (1.15) 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 A.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 A.1 and lemma A.2 together gives an important characterization of the on-equilibrium-path value function U_S:

Proposition 1.3. *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*

$$0 = 2(r+\lambda_H\rho q)(U_S(\rho)-s)+\lambda_H\rho(1-\rho)U_S'(\rho)+(r+\lambda_H\rho)g(1-\rho)q(\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)})^{r/\lambda_H}$$
$$-\lambda_H g\rho(1-\rho)q-\left[\frac{r+\lambda_H\rho_S^\star}{1-\rho_S^\star}(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H}-\lambda_H(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H}\right]g(1-\rho)^2q(\frac{1-\rho}{\rho})^{r/\lambda_H}. \tag{1.17}$$

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

$$0 = 2(r+\lambda_H\rho q)(U_S(\rho)-s)+\lambda_H\rho(1-\rho)U_S'(\rho)$$
$$+\frac{r\lambda_H g}{r+\lambda_H}\frac{(1-\rho)^2q\rho_S^\star}{1-\rho_S^\star}(\frac{(1-\rho)\rho_S^\star}{\rho(1-\rho_S^\star)})^{r/\lambda_H}-\frac{rg}{r+\lambda_H}\lambda_H\rho(1-\rho)q \tag{1.18}$$

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

$$0 = 2(r+\lambda_H\rho q)(U_S(\rho)-s)+\lambda_H\rho(1-\rho)U_S'(\rho)+(r+\lambda_H\rho)g(1-\rho)q(\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)})^{r/\lambda_H}-\lambda_H g\rho(1-\rho)q$$
$$-r\left[\frac{r+\lambda_H+\lambda_H\rho_I^\star}{(r+\lambda_H)(1-\rho_I^\star)}(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H}-\frac{\lambda_H}{r+\lambda_H}(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H}\right]g(1-\rho)^2q(\frac{1-\rho}{\rho})^{r/\lambda_H} \tag{1.19}$$

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

The necessity of proposition 1.3 just comes from combining lemma A.1 and lemma A.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 1.3 can be solved by using observation A.1 in the appendix. In a mass market, for any $\rho > \rho_S^\star$, the value function $U_S(\rho)$ is given by

$$\begin{aligned} U_S(\rho) &= s+\frac{\lambda_H}{2r+\lambda_H}g\rho(1-\rho)q-g(1-\rho)q[\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}]^{r/\lambda_H} \\ &+ \left[\frac{r+\lambda_H\rho_S^\star}{r(1-\rho_S^\star)}(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H}-\frac{\lambda_H}{r}(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H}\right]g(1-\rho)^2q(\frac{1-\rho}{\rho})^{r/\lambda_H} \\ &+ C(1-\rho)^2q(\frac{1-\rho}{\rho})^{2r/\lambda_H}. \end{aligned} \tag{1.20}$$

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

$$\begin{aligned} U_S(\rho) &= s + \frac{r\lambda_H}{(2r+\lambda_H)(r+\lambda_H)} g\rho(1-\rho)q - \frac{\lambda_H g}{r+\lambda_H} \frac{\rho_S^\star(1-\rho)^2 q}{1-\rho_S^\star} (\frac{(1-\rho)\rho_S^\star}{\rho(1-\rho_S^\star)})^{r/\lambda_H} \\ &+ D(1-\rho)^2 q(\frac{1-\rho}{\rho})^{2r/\lambda_H}; \qquad (1.21) \end{aligned}$$

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

$$\begin{aligned} U_S(\rho) &= s + \frac{\lambda_H}{2r+\lambda_H} g\rho(1-\rho)q - g(1-\rho)q[\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}]^{r/\lambda_H} \\ &+ \left[\frac{r+\lambda_H+\lambda_H\rho_I^\star}{(r+\lambda_H)(1-\rho_I^\star)}(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H} - \frac{\lambda_H}{r+\lambda_H}(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H}\right] g(1-\rho)^2 q(\frac{1-\rho}{\rho})^{r/\lambda_H} \\ &+ (D - \frac{2\lambda_H g}{2r+\lambda_H}(\frac{\rho_I^\star}{1-\rho_I^\star})^{1+2r/\lambda_H})(1-\rho)^2 q(\frac{1-\rho}{\rho})^{2r/\lambda_H}. \qquad (1.22) \end{aligned}$$

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:

$$\begin{aligned} rU_S(\rho) = r(\rho q(\rho)g - P_S(\rho)) + \lambda_H\rho q(\rho)(U_I(\rho) - U_S(\rho)) + \lambda_H\rho q(\rho)(V_I(\rho) - U_S(\rho)) \\ - \lambda_H\rho(1-\rho)U_S'(\rho) \qquad (1.23) \end{aligned}$$

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 (1.13).

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

$$rJ_S(\rho) = 2rP_S(\rho) + 2\lambda_H\rho q(\rho)(J_I(\rho) - J_S(\rho)) - \lambda_H\rho(1-\rho)J_S'(\rho). \qquad (1.24)$$

where $J_I(\rho)$ is given by proposition 1.2.

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

Equations (1.23) and (1.24) are value functions if both unknown buyers purchase the risky product. The RHS of equation (1.23) 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 (1.24) 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,

$$P_S(\rho) = \rho q(\rho)g - s + \frac{\lambda_H}{2r+\lambda_H} g\rho(1-\rho)q(\rho) + Cq(\rho)(1-\rho)^2(\frac{1-\rho}{\rho})^{2r/\lambda_H} \tag{1.25}$$

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

$$P_S(\rho) = \rho q(\rho)g - s - \frac{\lambda_H}{2r+\lambda_H} g\rho(1-\rho)q(\rho) + Dq(\rho)(1-\rho)^2(\frac{1-\rho}{\rho})^{2r/\lambda_H} \tag{1.26}$$

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

$$\begin{aligned} P_S(\rho) = \rho q(\rho)g - s &+ \frac{\lambda_H}{2r+\lambda_H} g\rho(1-\rho)q(\rho) \\ &+ (D - \frac{2\lambda_H g}{2r+\lambda_H}(\frac{\rho_I^\star}{1-\rho_I^\star})^{1+2r/\lambda_H})q(\rho)(1-\rho)^2(\frac{1-\rho}{\rho})^{2r/\lambda_H} \end{aligned} \tag{1.27}$$

for $\rho > \rho_I^\star$. In the above equations, C and D are constants in equations (1.20) to (1.22). Notice in equations (1.26) and (1.27), 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 1.2, 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 (1.24) 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 1.4. *(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

$$\frac{1-q_0}{q_0(1-\rho_0)^2} > \frac{g}{(1-\rho_I^\star)s}. \tag{1.28}$$

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

$$\rho q(\rho) = \frac{rs}{rg + \lambda_H(V_I(\rho) + J_I(\rho)) - \lambda_H s}. \tag{1.29}$$

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 1.4. 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 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

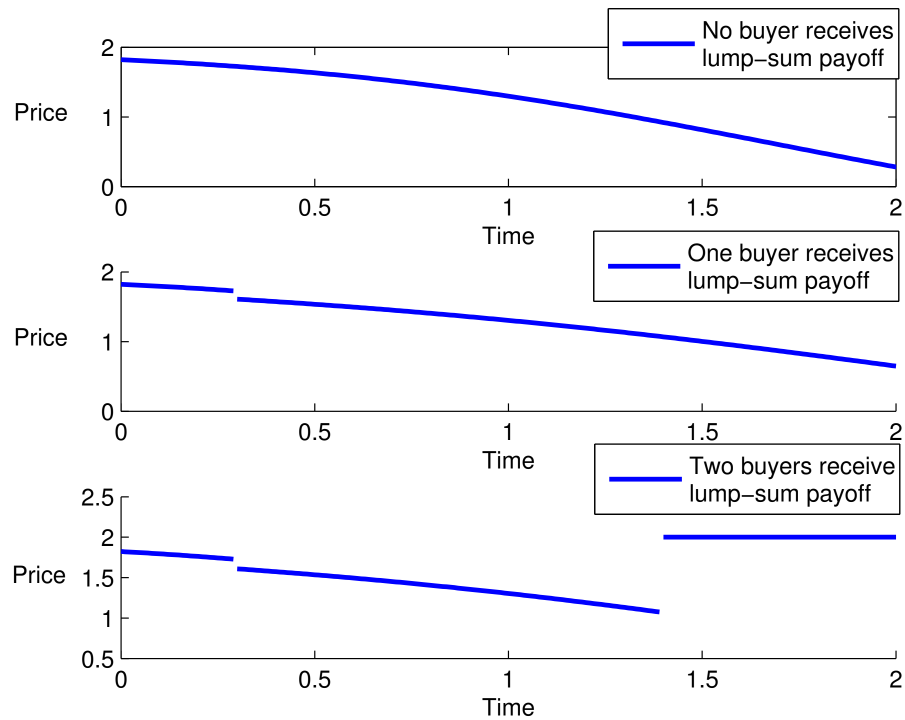

Figure 1.2: Equilibrium Price Dynamics

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.

Equilibrium Price Path

After solving for the equilibrium cutoff $\rho_S^{\star}$, the constants C and D in equations (1.20) and (1.21) can be pinned down from the value matching condition and then the expression for the equilibrium prices can be derived. Figure 1.2 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.

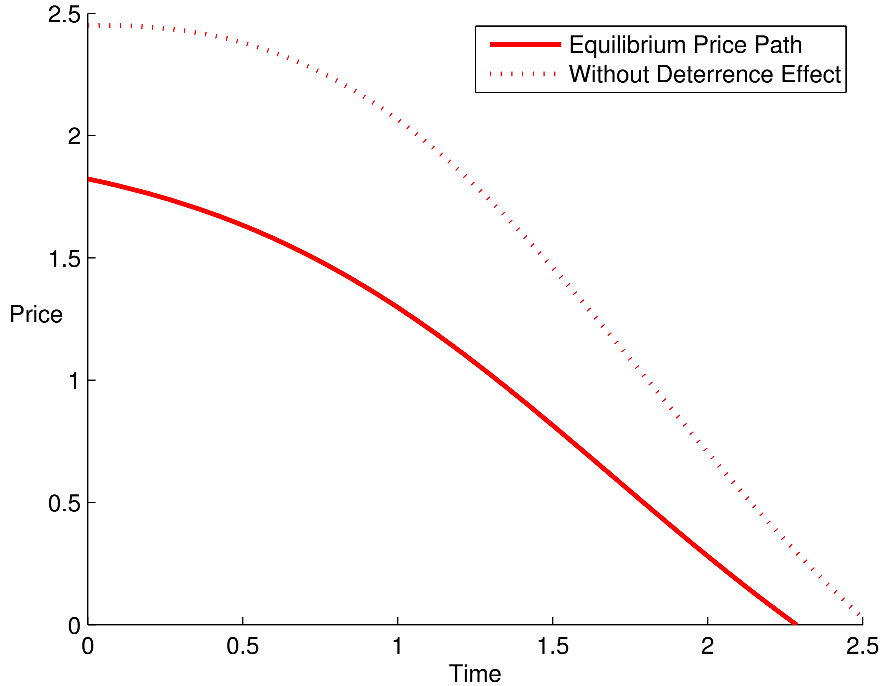

Figure 1.3: Deterrence Effect

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 1.3 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 1.2. 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 1.1. *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 1.4 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.

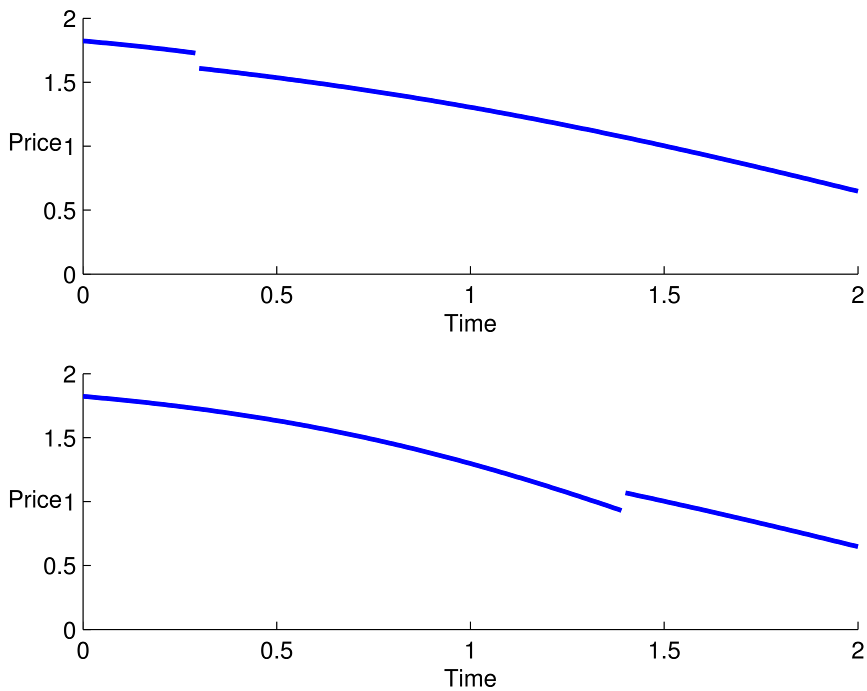

Figure 1.4: Instantaneous Price Response to the First Arrival of Good News

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

Proposition 1.5. *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 1.4 can be extended to a general n case. The induction argument is used to avoid solving for every value function explicitly.

Theorem 1.1. *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.

1.4 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.

1.4.1 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

$$W(\rho) = \max \left\{ s, A - \rho B + \frac{1}{r} \left[\lambda_H \rho(s - W(\rho)) - \lambda_H \rho(1-\rho)W'(\rho) \right] \right\}. \tag{1.30}$$

Solve the optimal starting problem defined by equation (1.30) and we get the following result:

Proposition 1.6. *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:

$$\Omega_S(\rho) = \max\left\{ns, n(A-\rho q(\rho)B) + \frac{1}{r}[\lambda_H n\rho q(\rho)(\Omega_1(\rho)-\Omega_S(\rho)) - \lambda_H\rho(1-\rho)\Omega'_S(\rho)]\right\}. \tag{1.31}$$

The optimal starting problem (1.31) is solved by solving differential equation

$$(r+\lambda_H n\rho q)\Omega_S(\rho) = rn(A-\rho qB) + \lambda_H n\rho q[(n-1)W(\rho)+s] - \lambda_H\rho(1-\rho)\Omega'_S(\rho), \tag{1.32}$$

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:

Proposition 1.7. *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 $\rho \le \rho_S^e(q_0)$.*

Proof. In the appendix. □

1.4.2 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 1.1 implies that the triple of $(P_k, \alpha_k^0, \alpha_k^1)$ is a symmetric Markov perfect equilibrium if:

[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.

- $\alpha_k^0 = 1$ 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 (1.1); $q_k(\rho_t) = 1$ for $k \geq 1$ and $q_k(\rho_t)$ is given by equation (1.6) 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 1.1. *In equilibrium, the most pessimistic unknown buyer's value is always s.*

Claim 1.1 implies that the on-equilibrium-path value for each unknown buyer is always s since they are equally pessimistic. This is different from proposition 1.3 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 1.1: 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 1.1. 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 1.2. *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.

1.5 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 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

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

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.

Chapter 2

Assortative Learning (Joint with Jan Eeckhout)

2.1 Introduction

High ability workers sort into more productive jobs. Due to complementarities in production, their higher marginal product allows them to command higher wages. The Beckerian model of assortative matching is very well suited to explain those patterns of sorting. Unfortunately, it is mute on the issue of turnover of workers between different jobs. Instead, the Jovanovic (1979) learning model has long been the canonical framework for analyzing turnover in the labor market[1] over the life cycle. Workers and firms learn about match-specific human capital and will tend to stay in a match if learning reveals the match is good. Experimentation occurs early on which leads to decreasing turnover over the life cycle. Because in Jovanovic (1979) learning is about the match and not about the worker, there is neither worker heterogeneity nor sorting. In this paper, we offer a unified approach of learning and sorting. We establish a solution method for a market equilibrium in a continuous time economy with multiple learning opportunities (multi-armed bandit) and derive a no-deviation condition, a condition hitherto unknown. We show that under supermodularity, positive assortative matching obtains in equilibrium, even if learning rates differ across firms.

[1] Of course, also the search model inherently exhibits turnover, but with observable types turnover is constant over the life cycle. Moscarini (2005) brings together search and learning in the Jovanovic framework.

In the labor market, the learning experiences of workers are most likely to differ across different firms. Starting in a top law firm or a multinational will induce different paths of information revelation than working in a local family business. The worker now faces a trade-off between different experimentation experiences: take a lower wage at a high productivity firm where information may be revealed at a different rate or accept higher wage and learn more slowly. It is intuitive that sorting and learning are intimately connected.

Modelling the labor market as a multi-armed bandit problem and solving it is challenging. Most existing learning models and continuous time games are tractable because they are essentially one-armed bandit problems with a fixed outside option that acts as an absorbing state. One-armed bandit problems typically have attractive properties, including reservation strategies. Instead, multi-armed bandits in general do not have reservation strategies when arms are correlated, even if the learning rate is the same across firms. But our labor market is not exactly identical to the canonical bandit problem. First, there are a continuum of experimenters. Second, because of competitive wage determination à la Jovanovic (1979), the payoffs are endogenous. Finally, because workers learn about general human capital instead of match-specific human capital, the arms are positively correlated.

We find that it is the combination of competitive wage determination (endogenous pay-offs)and the incentives needed to avoid a deviation that give rise to a new condition which we call the *no-deviation condition*. This condition must be satisfied in addition to the common equilibrium conditions of value-matching and smooth-pasting. The no-deviation condition can be interpreted as the continuous time version of the one-shot deviation principle.[2] We prove that the no-deviation condition implies that the second derivative of worker's value function at the cut-off belief is the same in the high type as well as in the low type firms.

[2]The idea of sequential rationality is of course not new and has also been employed in continuous time games by Sannikov (2007) who uses the concept of self generation. And Cohen and Solan (2009) use dependence of strategies on a small interval dt to restrict the set of Markovian strategies, in the spirit of our dt-shot deviation. It is precisely the one-shot deviation *in conjunction* with endogenous payoffs that leads to the equalization of the second derivative of the value functions.

Recall that value matching requires that at the cut-off the worker's value functions take the same value in both firms, the smooth-pasting condition requires that the first derivative is the same, and now the no-deviation requires equal second derivatives as well.

We show that supermodularity of the production technology is a necessary and sufficient condition for positive assortative matching, and that the equilibrium allocation is unique. Those workers with the highest beliefs about their ability will in equilibrium sort into those firms that are most productive. Moreover, we can analytically solve for the equilibrium allocation in terms of the cut-off belief, and we derive in closed form the stationary distribution of beliefs.

While in most of the analysis we consider common variance across firms, it turns out that the sorting result holds for different learning rates (noise) across firms, even if the rate of learning is slower in the high type firm. It is conceivable that with supermodularity and a learning rate no smaller in high types firms there will be positive sorting. The high type firm is both superior in the learning rate and in productive efficiency. But if high type firms learn at a sufficiently slower rate (the noise is sufficiently high), then the signal-to-noise ratio in the high type firm may well be lower. The reason why this nonetheless does not affect the learning is that the value of learning also depends on the degree of convexity of the value function (from Ito's Lemma), in addition to the signal-to-noise ratio. But by the no-deviation condition, at the cut-off belief, the degree of convexity is the same in both firms and therefore the equilibrium value of learning is the same, no matter the difference in signal-to-noise ratios. Key here is that wages are endogenous and determined competitively. That is why this property does not necessarily hold in the canonical multi-armed bandit problem.

We analyze the planner's problem and show that a planner's stationary allocation coincides with the decentralized equilibrium allocation, even if learning rates differ across different firms. This is surprising since there is a market incompleteness: wages are spot market prices only and cannot be made contingent on future realizations. It turns out that the efficiency result and proof crucially hinges on the martingale property inherent in Bayesian learning. The martingale property implies that no matter how fast workers learn, the expected beliefs about their ability will stay the same. Since under strict supermodularity, the differential in expected output between working in high and low productivity firms is monotonically increasing in the likelihood that the worker has high ability, reallocating a group of low belief workers to a better match will decrease expected outputs no matter how fast they learn.

We extend our analysis of Bayesian learning to allow for *observable* human capital accumulation. This adds realism in the sense that workers learn on the job and increase their productivity with tenure, yet we do not resort to non-Bayesian updating. Now cut-off types that characterize the equilibrium allocation depend on the degree of observable experience, and beliefs continue to follow a martingale process. The properties of our equilibrium extend to this more general human capital accumulation case.

The motivation of our analysis and the results are obviously closest related to the labor market learning literature (Jovanovic (1979), Harris and Holmström (1982), Moscarini (2005) and Papageorgiou (2009)).[3] Yet, there is a close relation to both the experimentation literature (Bolton and Harris (1999), Keller, Rady, and Cripps (2005), Strulovici (2010)) and the literature on continuous time games (Sannikov (2007), Faingold and Sannikov (2007)). Most models of learning have a finite set of players and have an absorbing state. Ours has a continuum of agents and there is learning in all states. Moreover, it is essentially a competitive model with equilibrium prices and therefore payoffs from learning are endogenous.

The idea of analyzing a matching model where the current allocation determines the future type is first explored in Anderson and Smith (2000). They find the opposite result of

[3]Papageorgiou (2009) analyzes a learning model with heterogeneity. He estimates the version of Moscarini's search model with two-sided heterogeneity. With search frictions, wage setting is non-competitive and as a result, the no-deviation condition is not imposed in addition to value matching and smooth pasting. Nonetheless, his findings provide us with realistic estimates of the labor market characteristics of our model. See also Groes, Kircher, and Manovskii (2009) for estimates of a different learning model.

ours: positive assortative matching fails even under supermodularity. They analyze a two-sided matching model of reputations with imperfect information about both matched types.[4] Our setup differs substantially, but the main difference is in the information extraction. Their agents infer the type of each of the matched partners from the realization of a *joint* signal.[5]

Another key characteristic of our model is that it is a pure Bayesian learning model where beliefs follow a martingale. In Section 2.8 we show that our result holds for Bayesian updating processes other than the Brownian motion (we extend our result to a generalized Lévy process), and we also establish that positive assortative matching can fail if the updating process is not Bayesian (this can be interpreted for example as a technology of *unobserved* human capital accumulation in addition to the information extraction).

2.2 The Model Economy

Population of Firms and Workers. The economy is populated by a unit measure of workers and a unit measure of firms. Both firms and workers are *ex ante* heterogeneous. The firm's type $y \in \{H, L\}$ represents its productivity. The type y is observable to all agents in the economy. The fraction of H type firms is π and all firms are infinitely lived. The worker ability $x \in \{H, L\}$ is not observable, both to firms and workers, i.e., information is symmetric.[6] Nonetheless, both hold a common belief about the worker type, denoted by $p \in [0, 1]$. Upon entry, a newly born worker is of type H with probability p_0 and of type L with probability $1 - p_0$. Workers die with exogenous probability δ. New workers are born at

[4]Our model is more closely related to the standard firm-worker model to which they compare their two-sided model in the discussion. There is only a one-sided inference problem in that model and they find that positive assortative matching arises for extreme beliefs $p = 0$ and 1, but conjecture it does not in the interior.

[5]The difficulty is to account for agents switching partners. Anderson and Smith (2000) resolve this by assuming symmetric learning in discrete time. Both sides of the market update in an identical fashion and under PAM their new matched partner coincides exactly with the updated type of their old partner. As a result, in a candidate PAM equilibrium there is never any switching.

[6]This substantially simplifies the problem at hand. With private signals Cripps, Ely, Mailath, and Samuelson (2008) show that with a finite signal space there will be common learning, but not necessarily with an infinite signal space as is the case in our model here.

the same rate.[7]

Preferences and Production. Workers and firms are risk-neutral and discount future payoffs at rate $r > 0$. Utility is perfectly transferable. Output is produced in pairs of one worker and one firm (x, y). Time is continuous. Positive output produced consists of a divisible consumption good and is denoted by μ_{xy}. We assume that more able workers are more productive in any firm, $\mu_{Hy} \geq \mu_{Ly}, \forall y$ and refer to it as worker monotonicity. While it is often useful, we do not in general assume firm monotonicity, which would be $\mu_{xH} \geq \mu_{xL}, \forall x$. Strict supermodularity is defined in the usual way:

$$\mu_{HH} - \mu_{LH} > \mu_{HL} - \mu_{LL}, \tag{2.1}$$

and with the opposite sign for strict submodularity. In the entire paper, we will refer to strict supermodularity when we just mention supermodularity, likewise for submodularity.

Information. Because worker ability is not observable to both the worker and the firm, parties face an information extraction problem. They observe a noisy measure of productivity, denoted by X_t. Cumulative output is assumed to be a Brownian motion with drift μ_{xy} and common variance σ^2

$$X_t = \mu_{xy} t + \sigma Z_t \tag{2.2}$$

where Z_t is a standard Wiener process and as a result, X_t is normally distributed with mean $\mu_{xy} t$ and variance $\sigma^2 t$. By Girsanov's Theorem the probability measures over the paths of two diffusion processes with the same volatility but different bounded drifts are equivalent, that is, they have the same zero-probability events. Since the volatility of a continuous-time diffusion process is effectively observable, the worker's type could be learned directly from the observed volatility if σ depends on workers' types.[8]

[7]Without death, we know the posterior belief will converge with probability one to $p = 1$ or $p = 0$. Death here actually acts as a shuffling device to guarantee a non-trivial stationary distribution of posterior beliefs.

[8]However, we can allow σ to be firm-specific. In section 2.8 we analyze the general case of firm-dependent σ_y.

Equilibrium. We consider a stationary competitive equilibrium in this economy. With two types of firms and a continuum of p's in this market, take a competitive wage schedule $w_y(p)$ as given which specifies wage for every possible type p worker working in firm y.[9] Denote by V_y the stationary discounted present value of the competitive profits for firm y. The flow profit can be written as rV_y.[10] Now we are ready to define the notion of competitive equilibrium:

Definition 2.1. *A stationary competitive equilibrium consists of a competitive wage schedule $w_y(p) = \mu_y(p) - rV_y$, where $\mu_y(p) = p\mu_{Hy} + (1-p)\mu_{Ly}$ denotes worker p's expected productivity in firm $y = H, L$ and worker p chooses the firm y with the highest discounted present value. The market clears such that the measure of workers in L firms is $1 - \pi$ and the measure of workers in H firms is π.*

2.3 Preliminaries

2.3.1 Benchmark: No Learning

Workers differ in the common beliefs p of being a high type. We shut down learning so that beliefs are invariant. This can be viewed as a special case of the learning model with the variance σ^2 going to infinity. We assume that there is no birth or death so we essentially have a static problem. Suppose without loss of generality that p is uniformly distributed on $[0, 1]$. We continue to maintain the assumption that the worker does not know her true type or that she has no private information about it. Denote $\mu_y(p) = p\mu_{Hy} + (1-p)\mu_{Ly}$ for $y = H, L$ and r as the discount rate.

[9] Bergemann and Välimäki (1996) and Felli and Harris (1996) consider a two-firm, one-worker/buyer model with strategic price setting in a world with independent arms. With ex ante heterogeneous firms and workers and correlated arms, we instead focus on competitive price setting which is closest in spirit to the Beckerian benchmark.

[10] Notice since there is no free entry, V_y need not to be zero. We could model free entry as long as in equilibrium there is a non-degenerate distribution of firm types in the economy. We consider this does not add to the insights of our model.

Under the above notion of competitive equilibrium, it is easy to verify the following claim (All of the results in this paper are in the sense of "almost surely" because we allow a zero measure of agents to behave differently):

Claim 2.1. *Under strict supermodularity, PAM is the unique (stationary) competitive equilibrium allocation: H firms match with workers $p \in [1-\pi, 1]$, L firms match with workers $p \in [0, 1-\pi)$. The opposite (NAM) holds under strict submodularity: H firms match with workers in $[0, \pi)$.*

Since there is no learning, essentially this result is identical to Becker's (1973) result, but with uncertainty. Noteworthy about this version of Becker is that even though for PAM there is supermodularity of the ex-post payoffs ($\mu_{HH} + \mu_{LL} > \mu_{HL} + \mu_{LH}$), there need not be monotonicity in expected payoffs, i.e., $\mu_H(1-\pi)$ may be smaller than $\mu_L(1-\pi)$. In fact, that will be reflected in the firm's equilibrium payoffs: $V_H \geq V_L$ if and only if $\mu_H(1-\pi) \geq \mu_L(1-\pi)$.

As in Becker, the equilibrium allocation is unique, but there may be multiple splits of the surplus. In the case of PAM, we only require at the cutoff type $\underline{p} = 1 - \pi$ that $w_H(\underline{p}) = w_L(\underline{p})$. There are multiple equilibrium payoffs if the surplus of a match between L and $p = 0$ is positive. Instead, if $\mu_L(0) = 0$,[11] there is a unique equilibrium payoff.

2.3.2 Belief Updating

In the presence of learning we can now derive the beliefs and subsequently the value functions. The posterior belief p_t that the worker has a high productivity is a sufficient statistic for the output history. Now, we can use the following well-known result: conditional on the output process $(X_t)_{t\geq 0}$, $(p_t)_{t\geq 0}$ is a martingale diffusion process. Moreover, this process can be represented as a Brownian motion. Based on the framework of our model, denote $s_y = (\mu_{Hy} - \mu_{Ly})/\sigma,\ y = H, L,\ \Sigma_y(p) = \frac{1}{2}p^2(1-p)^2 s_y^2$ and then we get:

[11] And there is limited liability, i.e., workers and firms cannot receive negative payoffs.

Lemma 2.1. *(Belief Consistency) Consider any worker who works for firm y between t_0 and t_1. Given a prior $p_{t_0} \in (0,1)$, the posterior belief $(p_t)_{t_0<t\leq t_1}$ is consistent with the output process $(X_{y,t})_{t_0<t\leq t_1}$ if and only if it satisfies*

$$dp_t = p_t(1-p_t)s_y d\bar{Z}_{y,t}$$

where

$$d\bar{Z}_{y,t} = \frac{1}{\sigma}[dX_{y,t} - (p_t\mu_{Hy} + (1-p_t)\mu_{Ly})dt].$$

The proof of this Lemma is in Faingold and Sannikov (2007) or Daley and Green (2008). The basic idea behind the proof is a combination of Bayes' rule and Ito's lemma. Given the period t posterior belief p_t and dX_t, we know the posterior belief at period $t+dt$ is:

$$p_{t+dt} = \frac{p_t \exp\{-\frac{[dX_t-\mu_{Hy}dt]^2}{2\sigma^2 dt}\}}{p_t \exp\{-\frac{[dX_t-\mu_{Hy}dt]^2}{2\sigma^2 dt}\} + (1-p_t)\exp\{-\frac{[dX_t-\mu_{Ly}dt]^2}{2\sigma^2 dt}\}}.$$

Hence,

$$dp_t = p_{t+dt} - p_t = p_t(1-p_t)\frac{\exp\{-\frac{[dX_t-\mu_{Hy}dt]^2}{2\sigma^2 dt}\} - \exp\{-\frac{[dX_t-\mu_{Ly}dt]^2}{2\sigma^2 dt}\}}{p_t \exp\{-\frac{[dX_t-\mu_{Hy}dt]^2}{2\sigma^2 dt}\} + (1-p_t)\exp\{-\frac{[dX_t-\mu_{Ly}dt]^2}{2\sigma^2 dt}\}}.$$

Apply Ito's Lemma and we obtain the above result.

Lemma 2.1 establishes that dp depends on three elements: $p(1-p)$, which peaks at $1/2$; the signal-to-noise ratio of output, $s_y = (\mu_{Hy} - \mu_{Ly})/\sigma$ and $d\bar{Z}_y$, the normalized difference between realized and unconditionally expected flow output, which is a standard Wiener process with respect to the filtration $\{X_{y,t}\}$. Obviously, beliefs move faster the more uncertainty about worker's quality (p close to $1/2$); the less variation in the output process (smaller σ) and the larger the productivity difference (higher $\mu_{Hy} - \mu_{Ly}$).

Learning considerations will change the benchmark results. Moreover, supermodularity not only affects the value of the static output created as in the standard Beckerian model, but it also has dynamic effect by changing the speed of learning. For example, under supermodularity ($\mu_{HH} - \mu_{HL} > \mu_{LH} - \mu_{LL}$), the learning speed is faster in the high type firm, which is especially significant for p close to $1/2$. Intuitively speaking, learning makes it more attractive to match with a high type firm even though statically it is better for her to match with a low type firm without learning.

2.3.3 Value Functions

Given the wage schedule, each worker is facing a two-armed bandit problem. We restrict the workers' strategies to be Markovian:

$$a : [0,1] \to \{H, L\}.$$

The value function of a type p worker can be written as:

$$W(p) = \sup_{a:[0,1]\to\{H,L\}} \left\{ \mathbb{E} \int_{t=0}^{\infty} e^{-(r+\delta)t} w_{a_t}(p_t) dt \right\}$$
$$s.t. dp_t = p_t(1-p_t) s_{a_t} d\bar{Z}_{a_t,t} \quad \text{and} \quad a_t \triangleq a(p_t).$$

Denote $W_y(p)$ to be the value function of a worker with posterior in a neighborhood of p optimally choosing firm y.

The value function $W_y(p)$ is given by[12]:

$$rW_y(p) = \mu_y(p) - V_y + \Sigma_y(p) W_y^{''}(p) - \delta W_y(p), \tag{2.3}$$

from Ito's Lemma. The term $\mu_y(p) - V_y$ is equal to the flow wage payoff and corresponds to the deterministic component of the diffusion $X_{y,t}$, and the term $\Sigma_y(p)W_y^{''}(p)$ is the second-order term from the transformation W of the diffusion process $X_{y,t}$. First-order and all higher-order terms vanish as the time interval shrinks to zero. The general solution to this differential equation is:

$$W_y(p) = \frac{\mu_y(p) - V_y}{r+\delta} + k_{y1} p^{1-\alpha_y}(1-p)^{\alpha_y} + k_{y2} p^{\alpha_y}(1-p)^{1-\alpha_y}, \tag{2.4}$$

[12]Note that we critically need the assumption that the worker does not have any private information about his type. If this assumption is violated, the worker's value functions could not be written like this.

where

$$\alpha_y = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{2(r+\delta)}{s_y^2}} \geq 1.$$

First notice that the boundedness of the value function implies that if 0 is included in the domain, then $k_{y1} = 0$ and if 1 is included in the domain, then $k_{y2} = 0$. If not, with $\alpha_y > 1$ the value of W shoots off to infinity. Second, $\Sigma_y(p) W_y''(p)$ is the value of learning and this is an option value in the sense that the worker has the choice to change his job as he learns his type p. It is easy to verify that this value is zero if the worker never changes his job.[13] From the Martingale property of the Brownian motion, at any p the expected value of p in the next time interval is equal to p. There is as much good news as bad news to be expected in the next period. It is the option value of switching to a *more suitable* match that generates the value of learning. Equation (2.4) implies that this option value can be decomposed into two parts: $k_{y1} p^{1-\alpha_y}(1-p)^{\alpha_y}$ ($k_{y2} p^{\alpha_y}(1-p)^{1-\alpha_y}$) denotes the option value of switching to a *more suitable* match when p goes down (up). The option value $k_{y1} p^{1-\alpha_y}(1-p)^{\alpha_y}$ ($k_{y2} p^{\alpha_y}(1-p)^{1-\alpha_y}$) must be zero if 0 (1) is included since no switch happens as p goes down (up).

2.4 Analysis and Results

2.4.1 Characterization of the Equilibrium Allocation

Now consider any candidate stationary equilibrium where a type p worker switches from firm y to y'. Since the worker is essentially facing a two-armed bandit problem given the wage schedule, optimality in stopping time requires the value-matching condition (the worker gets the same value at the cutoff) and the smooth-pasting condition (the marginal of both value functions is identical) (see Dixit (1993)). For example, if for $p \in [p_1, p_2)$, the worker works in the low type firm and for $p \in [p_2, p_3)$, the worker works in the high type firm, then we

[13] In that case, p can take both the values 0 and 1. So the boundedness of the value function requires that both k_{y1} and k_{y2} are zero and hence $W_y''(p) = 0$ for every p.

must have:[14]

$$W_L(p_2) = W_H(p_2) \quad \text{and} \quad W_L{}'(p_2) = W_H{}'(p_2). \tag{2.5}$$

Notice that workers are price takers. As a result, there is no strategic interaction between players where equilibrium solves for the fixed point of individual strategies. It is also important to point out that both the value-matching condition and the smooth-pasting condition are on-equilibrium path conditions. They have nothing to do with the off-equilibrium path (i.e., instead of accepting offers from low type firms, workers with $p \in [p_1, p_2)$ are tempted to accept offers from high type firms). In the following lemmas we characterize the value functions establishing convexity and monotonicty:

Lemma 2.2. *The equilibrium value functions W_y are strictly convex for $p \in (0, 1)$.*

Proof. In Appendix. □

The intuition for this Lemma is the following. Preferences and output are linear in p, and the option value of learning is strictly positive, hence the value function with the option of learning is convex. To see this, observe that since the measure of both types of firms are strictly positive, market clearing requires that workers with some p's will be employed by high type firms while workers with other p's will be employed by low type firms. This implies that some worker has to change jobs at some point and the option value of learning $\Sigma_y(p)W_y{}''(p)$ is strictly positive. Hence we have $W_y''(p) > 0$, for all $p \in (0, 1)$ since $\Sigma_y(p) > 0$. On the other hand, when $p = 0$ or 1, the posterior belief will always stay at 0 or 1 by Bayes' rule such that learning never happens. It is easy to verify that $W_y''(p) = 0$ for $p = 0$ or 1.

Given the strict convexity of equilibrium value functions and the smooth pasting condition, we can immediately derive the following Lemma:

[14]We slightly abuse notation hers since W_L is not defined on p_2. A more precise way of writing the equations is $W_L(p_2+) = W_H(p_2)$ and $W_L{}'(p_2+) = W_H{}'(p_2)$. In what follows, we will continue to use the expression in the text in order to economize on notation.

Lemma 2.3. *The equilibrium value functions W_y are strictly increasing.*

Proof. In Appendix. $\square$

One important implication is that if we define $\mathcal{W}(p)$ as the envelope of all equilibrium value functions $W_y(p)$, then this envelope function $\mathcal{W}(p)$ is continuous, strictly increasing and strictly convex for $p \in (0,1)$. Suppose workers with $p \in [0, \underline{p})$ are employed by type y firm and workers with $p \in (\bar{p}, 1]$ are employed by type $-y$ firm. Then we should have: $W_y'(0) = \frac{\mu_{Hy} - \mu_{Ly}}{r+\delta} < W_{-y}'(1) = \frac{\mu^{H,-y} - \mu^{L,-y}}{r+\delta}$. This gives us another result:

Lemma 2.4. *Under supermodularity, in any equilibrium $p = 0$ workers match with L firms; $p = 1$ workers match with H firms. The opposite under strict submodularity. Moreover,*

$$\frac{\min(\Delta_H, \Delta_L)}{r+\delta} < W'(p) < \frac{\max(\Delta_H, \Delta_L)}{r+\delta},$$

where $\Delta_H = \mu_{HH} - \mu_{LH}$ and $\Delta_L = \mu_{HL} - \mu_{LL}$.

Intuitively this result is best understood by using the standard sorting argument from Becker (1973). At $p = 0$ and $p = 1$ there is no value of learning. As a result, there the value function can be interpreted as being determined by the no-learning allocation.

The properties derived above are mainly concerned with on-equilibrium path behavior. We also need to specify what happens in the event of deviations and consider behavior off-equilibrium path. We contemplate the equivalence of a one-shot deviation in continuous time because we think of the continuum as an idealization of discrete time. This amounts to a worker playing the deviant action over an interval $[t, t + dt)$ according to the belief p at time t, and considering the limit as $dt \to 0$.[15] This is very important because it allows us to derive the value function for deviation. On the contrary, if the deviation only takes place at a single point in time t, then the value function for deviation is essentially the same as

[15]This notion is also implicitly used in Proposition 2 of Sannikov (2007), and also in Cohen and Solan (2009) who consider deviations from Markovian strategies in bandit problems.

the one without deviation because no information will be extracted from just a single time point.

The next Lemma establishes that if we consider off-the-equilibrium path deviations, we actually derive one additional condition, which we call the *no-deviation* condition.

Lemma 2.5. *To deter possible deviations, a necessary condition is:*

$$W_H''(\underline{p}) = W_L''(\underline{p}) \quad \textit{(No-deviation condition)} \tag{2.6}$$

for any possible cutoff $\underline{p}$.

Proof. Without loss of generality, we assume that on the equilibrium path, a worker in a neighborhood right of $\underline{p}$ accepts offers from H firms (say, $p \in (\underline{p}, \bar{p})$) and a worker in a neighborhood left of $\underline{p}$ accepts offers from L firms. Consider one possible one-shot deviation: at time t, a $p > \underline{p}$ worker chooses a low type firm for dt length of time and then switches back. On the equilibrium path, the value function is defined as before (from Hamilton-Jacobi-Bellman equation):

$$(r+\delta)W(p) = (r+\delta)W_H(p) = w_H(p) + \Sigma_H(p)W_H''(p).$$

The deviator's new value could be written as:

$$\tilde{W}_L(p) = \mathbb{E}\left\{\int_t^{t+dt} e^{-(r+\delta)(s-t)} w_L(p_s)ds + e^{-(r+\delta)dt} W(p_{t+dt})\right\}. \tag{2.7}$$

Potentially, p_{t+dt} can take any value between 0 and 1. We have to show that as dt becomes very small, almost surely, p_{t+dt} will be close to p such that it is in the region where the worker will still accept offers from high type firms: $\Pr(p_{t+dt} \notin (\underline{p}, \bar{p})) = o(dt)$.[16]

[16]Since the deviator's belief updating follows a Brownian motion: $dp_t = s_L p(1-p)d\bar{Z}_{L,t}$, the probability that a worker $p > \underline{p}$ will have belief $p_{t+dt} \le \underline{p}$ is given by $\Phi\left(\frac{\underline{p}-p}{s_L p(1-p)\sqrt{dt}}\right)$, where $\Phi(\cdot)$ is the cumulative distribution function for a standard normal distribution. Apply L'Hopital's rule and it is straightforward to see that

$$\lim_{dt\to 0} \frac{\Phi\left(\frac{\underline{p}-p}{s_L p(1-p)\sqrt{dt}}\right)}{dt} = 0.$$

Use the same logic and it is easy to see that $\Pr(p_{t+dt} > \bar{p}) = o(dt)$.

Notice that for any $dt > 0$,

$$\begin{aligned} W_H(p) &> \tilde{W}_L(p) > \mathbb{E}\left\{\int_t^{t+dt} e^{-(r+\delta)(s-t)} w_L(p_s) ds\right\} \\ &+ \mathbb{E}e^{-(r+\delta)dt}\left[W_H(p_{t+dt})(1-\Pr(p_{t+dt} \notin (\underline{p},\bar{p}))) + \Pr(p_{t+dt} \notin (\underline{p},\bar{p}))W(0)\right]. \end{aligned} \tag{2.8}$$

The first inequality comes from the fact that there should be no profitable deviation. The second inequality is true because we replace the value for $p_{t+dt} \notin (\underline{p},\bar{p})$ with the lowest value $W(0)$ ($W(\cdot)$ is an increasing function by Lemma 2.3). From Ito's Lemma, we can get for the deviator:

$$\mathbb{E}W_H(p_{t+dt}) = W_H(p) + \Sigma_L(p)W_H''(p)dt + o(dt).$$

For any $dt > 0$, the no deviation condition implicit in equation (2.8) implies:

$$\begin{aligned} &\frac{\mathbb{E}\{\int_t^{t+dt} e^{-(r+\delta)(s-t)} w_L(p_s) ds\}}{dt} \\ + &\frac{\mathbb{E}\left\{e^{-(r+\delta)dt}[W_H(p_{t+dt})(1-\Pr(p_{t+dt} \notin (\underline{p},\bar{p}))) + \Pr(p_{t+dt} \notin (\underline{p},\bar{p}))W(0)]\right\} - W_H(p)}{dt} < 0. \end{aligned}$$

Let $dt \to 0$ and first, it follows immediately that:

$$\lim_{dt\to 0} \frac{\mathbb{E}\left\{\int_t^{t+dt} e^{-(r+\delta)(s-t)} w_L(p_s) ds\right\}}{dt} = w_L(p).$$

Second, as proved earlier,

$$\lim_{dt\to 0} \frac{\Pr(p_{t+dt} \notin (\underline{p},\bar{p}))}{dt} = 0.$$

Finally,

$$\begin{aligned} &\lim_{dt\to 0} \frac{\mathbb{E}\left\{e^{-(r+\delta)dt} W_H(p_{t+dt})(1-\Pr(p_{t+dt} \notin (\underline{p},\bar{p})))\right\} - W_H(p)}{dt} \\ = &\lim_{dt\to 0} \frac{(e^{-(r+\delta)dt}-1)W_H(p) + \Sigma_L(p)W_H''(p)dt + o(dt)}{dt} = \Sigma_L(p)W_H''(p) - (r+\delta)W_H(p). \end{aligned}$$

Therefore, the necessary condition such that a $p > \underline{p}$ worker has no incentive to deviate can be written as:

$$w_L(p) + \Sigma_L(p)W_H''(p) - (r+\delta)W_H(p) = w_L(p) + \Sigma_L(p)W_H''(p) - w_H(p) - \Sigma_H(p)W_H''(p) < 0. \tag{2.9}$$

The above inequality must hold for any $p \in (\underline{p}, \bar{p})$. Let $p \to \underline{p}$ and we have:[17]

$$w_L(\underline{p}) - w_H(\underline{p}) + [\Sigma_L(\underline{p}) - \Sigma_H(\underline{p})]W_H''(\underline{p}) \leq 0$$
$$\Rightarrow w_L(\underline{p}) + \Sigma_L(\underline{p})W_L''(\underline{p}) - (w_H(\underline{p}) + \Sigma_H(\underline{p})W_H''(\underline{p})) + (W_H''(\underline{p}) - W_L''(\underline{p}))\Sigma_L(\underline{p}) \leq 0$$
$$\Rightarrow W_H''(\underline{p}) \leq W_L''(\underline{p}). \tag{2.10}$$

Similarly, we can consider another possible one-shot deviation: a $p < \underline{p}$ worker matches with a high type firm for dt and then switches back. The same logic establishes that to deter such deviation, it must be the case that:

$$w_H(p) - w_L(p) + [\Sigma_H(p) - \Sigma_L(p)]W_L''(p) < 0 \tag{2.11}$$

for any $p < \underline{p}$. As p goes to $\underline{p}$, we should have:

$$w_H(\underline{p}) - w_L(\underline{p}) + [\Sigma_H(\underline{p}) - \Sigma_L(\underline{p})]W_L''(\underline{p}) \leq 0 \Rightarrow W_H''(\underline{p}) \geq W_L''(\underline{p}). \tag{2.12}$$

(2.10) and (2.12) imply that $W_H''(\underline{p}) = W_L''(\underline{p})$. □

This no-deviation condition is quite unique for the two-armed bandit problem. This condition is absent in an one-armed bandit problem. Most of the models in the literature on continuous time learning models (Jovanovic (1979) and Moscarini (2005)) and continuous time games (see amongst others, Sannikov (2008)) are essentially investigating a one-armed bandit problem. There, we can directly look at equilibria in cutoff strategies. In the one-armed bandit problems, the safe arm essentially is an absorbing state so we only need to worry about the potential deviation from the risky arm to the safe arm.[18] Then the no-deviation condition becomes $W_H''(\underline{p}) \geq W_L''(\underline{p}) = 0$ but this is already implied by the

[17] As p goes to $\underline{p}+$, notice that $w_L(\underline{p}-) = w_L(\underline{p}+), \Sigma_L(\underline{p}-) = \Sigma_L(\underline{p}+)$. Hence, we will have: $w_L(\underline{p}-) + \Sigma_L(\underline{p}-)W_L''(\underline{p}-) - (w_H(\underline{p}+) + \Sigma_H(\underline{p}+)W_H''(\underline{p}+)) + (W_H''(\underline{p}+) - W_L''(\underline{p}-))\Sigma_L(\underline{p}-) \leq 0$.

[18] For example, in our model assume $\mu_{HL} = \mu_{LL}$ and the return in the low type firm is deterministic.

convexity property.[19]

We provide some intuition for the no-deviation condition. By assuming Sequential Rationality, i.e., the equilibrium is robust to a one-shot deviation, we basically impose that the equilibrium wage is self-enforcing. There is no commitment to future realizations of X_t and therefore of future beliefs p. Now we can interpret W'' as the marginal value of learning: W' is the marginal change of W with respect to the posterior p, and learning changes p and is therefore quantified by the change in W' which is W''. The condition states that there is no deviation if the marginal value of learning at $\underline{p}$ is the same in both firms.

Now in our two-armed bandit problem, we first need to answer the question whether there exist non-cutoff stationary equilibria, i.e., a worker with $p \in [p_1, p_2)$ accepts the offer from a high type firm, with $p \in [p_2, p_3)$ accepts the offer from a low type firm and with $p \in [p_3, p_4)$ accepts the offer from a high type firm again. Surprisingly, Lemmas 2.2–2.5 imply that all possible stationary competitive equilibria must be in cutoff strategies. The next theorem therefore establishes uniqueness and sorting under supermodularity. It does not shown existence yet, which we do in Theorem 2.3 below.

Theorem 2.1. *If an equilibrium exists, PAM is the unique stationary competitive equilibrium allocation under strict supermodularity. Likewise for NAM under strict submodularity.*

To prove this theorem, we only need to prove the following Claim:

Claim 2.2. *Under strict supermodularity, it is impossible to have $p_1 < p_2$ and equilibrium value functions W_H (for $p \in [p_1, p_2]$), W_{L1} (for $p < p_1$), W_{L2} (for $p > p_2$) such that:*

$$W_H(p_1) = W_{L1}(p_1) \quad \text{and} \quad W''_H(p_1) = W''_{L1}(p_1)$$

$$W_H(p_2) = W_{L2}(p_2) \quad \text{and} \quad W''_H(p_2) = W''_{L2}(p_2)$$

[19] In a model of option pricing by Dumas (1991), there does exist a condition on the second derivative called the "super contact" condition, which is of a very different nature. It arises as the optimal solution to the option pricing problem with proportional cost. More discussions about this no-deviation condition can be found in Eeckhout and Weng (2010)

are satisfied simultaneously.

Under strict submodularity, it is impossible to have $p_1 < p_2$ and equilibrium value functions W_L (for $p \in [p_1, p_2]$), W_{H1} (for $p < p_1$), W_{H2} (for $p > p_2$) such that:

$$W_L(p_1) = W_{H1}(p_1) \quad \text{and} \quad W''_L(p_1) = W''_{H1}(p_1)$$

$$W_L(p_1) = W_{H2}(p_2) \quad \text{and} \quad W''_L(p_2) = W''_{H2}(p_2)$$

are satisfied simultaneously.

Proof. In Appendix. □

This result states that it is not benefial for a worker of type p to learn in the high type firm H in the middle as long as there there are still types p on both sides who work in the low type firms. Given the above claim, it is easy to prove the theorem:

Proof. Under supermodularity, by Lemma 2.5, workers with sufficiently low p's will accept a low type firm's wage offer and workers with sufficiently high p's will accept a high type firm's offer. But Claim 2.2 implies it is impossible to have worker first accept low type firm's offer, then accept high type firm's offer and finally accept low type firm's offer again. Hence, we must have some cutoff $\underline{p}$ such that $p < \underline{p}$ will accept low type firm's offer and $p > \underline{p}$ will accept high type firm's offer. This is exactly a PAM allocation. Use the same logic, NAM is the only possible stationary competitive equilibrium allocation under strict submodularity. □

Before we turn to the equilibrium distribution, we show that the no-deviation condition in Lemma 2.5 is not just necessary but also sufficient under strict supermodularity:

Lemma 2.6. *Under strict supermodularity, $W''_H(\underline{p}) = W''_L(\underline{p})$ implies that no deviation will happen for the PAM equilibrium allocation.*

Proof. In Appendix. □

2.4.2 The Equilibrium Distribution

The previous section shows that under strict supermodularity (submodularity), PAM (NAM) is the unique candidate stationary competitive equilibrium allocation. Note that this doesn't necessarily mean the equilibrium exists. We still need to construct such an equilibrium. To do that, we assume strict supermodularity and worker and firm monotonicity: ($\mu_{HH} > \mu_{HL}$ and $\mu_{LH} > \mu_{LL}$).[20] Now consider a strictly positive assortative matching equilibrium such that workers with beliefs less than $\underline{p}$ will choose L firms and workers with beliefs higher than $\underline{p}$ will choose H firms. From equation (2.4) we hence have $k_{L1} = 0$ and $k_{L2} > 0$ for $y = L$ and $k_{H2} = 0$ and $k_{H1} > 0$ for $y = H$. Let $k_L = k_{L2}$, $k_H = k_{H1}$ and worker's value functions become:

$$W_L(p) = \frac{w_L(p)}{r+\delta} + k_L p^{\alpha_L}(1-p)^{1-\alpha_L} \tag{2.13}$$

and

$$W_H(p) = \frac{w_H(p)}{r+\delta} + k_H p^{1-\alpha_H}(1-p)^{\alpha_H}, \tag{2.14}$$

where

$$\alpha_y = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{2(r+\delta)}{s_y^2}} \geq 1.$$

To discuss market clearing conditions, we need to consider the ergodic distribution of p's. From the Fokker-Planck (Kolmogorov forward) equation, the stationary and ergodic density f_y should satisfy the following differential equation:

$$0 = \frac{df_y(p)}{dt} = \frac{d^2}{dp^2}[\Sigma_y(p) f_y(p)] - \delta f_y(p). \tag{2.15}$$

---

[20]Monotonicity is just to help us find one particular way to divide the surplus. The whole construction of equilibrium also goes through if we do not make this assumption.

The general solution to this differential equation is (see also Moscarini (2005)):[21]

$$f_y(p) = [f_{y0}p^{\gamma_{y1}}(1-p)^{\gamma_{y2}} + f_{y1}(1-p)^{\gamma_{y1}}p^{\gamma_{y2}}] \tag{2.16}$$

where

$$\gamma_{y1} = -\frac{3}{2} + \sqrt{\frac{1}{4} + \frac{2\delta}{s_y^2}} > -1$$

and

$$\gamma_{y2} = -\frac{3}{2} - \sqrt{\frac{1}{4} + \frac{2\delta}{s_y^2}} < -2.$$

First, the integrability of f_y requires that $f_{y1} = 0$ if 0 is included in the domain and $f_{y0} = 0$ if 1 is included in the domain. Second, the Fokker-Planck (Kolmogorov forward) equation is only valid for $p \neq p_0$. Since there is a flow in of new workers, for $p = p_0$ we should have a kink in the density function. This also raises the issue of the relative position between p_0 and $\underline{p}$. We first consider the case where $\underline{p} < p_0$. We then derive in abbreviated format the result when $\underline{p} > p_0$.

Given any $p_0 \in (0,1)$, if $\underline{p} < p_0$, then the density functions are:

$$f_H(p) = [f_{H0}p^{\gamma_{H1}}(1-p)^{\gamma_{H2}} + f_{H1}(1-p)^{\gamma_{H1}}p^{\gamma_{H2}}]\mathbb{I}(\underline{p} < p \leq p_0) + f_{H2}(1-p)^{\gamma_{H1}}p^{\gamma_{H2}}\mathbb{I}(p > p_0) \tag{2.17}$$

and

$$f_L(p) = f_{L0}p^{\gamma_{L1}}(1-p)^{\gamma_{L2}}. \tag{2.18}$$

The density functions are subject to the following boundary conditions. The derivations of these boundary conditions are shown in the appendix. First, once the posterior belief reaches the equilibrium separation point $\underline{p}$, we should have the cutoff condition:

$$\Sigma_H(\underline{p}+)f_H(\underline{p}+) = \Sigma_L(\underline{p}-)f_L(\underline{p}-). \tag{2.19}$$

[21]Here the assumption that there is no heterogeneity in the prior p_0 substantially simplifies the solution to this differential equation. While there is no solution for a general distribution of priors, we have been able to solve the stationary distribution if the priors are drawn from a beta distribution. See also Papageorgiou (2009).

This condition guarantees that the flow speed of agents who cross $\underline{p}$ from below is equal to the flow speed of agents who cross from above. The implication is that since the speed from above Σ_H is larger than Σ_L, the densities are not continuous: $f_H(\underline{p}+) < f_L(\underline{p}-)$. It is worth comparing this condition to the standard condition when there is an absorbing state (Cox and Miller (1965), Dixit (1993), and Moscarini (2005)). In the case with only one Brownian motion and an absorbing state, what is required is $\Sigma(\underline{p}+)f(\underline{p}+) = 0$ because the probability of absorption in a time interval dt must equal the flow-in speed of the Brownian motion which is proportional to $\sqrt{dt}$ (see Cox and Miller (1965, p.220)).

Second, total flows in and out of the high type firms must balance:

$$\Sigma_H(p_0)[f'_H(p_0-) - f'_H(p_0+)] = \delta\pi + \frac{d}{dp}[\Sigma_H(p)f_H(p)]|_{\underline{p}+}.$$

The left-hand side of the above equation is the total inflow into high type firms, which are new workers who enter into this economy. The right-hand side of the above equation is the total outflows from the high type firms, which include workers who reach $\underline{p}$ and transfer to low type firms and workers who are hit by the death shock. We manage to show that this equation will further imply:

$$\frac{d}{dp}[\Sigma_L(p)f_L(p)]|_{\underline{p}-} = \frac{d}{dp}[\Sigma_H(p)f_H(p)]|_{\underline{p}+}$$

Third, the density function has to be continuous at p_0:

$$f_H(p_0-) = f_H(p_0+).$$

It is customary to impose this condition as it approximates entry from a non-degenerate distribution instead of entry of identical types p_0.

Finally, usual market clearing conditions apply:

$$\int_{\underline{p}}^1 f_H(p)dp = \pi \quad \text{and} \quad \int_0^{\underline{p}} f_L(p)dp = 1 - \pi.$$

In summary, when $\underline{p} < p_0$, the equilibrium is characterized by a system of eight equations with nine unknowns $(V_L, V_H, k_L, k_H, \underline{p}, f_{H0}, f_{H1}, f_{H2}, f_{L0})$:[22]

$$W_H(\underline{p}) = W_L(\underline{p}) \quad \text{(Value-matching condition)} \tag{2.20}$$

$$W_H'(\underline{p}) = W_L'(\underline{p}) \quad \text{(Smooth-pasting condition)} \tag{2.21}$$

$$W_H''(\underline{p}) = W_L''(\underline{p}) \quad \text{(No-deviation condition)} \tag{2.22}$$

$$\Sigma_H(\underline{p}+) f_H(\underline{p}+) = \Sigma_L(\underline{p}-) f_L(\underline{p}-) \quad \text{(Boundary condition)} \tag{2.23}$$

$$\int_{\underline{p}}^1 f_H(p)dp = \pi \quad \text{(Market clearing } H\text{)} \tag{2.24}$$

$$\int_0^{\underline{p}} f_L(p)dp = 1 - \pi \quad \text{(Market clearing } L\text{)} \tag{2.25}$$

$$\frac{d}{dp}[\Sigma_L(p) f_L(p)]|_{\underline{p}-} = \frac{d}{dp}[\Sigma_H(p) f_H(p)]|_{\underline{p}+} \quad \text{(Flow equation at } \underline{p}\text{)} \tag{2.26}$$

$$f_H(p_0-) = f_H(p_0+) \quad \text{(Continuous density at } p_0\text{)} \tag{2.27}$$

Fortunately, Equations (2.23)–(2.27) can be solved separately from Equations (2.20)–(2.22). In other words, the procedure of solving this system of equation could be: first we solve $\underline{p}$ jointly with $f_{H0}, f_{H1}, f_{H2}, f_{L0}$ from Equations (2.23)–(2.27) and then we plug $\underline{p}$ into Equations (2.20)–(2.22) to pin down other unknowns.

Proposition 2.1. *Equations (2.23)-(2.27) imply $\underline{p} < p_0$ if and only if:*

$$\left(\frac{p_0}{1-p_0}\right)^{\gamma_{H1}-\gamma_{L2}} \frac{\delta/s_H^2}{\delta/s_L^2} \frac{\int_{p_0}^1 p^{\gamma_{H2}}(1-p)^{\gamma_{H1}}dp}{\int_0^{p_0} p^{\gamma_{L1}}(1-p)^{\gamma_{L2}}dp} < \frac{\pi}{1-\pi}. \tag{2.28}$$

Moreover, if such $\underline{p}$ exists, it must be unique.

Proof. In Appendix. □

[22] Observe that with more unknowns than variables, the solution to our system is indeterminate. In fact, there are potentially a continuum of wages that can be supported in equilibrium, though the allocation will be unique. This indeterminacy is as in Becker: the allocation is unique, but there may be multiple ways to split the surplus. In all that follows, when we use the term uniqueness of equilibrium, we refer to the allocation, not to the wages.

The proof of Proposition 2.1 is quite straightforward. The idea of the proof is the following: since we have 5 equations with five unknowns, we can first express $f_{H0}, f_{H1}, f_{H2}, f_{L0}$ as functions of $\underline{p}$ and then use the last equation to pin down $\underline{p}$.

The existence and uniqueness of the solution to the system require that $f_{H0}, f_{H1}, f_{H2}, f_{L0}$ change monotonically with $\underline{p}$. Fortunately, this is the case as shown in the appendix. The monotonicity guarantees that if a solution exists, it must be unique. Furthermore, it enables us to only check the boundaries when determining whether a solution exists. Equation (2.28) given in the Proposition is thus derived.

In the second case, $\underline{p} \geq p_0$. Given any $p_0 \in (0,1)$, if $\underline{p} \geq p_0$, then the density functions are:

$$f_L(p) = f_{L0} p^{\gamma_{L1}} (1-p)^{\gamma_{L2}} \mathbb{I}(p < p_0) + [f_{L1} p^{\gamma_{L1}} (1-p)^{\gamma_{L2}} + f_{L2} (1-p)^{\gamma_{L1}} p^{\gamma_{L2}}] \mathbb{I}(p_0 \leq p \leq \underline{p}) \quad (2.29)$$

and

$$f_H(p) = f_{H0} (1-p)^{\gamma_{H1}} p^{\gamma_{H2}}. \quad (2.30)$$

Then the system of equations to determine the equilibrium is:

$$W_H(\underline{p}) = W_L(\underline{p}) \qquad \text{(Value-matching)} \quad (2.31)$$

$$W_H'(\underline{p}) = W_L'(\underline{p}) \qquad \text{(Smooth-pasting)} \quad (2.32)$$

$$W_H''(\underline{p}) = W_L''(\underline{p}) \qquad \text{(No-deviation)} \quad (2.33)$$

$$\Sigma_H(\underline{p}+) f_H(\underline{p}+) = \Sigma_L(\underline{p}-) f_L(\underline{p}-) \qquad \text{(Boundary condition)} \quad (2.34)$$

$$\int_{\underline{p}}^1 f_H(p) dp = \pi \qquad \text{(Market clearing } H\text{)} \quad (2.35)$$

$$\int_0^{\underline{p}} f_L(p) dp = 1 - \pi \qquad \text{(Market clearing } L\text{)} \quad (2.36)$$

$$\frac{d}{dp}[\Sigma_L(p) f_L(p)]|_{\underline{p}-} = \frac{d}{dp}[\Sigma_H(p) f_H(p)]|_{\underline{p}+} \qquad \text{(Flow equation at } \underline{p}\text{)} \quad (2.37)$$

$$f_L(p_0-) = f_L(p_0+) \qquad \text{(Continuous density at } p_0\text{)} \quad (2.38)$$

Based on the above equations, we can prove the following Proposition, the counterpart to Proposition 2.1, in a similar fashion:

Proposition 2.2. *Equations (2.34)-(2.38) imply $\underline{p} \geq p_0$ if and only if:*

$$(\frac{p_0}{1-p_0})^{\gamma_{H1}-\gamma_{L2}} \frac{\delta/s_H^2}{\delta/s_L^2} \frac{\int_{p_0}^1 p^{\gamma_{H2}}(1-p)^{\gamma_{H1}}dp}{\int_0^{p_0} p^{\gamma_{L1}}(1-p)^{\gamma_{L2}}dp} \geq \frac{\pi}{1-\pi}. \tag{2.39}$$

Moreover, if such $\underline{p}$ exists, it must be unique.

The idea for the proof of Proposition 2 is exactly the same as that for the proof of Proposition 1 and the proof is also shown in the appendix. Propositions 2.1 and 2.2 together provide the following existence and uniqueness result:

Theorem 2.2. *Under strict supermodularity, for any pair $(p_0, \pi) \in (0,1)^2$, there exists a unique PAM cutoff $\underline{p}$. Moreover, $\underline{p} < p_0$ if and only if:*

$$(\frac{p_0}{1-p_0})^{\gamma_{H1}-\gamma_{L2}} \frac{\delta/s_H^2}{\delta/s_L^2} \frac{\int_{p_0}^1 p^{\gamma_{H2}}(1-p)^{\gamma_{H1}}dp}{\int_0^{p_0} p^{\gamma_{L1}}(1-p)^{\gamma_{L2}}dp} < \frac{\pi}{1-\pi}. \tag{2.40}$$

One of the nice properties about Equation (2.40) is that the whole equation only depends on p_0, π, δ/s_H^2 and δ/s_L^2. This provides a feasible way to compute $\underline{p}$. Given p_0, π, δ/s_H^2 and δ/s_L^2, we first need to decide the sign of

$$(\frac{p_0}{1-p_0})^{\gamma_{H1}-\gamma_{L2}} \frac{\delta/s_H^2}{\delta/s_L^2} \frac{\int_{p_0}^1 p^{\gamma_{H2}}(1-p)^{\gamma_{H1}}dp}{\int_0^{p_0} p^{\gamma_{L1}}(1-p)^{\gamma_{L2}}dp} - \frac{\pi}{1-\pi}.$$

If this sign is negative, then we know that $\underline{p}$ is smaller than p_0 and we can use the system of equations in the first case to figure out $\underline{p}$. On the contrary, if this sign is not negative, then we know that $\underline{p}$ is larger than p_0 and we can use the system of equations in the second case to compute $\underline{p}$. This turns out to be a convenient way to determine the equilibrium cutoff numerically.

Before presenting the numerical results, we have a simple theoretical comparative static result:

Corollary 2.1. *$\underline{p}$ is strictly increasing in p_0 and decreasing in π.*

This corollary is proved in the appendix. But the intuition is quite straightforward: decreasing in π means there are more low type firms in the economy and hence $\underline{p}$ has to

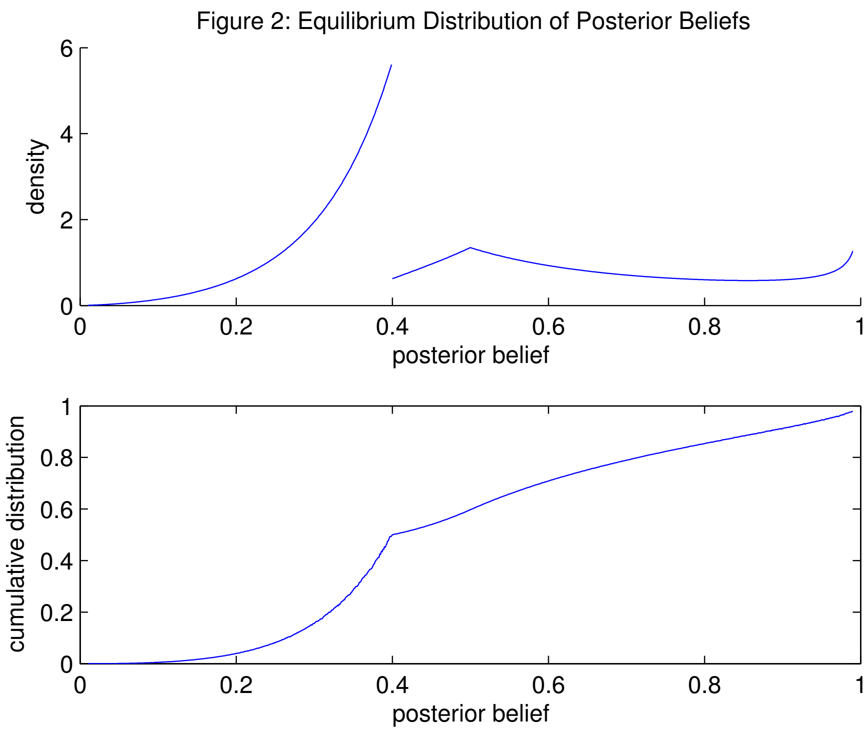

Figure 2.1: Equilibrium Distribution of Posterior beliefs.

become larger such that more workers are matched with low type firms; increasing in p_0 means the overall quality of the workers is becoming better in the economy and $\underline{p}$ has to go up to make sure that low type firms are also matched with better workers.

Mathematically, it is not easy to derive comparative statics between $\underline{p}$ and δ/s_H^2 or δ/s_L^2. But intuitively speaking, as s_L increases, the degree of supermodularity will be reduced while the speed of learning in low type firms will increase. Both of these factors make the low type firms more attractive and hence $\underline{p}$ should increase in s_L. On the other hand, as s_H becomes higher, both the degree of supermodularity and the speed of learning in high type firms will go up, which will lead to a reduction in $\underline{p}$.

Figure 2.1 plots the stationary distribution of beliefs p, for the case of PAM and with parameter values: $s_H = 0.15, s_L = 0.05, p_0 = 0.5, \pi = 0.5, \delta = 0.01$.

2.4.3 Equilibrium Analysis: Value Functions

Theorem 2.2 implies that under strict supermodularity, the PAM cutoff $\underline{p}$ can be uniquely determined. But given this $\underline{p}$, we still have the following conditions to satisfy:

$$W_H(\underline{p}) = W_L(\underline{p}) \qquad \text{(Value-matching condition)} \tag{2.41}$$

$$W'_H(\underline{p}) = W'_L(\underline{p}) \qquad \text{(Smooth-pasting condition)} \tag{2.42}$$

$$W''_H(\underline{p}) = W''_L(\underline{p}) \qquad \text{(No-deviation condition)} \tag{2.43}$$

Equations (2.41)-(2.43) are three equations for four unknowns. The equilibrium is indeterminate in the sense that although the allocation $\underline{p}$ is unique, there could be multiple ways to divide the surplus. To make the system determinate, we assume firm monotonicity and set $\mu_{LL} = 0$. Then limited liability requires that $w_L(0)$ has to be zero and hence $V_L = 0$. Equations (2.41)-(2.43) thus could be written as:

$$\frac{\mu_L(\underline{p})}{r+\delta} + k_L\underline{p}^{\alpha_L}(1-\underline{p})^{1-\alpha_L} = \frac{\mu_H(\underline{p}) - rV_H}{r+\delta} + k_H\underline{p}^{1-\alpha_H}(1-\underline{p})^{\alpha_H}$$

$$\frac{\mu_{HL}-\mu_{LL}}{r+\delta} + k_L\underline{p}^{\alpha_L}(1-\underline{p})^{1-\alpha_L}\left(\frac{\alpha_L-\underline{p}}{\underline{p}(1-\underline{p})}\right) = \frac{\mu_{HH}-\mu_{LH}}{r+\delta} + k_H\underline{p}^{1-\alpha_H}(1-\underline{p})^{\alpha_H}\left(\frac{1-\alpha_H-\underline{p}}{\underline{p}(1-\underline{p})}\right)$$

$$k_L\underline{p}^{\alpha_L-2}(1-\underline{p})^{-1-\alpha_L}\alpha_L(\alpha_L-1) = k_H\underline{p}^{-1-\alpha_H}(1-\underline{p})^{\alpha_H-2}\alpha_H(\alpha_H-1)$$

This system of equations will give us a unique formula for V_H:

$$rV_H = (\mu_{LH}-\mu_{LL}) + \frac{\alpha_H(\alpha_L-1)(\Delta_H-\Delta_L)\underline{p}}{\alpha_H(\alpha_L-1)-(1-\underline{p})(\alpha_L-\alpha_H)}. \tag{2.44}$$

As usual, $\Delta_H = \mu_{HH} - \mu_{LH}$ and $\Delta_L = \mu_{HL} - \mu_{LL}$. Furthermore, it is easy to check that both k_H and k_L are strictly larger than zero such that the option value of learning is strictly positive.

Therefore, we finally reach our main result:

Theorem 2.3. *Under strict supermodularity, the stationary competitive equilibrium is unique in the sense that all equilibria are PAM and the allocation is uniquely determined by Theorem 2.2. Moreover, assume firm monotonicity and normalize $V_L = 0$, we can get a unique formula for V_H given by equation (2.44).*

2.4.4 Wage Gap at the Cutoff

The analysis of the value functions allows us to determine equilibrium wages. We start with an interesting observation:

$$\begin{aligned} w_H(\underline{p}) = \mu_H(\underline{p}) - rV_H &= \Delta_H \underline{p} + \mu_{LL} - \frac{\alpha_H(\alpha_L - 1)(\Delta_H - \Delta_L)\underline{p}}{\alpha_H(\alpha_L - 1) - (1-\underline{p})(\alpha_L - \alpha_H)} \\ &< \Delta_L \underline{p} + \mu_{LL} = w_L(\underline{p}). \end{aligned}$$

This implies that the worker with posterior belief slightly higher than $\underline{p}$ will accept the high firm's offer even though the wage provided is lower than the wage at the low firm. This obviously comes from the fact that the learning speed in the high firm is higher and this would compensate the loss in the flow wages.

On the other hand, we can see that the difference in expected productivity at $\underline{p}$ is

$$\mu_H(\underline{p}) - \mu_L(\underline{p}) = (\mu_{HL} - \mu_{LL}) + (\Delta_H - \Delta_L)\underline{p} < rV_H.$$

This implies the high firm can enjoy a strictly positive rent from a higher learning speed. This above result actually does not depend on the assumption $V_L = 0$ and it can be generalized for any possible division of surplus.[23] This is illustrated by Figure 2.2:

Lemma 2.7. *Under strict supermodularity, we have:* $w_H(\underline{p}) < w_L(\underline{p})$ *and* $rV_H - rV_L > \mu_H(\underline{p}) - \mu_L(\underline{p})$.

2.5 Firm-dependent Volatility: σ_y

A valid criticism of our approach is that we give the H firms too much of an edge under supermodularity (likewise for the L firms under submodularity). Not only are they superior

[23]Generally, value matching and no-deviation conditions imply that

$$(r+\delta)W_H(\underline{p}) = w_H(\underline{p}) + \Sigma_H(\underline{p})W_H''(\underline{p}) = (r+\delta)W_L(\underline{p}) = w_L(\underline{p}) + \Sigma_L(\underline{p})W_L''(\underline{p})$$

and

$$W_H''(\underline{p}) = W_L''(\underline{p}).$$

These immediately mean that $w_H(\underline{p}) < w_L(\underline{p})$ and $rV_H - rV_L > \mu_H(\underline{p}) - \mu_L(\underline{p})$.

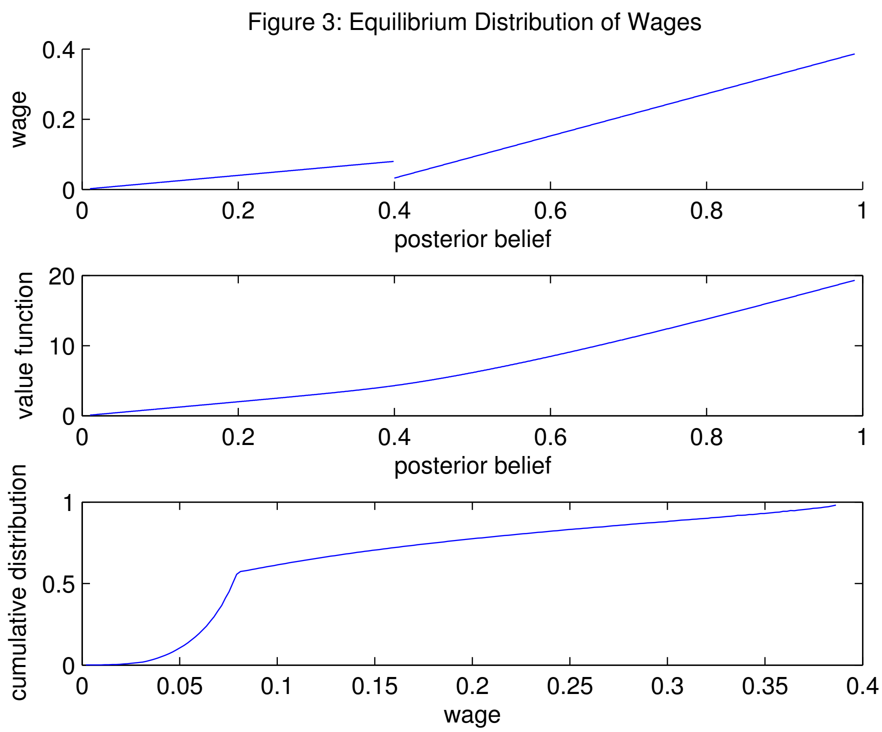

Figure 2.2: Equilibrium wage function and value function in terms of beliefs p; Stationary wage distribution.

in the production of output, by assuming that the volatility σ is common to both types of firms, effectively the signal-to-noise ratio is higher in H firms:

$$s_H = \frac{\mu_{HH} - \mu_{LH}}{\sigma} > \frac{\mu_{HL} - \mu_{LL}}{\sigma} = s_L,$$

from supermodularity. With firm-dependent volatility, that need not be the case. In particular, for σ_H sufficiently high, it may well be the case that $s_H < s_L$.

Mere observation of the value function in Equation (2.3), $rW_y(p) = \mu_y(p) - V_y + \Sigma_y(p)W_y''(p) - \delta W_y(p)$, reveals that firm-dependent volatility will play a crucial role here. Since $\Sigma_y = \frac{1}{2}p^2(1-p)^2 s_y^2$, for sufficiently high σ_H and therefore low s_H, it appears intuitive that the value W_H can be smaller than the value of W_L for high p. It turns out that this intuition is wrong. First, in this competitive equilibrium, wages are endogenous and therefore as the value of learning changes, so does $\mu_y(p) - V_y$. Second, the no-deviation condition requires that at the marginal type $\underline{p}$, $W_H'' = W_L''$. It turns out that as a result these two features, in equilibrium the learning effect is the same in both firms, no matter what the volatility σ_y is.

To make this argument formal, when $\sigma_H \neq \sigma_L$, we generally define $s_y = (\mu_{Hy} - \mu_{Ly})/\sigma_y,\ y = H, L$. It is trivial to show that belief updating also satisfies the formula:

$$dp_t = p_t(1-p_t)s_y d\bar{Z}_{y,t}.$$

Furthermore, Lemmas 2.2–2.5 still hold because none of these results depend explicitly on σ_y. As shown in the appendix, the statement in Claim 2.2 is generalized to *any* combination of (σ_H, σ_L).[24]

With the proof of Claim 2.2 in hand, the result of Theorem 2.1 immediately extends: PAM (NAM) is the unique candidate stationary competitive equilibrium allocation under strict supermodularity (submodularity) thus holds for *any* combination of (σ_H, σ_L). Surprisingly,

[24]The sufficiency of the no-deviation condition is also extended to include all of the combinations of (σ_H, σ_L) by proving a generalized version of Claim 2.2 and Lemma 2.6 in the appendix.

this implies that under strict supermodularity, even if we have an extremely high σ_H such that the learning rate in high type firms is smaller than that in low type firms, we still have PAM. It is equivalent to assert that the direct productivity consideration dominates the learning in our model. The reason comes from the fact that the equilibrium wage schedules adjust to offset the impact of change in learning rate. The key insight here is the no-deviation condition. At $\underline{p}$, the no-deviation condition requires that the second-order effect on the value function is the same in both firms. This second-order effect W_y'' exactly captures the effect of learning through $\Sigma_y(\underline{p})W_y''(\underline{p})$ where $\Sigma_y = \frac{1}{2}p^2(1-p)^2 s_y^2$. Because equilibrium wages adjust to satisfy the no-deviation condition at the cutoff, the impact of differential learning rates is completely offset by the change of wage schedule, and the equilibrium allocation is solely determined by the productivity consideration.

2.6 The Planner's Problem

A priori, we might expect the competitive equilibrium not to decentralize the planner's problem. Wage contracts cannot condition on future realizations or actions and are assumed to be self-enforcing. As a result of this lack of commitment, there is a missing market. With incomplete markets, the competitive equilibrium in general does not necessarily decentralize the planner's problem. It turns out however as we show below that this market incompleteness does not preclude the efficiency of the decentralized equilibrium. As will become apparent, this efficiency result is driven by the martingale property present in all models of learning.

We consider a planner's problem under stationarity, i.e., in the presence of an ergodic distribution. The planner chooses an allocation rule and as a consequence of the Kolmogorov forward equation, the ergodic distribution associated with this allocation rule. The objective is to maximize the aggregate flow of output. Given stationarity of the problem, the focus on output maximization yields the same outcome as maximization of aggregate values.

Before we state and prove the efficiency result, we need to derive the stationary distribution under multiple cutoffs. Consider any allocation with multiple cutoffs:

$$0 < \underline{p}_N < \cdots < \underline{p}_1 < 1, \quad N \text{ odd.}$$

Without loss of generality, we assume workers with $p \in (p_1, 1]$ are allocated to the high type firms while workers with $p \in [0, p_N)$ are allocated to the low type firms since for workers with $p = 0$ or 1, there is no need for learning and it is optimal to allocate them according to instantaneous production efficiency (PAM).[25] This also implies that generically N is odd. Denote by Ω_y the set of p's that match with firms of type y.

Formally, the planner will choose Ω_y to solve the problem:

$$\begin{aligned}
\max_{\Omega_y} S = \int_{\Omega_H} \mu_H(p) f_H(p) dp + \int_{\Omega_L} \mu_L(p) f_L(p) dp & \\
\text{s.t. } \frac{d^2}{dp^2}[\Sigma_y(p) f_y(p)] - \delta f_y(p) = \frac{df_y(p)}{dt} &= 0 \qquad \text{Kolmogorov forward equation} \\
\int_{\Omega_H} p f_H(p) dp + \int_{\Omega_L} p f_L(p) dp &= p_0 \qquad \text{Martingale property} \\
\int_{\Omega_L} f_L(p) dp = 1 - \pi, \quad \int_{\Omega_H} f_H(p) dp &= \pi. \qquad \text{Market clearing}
\end{aligned}$$

It turns out that the martingale property enables an easier way to compare different allocations, hence the following Lemma:

Lemma 2.8. *Consider two possible allocations with ergodic density functions $f_H(p)$, $f_L(p)$ (allocation 1) and $\tilde{f}_H(p)$, $\tilde{f}_L(p)$ (allocation 2) respectively. Then allocation 1 generates higher aggregate output than the allocation 2 if and only if $\int_{\Omega_H} p f_H(p) dp > \int_{\tilde{\Omega}_H} p \tilde{f}_H(p) dp$ or alternatively, $\int_{\Omega_L} p f_L(p) dp < \int_{\tilde{\Omega}_L} p \tilde{f}_L(p) dp$.*

Proof. In Appendix. □

[25] This property is also established in the one-sided model of Anderson and Smith (2010). Our results shows that not only at the extremes but also at the interior the planner's (and the equilibrium) allocation exhibit PAM.

To prove that the competitive equilibrium decentralizes the planner's stationary solution under supermodularity, it suffices to show that the PAM allocation is better than *any* allocation with multiple cutoffs because from Theorem 2.2, we know that PAM allocation is unique and will be the same as the competitive equilibrium allocation for any combination of (s_H, s_L). The key technical issue is that the ergodic distribution is endogenously determined by the allocation rule. It is infeasible to compute the ergodic density functions for each possible allocation. Our strategy of proof is therefore to use a variational argument to circumvent this difficulty.

The proof heavily uses the martingale property and works as follows. First we consider a candidate allocation with 3 cutoffs. Under this candidate allocation, there will be an interior interval of p's that are matched to L type firms associated with some ergodic distribution. We move the bounds of that interval slightly to the left, thus generating a new density in this interval while keeping all other cutoffs and distributions unchanged. The new interval is chosen by imposing market clearing conditions. Lemma 2.8 then shows that under supermodularity this experiment strictly increases aggregate output. This holds until cutoffs coincide such that the interior rang of p's matched with L firms disappears, thus reducing the number of cutoffs to $N = 1$. We use a similar argument to establish that output increases when moving from N to $N - 2$ cutoffs. The result then follows by induction. We derive the result under supermodularity. The same logic applies under submodularity.

Theorem 2.4. *The competitive equilibrium decentralizes the planner's stationary solution that maximizes the aggregate flow of output.*

Proof. In Appendix. □

2.7 On-the-job Human Capital Accumulation

On the job, workers and firms not only learn about their unknown innate skills, they also accumulate human capital. In reality, human capital accumulation is an ongoing, continuous process. The longer the tenure of a worker, the higher her productivity. This monotonically increasing relation between tenure and human capital experience is likely also to be concave. For modeling purposes, here we consider a very simple form that captures this relation. With probability λ, a worker transitions from being unexperienced to being experienced.[26] Once a worker is experienced, her productivity increases to $\mu_{xy} + \xi_x$ and the status of experience is complete information.[27] Now there are the same value functions for experienced workers as before W_y^e.

$$rW_y^e(p) \;\; = \;\; \mu_y(p) + \xi(p) - rV_y + \Sigma_y^e(p) W_y^{e''}(p) - \delta W_y^e(p)$$

where $\xi(p) = p\xi_H + (1-p)\xi_L$ is the expected experience.[28] For the unexperienced worker there is now one additional value function. As before, there are unexperienced workers who are matched with L firms, and who continue to match with an L firms; and there are those who match with H firms both when unexperienced as well as when experienced. We denote those values by W_{LL}^u, W_{HH}^u. There are now also some types p who match with an L firm when unexperienced and who switch to an H firm when they become experienced, the value of which is denoted by W_{LH}^u. This requires that the reservation type of an experienced worker ($\underline{p}^e$) is lower than that of the unexperienced worker ($\underline{p}^u$). We start from this premise and later verify that this is indeed the case. The value functions then are:

$$\begin{aligned}
rW_{yy}^u(p) &= \mu_y(p) - rV_y + \Sigma_y^u(p) W_{yy}^{u''}(p) + \lambda W_y^e(p) - (\delta+\lambda) W_{yy}^u(p) \\
rW_{LH}^u(p) &= \mu_L(p) - rV_L + \Sigma_L^u(p) W_{LH}^{u''}(p) + \lambda W_H^e(p) - (\delta+\lambda) W_{LH}^u(p)
\end{aligned}$$

Observe that even though experience is completely observable, it does affect the inference from learning in the sense that the signal-to-noise ratio changes to $[(\mu_{Hy}+\xi_H-\mu_{Ly}-\xi_L)]/\sigma^2$.

[26] Having a continuous relation between tenure and human capital renders the system of differential equations into a system of partial differential equations. Typically there is no solution. In the current setup, there is an additional state (experienced versus unexperienced) and the model remains tractable.

[27] Observe that experience is worker dependent, but not firm dependent. While it is likely a realistic feature to have experience dependent on the job type, the reason is that we would have a different level of experience for different histories which makes the problem non-tractible.

[28] In this section we maintain the earlier assumption that $\sigma_H = \sigma_L = \sigma$.

As a result, Σ_y depends on experience u, e.

$$
\begin{aligned}
W^u_{yy}(p) &= \frac{\mu_y(p) - rV_y}{r+\delta+\lambda} + k^u_{y1} p^{1-\alpha^u_y}(1-p)^{\alpha^u_y} + k^u_{y2} p^{\alpha^u_y}(1-p)^{1-\alpha^u_y} \\
&+ \frac{\lambda}{(r+\delta)(r+\delta+\lambda)}[\mu_y(p) + \xi(p) - rV_y] \\
&+ \frac{\lambda}{(\lambda+\delta+r) - \frac{(s^u_y)^2}{(s^e_y)^2}(r+\delta)}[k^e_{y1} p^{1-\alpha^e_y}(1-p)^{\alpha^e_y} + k^e_{y2} p^{\alpha^e_y}(1-p)^{1-\alpha^e_y}] \\
W^u_{LH}(p) &= \frac{\mu_L(p) - rV_L}{r+\delta+\lambda} + k^u_{L1} p^{1-\alpha^u_L}(1-p)^{\alpha^u_L} + k^u_{L2} p^{\alpha^u_L}(1-p)^{1-\alpha^u_L} \\
&+ \frac{\lambda}{(r+\delta)(r+\delta+\lambda)}[\mu_H(p) + \xi(p) - rV_H] \\
&+ \frac{\lambda}{(\lambda+\delta+r) - \frac{(s^u_L)^2}{(s^e_H)^2}(r+\delta)}[k^e_{H1} p^{1-\alpha^e_H}(1-p)^{\alpha^e_H} + k^e_{H2} p^{\alpha^e_H}(1-p)^{1-\alpha^e_H}] \\
W^e_y(p) &= \frac{\mu_y(p) + \xi(p) - rV_y}{r+\delta} + k^e_{y1} p^{1-\alpha^e_y}(1-p)^{\alpha^e_y} + k^e_{y2} p^{\alpha^e_y}(1-p)^{1-\alpha^e_y}
\end{aligned}
$$

where

$$
\alpha^u_y = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{2(r+\delta+\lambda)}{(s^u_y)^2}} \geq 1
$$
$$
\alpha^e_y = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{2(r+\delta)}{(s^e_y)^2}} \geq 1
$$

There are now two cut-offs $\underline{p}^u, \underline{p}^e$. Since we just want to compare $\underline{p}^u$ and $\underline{p}^e$, we can consider the following thought experiment. First, we assume that $\underline{p}^u = \underline{p}^e = \underline{p}$. Then we can get two systems of equations: one system is the set of value-matching, smooth-pasting and no-deviation conditions for the unexperienced workers and the other one is for the experienced workers. Second, we can solve $\Delta V = V_H - V_L$ the way we did previously but now we can get two possible values for ΔV. Denote them to be ΔV^e and ΔV^u. Notice that ΔV^e and ΔV^u are both increasing in the cutoff $\underline{p}$. Finally, we compare ΔV^e and ΔV^u under the assumption that $\underline{p}^u = \underline{p}^e = \underline{p}$. If $\Delta V^e > \Delta V^u$, this means that we should decrease $\underline{p}^e$ or increase $\underline{p}^u$ and hence $\underline{p}^u > \underline{p}^e$; on the contrary, if $\Delta V^e < \Delta V^u$, this means that we should decrease $\underline{p}^u$ or increase $\underline{p}^e$ and hence $\underline{p}^u < \underline{p}^e$. We derive this in the Appendix and can show this to hold when human capital accumulation is not too different for H and L types.

Proposition 2.3. *Assume supermodularity and $\xi_H \simeq \xi_L$. Then $\underline{p}^e < \underline{p}^u$.*

Proof. In Appendix. $\square$

With human capital accumulation, we can now characterize the entire equilibrium, including wage schedules and the ergodic distribution of types. Even though there are types who gradually learn they are of low productivity, wages need not decrease over the life cycle as they accumulate human capital.

Turnover and Tenure. We express the expected future duration of a match by tenure $\tau_y(p)$. Tenure relates inversely to turnover. For $p < \underline{p}^e$ and $p > \underline{p}^u$, $\tau_y(p)$ satisfies the following differential equation (see also Moscarini 2005):

$$\Sigma_y(p)\tau_y''(p) - \delta\tau_y(p) = -1,$$

with solutions:

$$\tau_H^u(p) = \frac{1}{\delta}\left\{1 - \left(\frac{p}{\underline{p}^u}\right)^{1/2-\sqrt{1/4+2\delta/(s_H^u)^2}} \left(\frac{1-p}{1-\underline{p}^u}\right)^{1/2-\sqrt{1/4-2\delta/(s_H^u)^2}}\right\};$$

$$\tau_L^u(p) = \frac{1}{\delta}\left\{1 - \left(\frac{p}{\underline{p}^u}\right)^{1/2-\sqrt{1/4-2\delta/(s_L^u)^2}} \left(\frac{1-p}{1-\underline{p}^u}\right)^{1/2-\sqrt{1/4+2\delta/(s_L^u)^2}}\right\};$$

$$\tau_H^e(p) = \frac{1}{\delta}\left\{1 - \left(\frac{p}{\underline{p}^e}\right)^{1/2-\sqrt{1/4+2\delta/(s_H^e)^2}} \left(\frac{1-p}{1-\underline{p}^e}\right)^{1/2-\sqrt{1/4-2\delta/(s_H^e)^2}}\right\};$$

$$\tau_L^e(p) = \frac{1}{\delta}\left\{1 - \left(\frac{p}{\underline{p}^e}\right)^{1/2-\sqrt{1/4-2\delta/(s_L^e)^2}} \left(\frac{1-p}{1-\underline{p}^e}\right)^{1/2-\sqrt{1/4+2\delta/(s_L^e)^2}}\right\}.$$

If $p \in (\underline{p}^e, \underline{p}^u)$, the only difference is that

$$\Sigma_y(p)\tau_L^{u\prime\prime}(p) - (\delta+\lambda)\tau_L^u(p) = -1,$$

since unexperienced workers will switch jobs once they become experienced. An immediate implication of the Proposition above is the following:

Proposition 2.4. *(Tenure) Assume supermodularity and $\xi_H \simeq \xi_L$. Then, $\tau_L^u(p) > \tau_L^e(p)$ for $p < \underline{p}^e$ and $\tau_H^u(p) < \tau_H^e(p)$ for $p > \underline{p}^u$. For $p \in (\underline{p}^e, \underline{p}^u)$, there is a cutoff such that $\tau_L^u(p) < \tau_H^e(p)$ for p higher than this cutoff and $\tau_L^u(p) > \tau_H^e(p)$ for p smaller than this cutoff.*

For the lowest types p, tenure for the unexperienced worker is longer as the experienced workers are more likely to be hired by an H firm given positive information revelation. The opposite is true for the highest p: the unexperienced types face a higher cut-off type and will therefore upon bad information be more likely to switch to an L firm. In the intermediate range, tenure depends on how close p is to either of the cut-offs.

2.8 Robustness

2.8.1 Generalized Lévy Processes

One may suspect that our results are exclusively driven by the specific assumptions of the Brownian motion. In the section, we illustrate that this is not the case by considering a generalized Lévy process, i.e., a compound Poisson process. Let λ_{xy} denote the expected arrival rate of jumps for a type x worker in a type y firm. Following Cohen and Solan (2009), the worker's value function can be written as:

$$\begin{aligned} W_y(p) = w_y(p)dt + (1 - rdt - \delta dt)\{[p\lambda_{Hy} + (1-p)\lambda_{Ly}]dtW_{y'}(p_h) \\ + (1 - [p\lambda_{Hy} + (1-p)\lambda_{Ly}]dt)W_y(p + dp) \end{aligned}$$

where $p_h = \frac{p\lambda_{Hy}}{p\lambda_{Hy}+(1-p)\lambda_{Ly}}$ and y' is the firm type which matches with worker p_h. If no jump occurs, the updating of the posterior belief in firm y follows:

$$dp = -p(1-p)(\lambda_{Hy} - \lambda_{Ly})dt + p(1-p)s_y d\bar{Z}.$$

As usual, the value function could be rewritten as a differential equation:

$$\begin{aligned} &(r + \delta + [p\lambda_{Hy} + (1-p)\lambda_{Ly}])W_y(p) \\ &\quad = w_y(p) + [p\lambda_{Hy} + (1-p)\lambda_{Ly}]W_{y'}(p_h) - p(1-p)(\lambda_{Hy} - \lambda_{Ly})W_y'(p) + \Sigma_y(p)W_y''(p). \end{aligned}$$

The no-deviation condition derived earlier still holds in this situation. The proof is similar and is omitted here.

Lemma 2.9. *To deter possible deviations, a necessary condition is:*

$$W''_H(\underline{p}) = W''_L(\underline{p}) \quad \textit{(No-deviation condition-Lévy)} \tag{2.45}$$

for any possible cutoff $\underline{p}$.

Consider the simplifying assumption that $\lambda_{Ly} = 0$ and denote λ_{Hy} by λ_y. Then p_h is always 1 and the value function becomes:

$$(r + \delta + p\lambda_y)W_y(p) = w_y(p) + p\lambda_y W_{y'}(1) - p(1 - p)\lambda_y W'_y(p) + \Sigma_y(p)W''_y(p).$$

The differential equation could be solved explicitly by guess and verify:

$$W_y(p) = A_y + B_y p + k_{y1}p^{\alpha_{y1}}(1-p)^{1-\alpha_{y1}} + k_{y2}p^{\alpha_{y2}}(1-p)^{1-\alpha_{y2}}$$

where $A_y = \frac{\mu_{Ly} - rV_y}{r+\delta}$, $B_y = \frac{\Delta_y + \lambda_y(W_{y'}(1) - A_y)}{r+\delta+\lambda_y}$ and

$$\begin{aligned}
\alpha_{y1} &= \frac{1}{2} + \frac{\lambda_y}{s_y^2} + \sqrt{(\frac{1}{2} + \frac{\lambda_y}{s_y^2})^2 + \frac{2(r+\delta)}{s_y^2}} > 1 + 2\frac{\lambda_y}{s_y^2} \\
\alpha_{y2} &= \frac{1}{2} + \frac{\lambda_y}{s_y^2} - \sqrt{(\frac{1}{2} + \frac{\lambda_y}{s_y^2})^2 + \frac{2(r+\delta)}{s_y^2}} < 0.
\end{aligned}$$

Obviously, the envelope of W_y is a strictly increasing and strictly convex function for $p \in (0, 1)$. First, we would like to argue that for $p = 1$, $y' = H$. Since the function is strictly convex, it must be the case that 0 and 1 workers are matched with different types of firms. Now suppose $y' = L$. Then since 0 workers are matched with H firms, $A_H > A_L$ and hence $W_L(1) = \frac{\Delta_L}{r+\delta} + A_L < \frac{\Delta_H}{r+\delta} + A_H = W_H(1)$. A contradiction.

Therefore, the value function could be rewritten as:

$$(r + \delta + p\lambda_y)W_y(p) = w_y(p) + p\lambda_y W_1(1) - p(1-p)\lambda_y W'_y(p) + \Sigma_y(p)W''_y(p). \tag{2.46}$$

with general solution:

$$W_y(p) = A_y + B_y p + k_{y1} p^{\alpha_{y1}} (1-p)^{1-\alpha_{y1}} + k_{y2} p^{\alpha_{y2}} (1-p)^{1-\alpha_{y2}}. \tag{2.47}$$

Notice that the equilibrium payoffs are such that $A_L > A_H$, $B_L < B_H$ and $A_L + B_L < A_H + B_H$. At any cutoff $\underline{p}$, the following three equations should hold simultaneously:

$$W_H(\underline{p}) = W_L(\underline{p}) \qquad \text{(Value-matching condition)} \tag{2.48}$$

$$W_H^{'}(\underline{p}) = W_L^{'}(\underline{p}) \qquad \text{(Smooth-pasting condition)} \tag{2.49}$$

$$W_H^{''}(\underline{p}) = W_L^{''}(\underline{p}) \qquad \text{(No-deviation condition)} \tag{2.50}$$

Then from Equation (2.46), it is immediate to get at $\underline{p}$,

$$\begin{aligned}&(\lambda_H - \lambda_L)\underline{p} W_H(\underline{p}) \\ &= w_H(p) - w_L(p) + (\lambda_H - \lambda_L)\underline{p} W_H(1) - (\lambda_H - \lambda_L)\underline{p}(1-\underline{p}) W_H'(\underline{p}) + (\Sigma_H(p) - \Sigma_L(p)) W_H''(p).\end{aligned}$$

Apply Equation (2.47) and the above equation could be simplified as:

$$0 = w_H(\underline{p}) - w_L(\underline{p}) + (r + \delta + \lambda_L)[A_L - A_H + (B_L - B_H)\underline{p}].$$

The RHS of the above equation is linear in $\underline{p}$. Therefore, if we can prove the slope is not zero then there cannot exist two $\underline{p}$'s satisfying the equation simultaneously. Fortunately, this is the case. The slope is

$$\Delta_H - \Delta_L + (r + \delta + \lambda_L)(B_L - B_H).$$

Notice that $B_H = \frac{\Delta_H}{r+\delta}$ and $(r + \delta + \lambda_L)B_L = \Delta_L + \lambda_L(W_H(1) - A_L)$. Hence,

$$\Delta_H - \Delta_L + (r + \delta + \lambda_L)(B_L - B_H) = \lambda_L(A_L - A_H) > 0.$$

The following result summarizes the findings above and corresponds to Theorem 2.1 in the Brownian motion case:

Proposition 2.5. *Given the Lévy process and provided an equilibrium exists, PAM is the unique stationary competitive equilibrium allocation under strict supermodularity.*

Under PAM, $k_{L1} > 0$, $k_{L2} = 0$ and $k_{H1} = 0$, $k_{H2} > 0$. We can use the procedure introduced in the previous sections to pin down the equilibrium cutoff $\underline{p}$ and derive value functions based on $\underline{p}$.

Notice also that under the Lévy process, beliefs are formed through Bayesian updating. We conjecture that PAM will always be the competitive equilibrium allocation under strict supermodularity for any stochastic process as long as there is Bayesian updating. This is because under Bayesian learning, the belief updating process is always a martingale. Of course, establishing this result for general information processes is impossible because it requires the explicit solution of the differential equations for the value function, which generally does not exist.

2.8.2 Non-Bayesian Updating

Suppose instead that the belief updating is not a martingale. Then it must be generated by some non-Bayesian learning process. We will now show for an example that the competitive equilibrium can be non-PAM even if there is supermodularity.

Suppose the belief updating process in firm y is given by: $dp = \lambda_y p dt$ for $p < 1$, with λ_y a constant, and once p reaches 1, $dp = 0$. We may think p as a special human capital with 1 as an upper bound on the accumulation. The value function of a worker is given by:[29]

$$(r + \delta)W_y(p) = w_y(p) + \lambda_y p W_y'(p)$$

with solution:

$$W_y(p) = C_y p^{\frac{r+\delta}{\lambda_y}} + \frac{\Delta_y}{r + \delta - \lambda_y} p + \frac{\mu_{Ly} - rV_y}{r + \delta}.$$

[29] We can write the value of a worker of type p in firm y as $W_y(p) = w_y(p)dt + (1 - (r + \delta)dt)W_y(p + dp)$. Using a Taylor expansion $W_y(p + dp) = W_y(p) + W_y'(p)dp + o(dt)$ and the fact that $dp = \lambda_y p dt$, we obtain the expression for $W_y(p)$.

Suppose PAM is the equilibrium allocation, then

$$\lim_{p\to 1} W_H(p) = W_H(1) = \frac{\Delta_H}{r+\delta}p + \frac{\mu_{LH} - rV_H}{r+\delta},$$

which implies that:

$$C_H = -\frac{\lambda_H \Delta_H}{(r+\delta)(r+\delta-\lambda_H)}.$$

At the cutoff $\underline{p}$ we have:

$$W_H(\underline{p}) = W_L(\underline{p}) \qquad \text{(Value-matching condition)} \tag{2.51}$$

$$W_H^{'}(\underline{p}) = W_L^{'}(\underline{p}), \qquad \text{(Smooth-pasting \& No-deviation condition)} \tag{2.52}$$

where it turns out that for this belief-updating process, the no-deviation condition coincides with the smooth-pasting condition. We derive the no-deviation condition in the Appendix.

This is a system of equations in C_L and $\underline{p}$. Substitute C_L and $\underline{p}$ could be expressed as:

$$\frac{\Delta_L}{r+\delta}\underline{p} + \frac{\mu_{LL} - rV_L}{r+\delta} = \frac{\lambda_L - \lambda_H}{r+\delta}\frac{\Delta_H}{r+\delta-\lambda_H}(\underline{p})^{\frac{r+\delta}{\lambda_H}} + (1 - \frac{\lambda_L}{r+\delta})\frac{\Delta_H}{r+\delta-\lambda_H}\underline{p} + \frac{\mu_{LH} - rV_H}{r+\delta}$$

or

$$\frac{\Delta_L - \Delta_H}{r+\delta}\underline{p} + \frac{\mu_{LL} - rV_L}{r+\delta} = \frac{\lambda_H - \lambda_L}{r+\delta}\frac{\Delta_H}{r+\delta-\lambda_H}[\underline{p} - (\underline{p})^{\frac{r+\delta}{\lambda_H}}] + \frac{\mu_{LH} - rV_H}{r+\delta}. \tag{2.53}$$

Notice that PAM requires that the $p = 0$ worker has incentive to be matched with L firms. Hence,

$$\frac{\mu_{LL} - rV_L}{r+\delta} > \frac{\mu_{LH} - rV_H}{r+\delta}$$

Also notice that

$$\frac{\lambda_H - \lambda_L}{r+\delta}\frac{\Delta_H}{r+\delta-\lambda_H}[\underline{p} - (\underline{p})^{\frac{r+\delta}{\lambda_H}}] < 0$$

if $\lambda_L > \lambda_H$ and $r + \delta > \lambda_H$.

If we can show that

$$\frac{\Delta_H - \Delta_L}{r+\delta}\underline{p} < \frac{\lambda_L - \lambda_H}{r+\delta}\frac{\Delta_H}{r+\delta-\lambda_H}[\underline{p} - (\underline{p})^{\frac{r+\delta}{\lambda_H}}],$$

then Equation (2.53) cannot hold as equality, which is the result we are looking for. First notice that the LHS of the inequality goes to zero as $\Delta_H - \Delta_L$ decreases to zero. Meanwhile, the belief updating process implies the ergodic distribution only depends on λ's and will not depend on Δ's. From previous sections, if PAM is indeed the equilibrium allocation, then $\underline{p}$ should not depend on Δ's. Therefore, fix any $\lambda_L > \lambda_H$ and $r + \delta > \lambda_H$ and we can derive some corresponding $\underline{p} \in (0, 1)$. Then, let $\Delta_H - \Delta_L$ decreases to zero and it is immediate to see that eventually we will have:

$$\frac{\Delta_H - \Delta_L}{r + \delta}\underline{p} < \frac{\lambda_L - \lambda_H}{r + \delta} \frac{\Delta_H}{r + \delta - \lambda_H}[\underline{p} - (\underline{p})^{\frac{r+\delta}{\lambda_H}}].$$

This implies that PAM cannot be an equilibrium if $\lambda_L > \lambda_H$ and the degree of supermodularity is sufficiently small.

2.9 Concluding Remarks

In this paper, we have proposed a competitive equilibrium model of the labor market that unifies frictionless sorting and a learning-based theory of turnover. In equilibrium under supermodularity, workers with better posteriors about their ability tend to sort into more productive jobs. The main technical contribution of this paper is that we find a new constraint on the worker's value function as a result of sequential rationality in the presence of competitively determined payoffs. At the cutoff type, the second derivative of the workers' value function must equate. In addition to the standard conditions of value-matching (zero-th derivative) and smooth-pasting (first derivative), we now also have the no-deviation condition (second derivative).

What is possibly most surprising is that the result of positive sorting under supermodularity is not determined by the speed of learning. In the trade-off between the learning speed and instantaneous productive efficiency, productive efficiency always takes the upper hand. As such, the equilibrium allocation does not depend on the signal-to-noise ratio (the ratio of the average payoff gain, which measures the efficiency, over the noise term). This seems to indicate in this competitive environment the sorting aspect dominates the learning. Quite surprisingly, this sorting result does not hinge on the particular information structure and is robust to general Bayesian learning processes.

Our analysis has certain limitations and several issues remain unanswered. First, like most experimentation models, payoffs are linear and agents are risk neutral. Non-linearity is desirable for the economic interpretation. However, it renders the solution to the differential equation of the value function much harder to solve.

Second, ideally we would like to extend the analysis to general distributions of worker and firm types. Like in much of the experimentation literature the realized type is either high or low on a risky arm. Here, in addition we have *two* risky arms that are correlated, since there is learning in both types of firms. The focus on the two firm-type case (two arms) keeps down the dimensionality of the continuous time problem. With more than two firm types, analyzing the Brownian motion process is mathematically substantially more demanding.

Finally, our result that PAM obtains under supermodularity and that the planner's problem can be decentralized, is established for a stationary equilibrium. While a solution of a general non-stationary equilibrium is too complex, one can easily construct a two-period counterexample in which PAM will not necessarily obtain in a non-stationary environment.

Chapter 3

Learning In War of Attrition Games

3.1 Introduction

Imagine a situation where two players are bargaining over a joint decision or two political parties are voting for a bill. The two individuals disagree with one another because of conflicting preferences. In particular, each of the two players must choose between sticking to his own favorable choice or conceding to the other player's favorable choice. The return to conceding decreases with time, but, at any time, a player earns a higher return if the other concedes first. War of attrition games are theoretical tools widely used to characterize how each of the two players chooses a time path of conceding in the event that the other player has not already conceded. Continuous-time war of attrition games have been investigated under both complete information (Hendricks, Weiss, and Wilson (1988)) and incomplete information (Abreu and Gul (2000)).

Delay is a key feature in war of attrition games. As shown by both Hendricks, Weiss, and Wilson (1988) and Abreu and Gul (2000), there at least exists an equilibrium such that rational players will randomize between conceding and staying.[1] As a result, it takes time to reach an agreement. However, in many realistic situations, each player is also receiving private information about how favorable the alternatives are while he is bargaining with the other player. Especially in the political environment, learning by political parties is a very

[1] Abreu and Gul (2000) show that this is the unique sequential equilibrium when information is incomplete.

common phenomenon. If a player learns that his opponent's alternative is quite favorable, he becomes more willing to concede. In this sense, the learning process may exogenously facilitate the players to reach an agreement. However, rational players will also respond to such a process, which may cause a longer delay. The natural question to ask is: if there is an exogenous information flow that facilitates an agreement, is it easier to reach an agreement taking into account the response of the rational players?

This paper develops a dynamic war of attrition model with learning to answer the above question. Learning is modelled in the following way: I assume at each point in time, each player may receive a private Poisson signal that reveals the payoff for conceding. Receiving the signal makes the player more willing to concede. This captures the idea that the flow of information exogenously facilitates an agreement. The main result of the paper is the following: compared to the model without learning, learning makes it more difficult to reach an agreement. Especially when the learning rate is low, the expected concession rate in the unique sequential equilibrium is always smaller than the expected concession rate without learning. When the learning rate is high, there also exist periods in which the expected concession rate is higher than the expected concession rate without learning. However, the paper shows that there will also be some periods in which it is harder to reach an agreement compared to the model without learning. In equilibrium, it may be easier to reach an agreement initially but it becomes more and more difficult over time. The later decrease in the concession rate will always offset the former increase and hence the expected expected delay becomes longer instead of shorter. I also consider a one-sided learning model where only one of the two players is able to learn. Interestingly, that model shows that to make the delay shorter, it is better to allow only one player to learn than to allow both to learn.

Due to private learning, each player may have two possible rational types at each point in time. The player could be either sure about his private payoff or still unsure. The sure player is more willing to concede than the unsure player. I show that in the equilibrium when concession still takes place, only one of the following three cases is possible: 1) the sure player is randomizing while the unsure player strictly prefers staying; 2) the unsure player is randomizing while the sure player strictly prefers conceding; and 3) the sure player strictly prefers conceding and the unsure player strictly prefers staying. The expected concession rate in the first (second) scenario is strictly higher (lower) than the expected concession rate without learning.

A player's strategy and the learning rate determine the expected concession rate of this player, which affects his opponent's equilibrium play. When the learning rate is sufficiently high, the first scenario will happen initially but eventually the second scenario will happen. Since the sure player always concedes (weakly) before the unsure player, the posterior belief that a player is unsure is (weakly) increasing as no concession happens. This increases delay since more weight has to be put on the second scenario, which has a lower expected concession rate. Interestingly, my paper shows that learning might decrease delay if learning did not change the weight because the expected concession rate is convex in posterior beliefs. However, delay is always increasing if I take into account the increasing in posterior beliefs. Although it is difficult to the derive the explicit expression of the expected equilibrium delay, a lower bound can be derived assuming the players choose the highest concession rate in all of the three scenarios. The paper shows that even this lower bound is higher than the expected delay without learning.

This paper is closely related to literature on bargaining and delay. The classical complete information bargaining game developed by Rubinstein (1982) has the feature that agreement is reached immediately. Although delay is possible in some variations of the Rubinstein bargaining framework (see e.g., Baron and Ferejohn (1989) and Merlo and Wilson (1995)), many authors have focused on incomplete information as the prime cause of delay (Kennan and Wilson (1993)). A non-exclusive list of sequential bargaining models with incomplete information includes Abreu and Gul (2000), Admati and Perry (1987), Chatterjee and Samuelson

(1987) and Damiano, Li, and Suen (2010a,b). In many of the above papers, the concession game structure is derived from a bargaining or political environment. In this paper, I assume a concession game theoretical framework by writing down the payoff matrix directly. Private learning generates multiple rational types on the equilibrium path. Compared to the standard incomplete information war of attrition model with only one rational type, this increases the difficulty of characterizing the equilibrium. Still, I am able to fully characterize the unique sequential equilibrium under some parameter values.

There are also several papers considering how public learning affects delay in the complete information Rubinstein bargaining framework. For example, Avery and Zemsky (1994) consider a situation where the players are allowed to wait for new public information about the size of the pie before accepting or rejecting an offer. In such an environment, the players may exercise their *option value of waiting*, yielding long delays with positive probability. In my model, learning is about player's private payoff state. For each player, there is no *option value of waiting* associated with learning since learning is a martingale process. The key driving force is that private learning generates more asymmetric information. The interaction between different private types leads to a longer delay. Yildiz (2004) considers a model where learning might increase delay in a complete information sequential bargaining model. However, to generate this result, the players have to be excessively optimistic about their bargaining power. Also, in that model, learning may not increase delay under some parameter values whereas in my model, learning *always* increases delay.

Recently, Kim and Xu (2011) also consider learning in war of attrition games. In their paper, learning is about the common payoff state, while, in this paper, learning is about the private payoff state. Both papers discuss the incentive of information acquisition. In Kim and Xu (2011), information acquisition is modelled as revealing the common payoff state immediately after paying a sunk cost. However, in this paper, information acquisition is modelled as choosing the learning rate by paying a flow cost. I show that if the maximum achievable learning rate is sufficiently low, then nobody has an incentive to acquire any information in the unique sequential equilibrium.

The remainder of this paper organizes as follows. Section 3.2 presents the concession game theoretical framework. Then Section 3.3 analyzes the benchmark model without learning about private payoff states. Section 3.4 characterizes the equilibrium where there is learning about private payoff states and compares the expected concessions with and without learning. Section 3.5 extends the model to investigate endogenous information acquisition. Finally, Section 3.6 concludes.

3.2 Model Setting

Two risk-neutral players ($i = 1, 2$) are playing a continuous-time war of attrition game. There is no discounting. At each point in time, both players have to choose simultaneously between one of two actions: to stay (S) or to quit (Q). Each player is either a commitment type or a normal type. The commitment type player will always choose to stay. For the normal type players, if neither of the two players chooses to quit, the game continues and each player has to incur a flow cost of c, which reflects the cost of delay. If at least one of the players chooses to quit, the payoff matrix is specified as the following (player 1 is the row player and player 2 is the column player):

	S	Q
S	(-,-)	(v_H, v_2)
Q	(v_1, v_H)	(M, M)

If player i stays while $-i$ quits, then player i is the winner of the game and gets a winning payoff of v_H. If player i quits first, then he is the loser and gets a losing payoff v_i. The payoff when both players quit simultaneously is M. There is common knowledge about v_H and $M < v_H$ but there is incomplete information about losing payoffs v_1 and v_2.

In particular, I assume that v_1 and v_2 follow independent and identical binary distributions. v_i can be either a positive number $v_L < v_H$ or zero. Throughout the paper, I will maintain the assumption that $v_L \leq \frac{1}{2}v_H$. The reason for making this technical assumption is to guarantee the equilibrium expected concession rate is monotonic over time. Each player i initially does not know the exact value of v_i. It is common knowledge that $v_i = v_L$ happens with prior p_0. It is also common knowledge that a player is normal with probability γ_0.

Remark 3.1. *In the current model, the flow cost of delay is fixed. An alternative way of modelling the cost of delay is to introduce discounting. However, discounting has an undesirable feature if I maintain the same assumption on the losing payoff v_i. In particular, if the player i knows for sure that $v_i = 0$, there will be no cost of delay and hence the expected delay is infinity. To avoid this issue, I have to assume that v_i can be either $\bar{v}_L$ or $\underline{v}_L$, where $0 < \underline{v}_L < \bar{v}_L < v_H$. Under that hypothesis, I conjecture that there is a unique sequential equilibrium, and the equilibrium has the same qualitative feature as characterized by Theorem 3.4.*

3.3 Benchmark Case: No Learning

I will first discuss the case without learning as a benchmark. Without learning, each normal player's belief that $v = v_L$ will stay at p_0. But there is incomplete information for each player i, since he is unsure whether his opponent is a normal player. The key is to characterize how a normal type player chooses a time path of conceding in the event that the other player has not already conceded. In the future, I will refer to a normal type player whenever I use the term "player."

A strategy for normal player 1 (2) is denoted as $X^1(t)$ ($X^2(t)$) where $X^i(t)$ denotes the probability that player i concedes to player $-i$ by time t (inclusive). $X^i(0)$ is allowed to be strictly positive such that player i concedes to player $-i$ immediately. I use $F^i(t)$ to denote player $-i$'s expected probability that player i concedes to player $-i$ by time t. Obviously, $F^i(t) = \gamma_0 X^i(t)$. Therefore, I can use either $F^i(t)$ or $X^i(t)$ to denote player i's strategy. Given player 2's strategy F^2, player 1's expected payoff by conceding at time t is given by:

$$U^1(t,F^2)=\int_{s<t}(v_H-cs)dF^2(s)+(M-ct)(F^2(t)-F^2(t-))+(p_0v_L-ct)(1-F^2(t)).$$

Here $F^2(t-)=\lim_{\tau\nearrow t}F^2(\tau)$. The expected payoff from never conceding is given by:

$$U^1(\infty,F^2)=\int_{s<\infty}(v_H-cs)dF^2(s).$$

Finally, define $U^1(F^1,F^2)$ to be player 1's expected discounted value by playing the profile (F^1,F^2). Formally, $U^1(F^1,F^2)$ can be written as:

$$U^1(F^1,F^2)=\int_{t\in[0,\infty]}U^1(t,F^2)dX^1(t)=\frac{1}{\gamma_0}\int_{t\in[0,\infty]}U^1(t,F^2)dF^1(t).$$

$U^2(F^1,F^2)$ can be defined similarly. A Nash equilibrium is defined as a profile of $F=(F^1,F^2)$ such that $F^i\in\text{argmax}U^i(\cdot,F^{-i})$.

The set of sequential equilibria is characterized by the following proposition:

Proposition 3.1. *Without learning, there exists a unique sequential equilibrium such that:*

(1) each normal type player concedes at a positive rate between time 0 and T where

$$T=-\frac{(v_H-p_0v_L)\log(1-\gamma_0)}{c}.$$

After time T, only the commitment type player stays;

(2) for each player at time $t\in[0,T]$, the expected concession rate $f_t=\frac{dF(t)/dt}{1-F(t)}$ is a constant $\frac{c}{(v_H-p_0v_L)}$; the normal type player's concession rate $x_t=\frac{dX(t)/dt}{1-X(t)}$ satisfies:

$$x_t[1-(1-\gamma_0)e^{\frac{ct}{v_H-p_0v_L}}]=\frac{c}{(v_H-p_0v_L)}.$$

Sketch of the proof. The proof of the above proposition is similar to the proof of proposition 1 in Abreu and Gul (2000). In particular, the key features of the candidate equilibrium are the same as those in Abreu and Gul (2000):

(1) A normal type player will not delay conceding once he knows that his opponent will never concede.

(2) F^i is continuous and strictly increasing.

(3) At time 0, neither of the two players concedes with a positive probability.

(4) After time 0, each player is indifferent between conceding and staying for any t before T.

The last property implies that the expected utility of a normal type player $-i$ who concedes at time t is the same as $p_0 v_L$ for all $t \in [0, T]$:

$$p_0 v_L = \int_0^t (v_H - cs) dF^i(s) + (p_0 v_L - ct)(1 - F^i(t)).$$

As a result, $F^i(\cdot)$ is differentiable and $f_t^i \triangleq \frac{dF^i(t)/dt}{1-F^i(t)}$ satisfies:

$$(v_H - p_0 v_L) f_t^i = c.$$

It is straightforward to see the expected concession rate $f_t^i = \frac{c}{(v_H - p_0 v_L)}$ for $i = 1, 2$.

On the other hand, since $F^i(t) = \gamma_0 X^i(t)$, f_t^i can be shown to be $x_t^i \gamma_t^i$, where $x_t^i \triangleq \frac{dX^i(t)/dt}{1-X^i(t)}$ is normal type player i's concession rate at time t and γ_t^i denotes the posterior belief that player i is normal given that player i does not concede until time t.

Obviously, the rate x_t^i is chosen such that player i's normal opponent is indifferent between staying and quitting and hence

$$\gamma_t^i x_t^i = \frac{c}{v_H - p_0 v_L}.$$

γ_t^i is updated by Bayes rule:

$$\gamma_t^i = \frac{\gamma_0 - F^i(t)}{1 - F^i(t)}$$

and the law of motion for γ_t^i satisfies: $\dot{\gamma}_t^i = -x_t^i \gamma_t^i (1 - \gamma_t^i)$, which implies that beginning from γ_0, $\gamma_t^1 = \gamma_t^2 = \gamma_t$ for all t such that

$$\dot{\gamma}_t = -\frac{c}{v_H - p_0 v_L}(1 - \gamma_t).$$

The solution to the above differential equation is given by $\gamma_t = 1 - (1 - \gamma_0) e^{\frac{ct}{v_H - p_0 v_L}}$. The normal type players will concede for sure if γ_t reaches zero. Therefore, for the normal type players, the game will last for at most $T = -\frac{(v_H - p_0 v_L)\log(1-\gamma_0)}{c}$ length of time.[2] $\blacksquare$

The expected delay is infinity since the commitment type players will always choose to stay. However, the expected delay Ω conditional on at least one of the two players being normal is finite.Conditional on at least one of the two players being normal, with probability $\frac{2\gamma_0 - 2\gamma_0^2}{2\gamma_0 - \gamma_0^2}$, the conceding times follow a truncated exponential distribution

$$F(t) = \frac{1 - e^{-\frac{ct}{v_H - p_0 v_L}}}{\gamma_0};$$

with probability $\frac{\gamma_0^2}{2\gamma_0 - \gamma_0^2}$, the conceding times follow the distribution

$$\tilde{F}(t) = 1 - (1 - F(t))^2.$$

Therefore, the expected delay Ω is given by:

$$\Omega = \frac{2\gamma_0 - 2\gamma_0^2}{2\gamma_0 - \gamma_0^2} \int_0^T t dF(t) + \frac{\gamma_0^2}{2\gamma_0 - \gamma_0^2} \int_0^T t d[1 - (1 - F(t))^2]$$
$$= \frac{1}{2\gamma_0 - \gamma_0^2} \int_0^T t d(1 - e^{-\frac{2ct}{v_H - p_0 v_L}}) = \frac{v_H - p_0 v_L}{c(2\gamma_0 - \gamma_0^2)} [\frac{1}{2}(1 - (1-\gamma_0)^2) + \log(1-\gamma_0)(1-\gamma_0)^2].$$

As γ_0 goes to one, the limiting equilibrium is the following equilibrium in the complete information game: each player concedes with rate $\frac{c}{v_H - p_0 v_L}$, the maximum delay time is infinity and the expected delay is $\frac{v_H - p_0 v_L}{2c}$.

3.4 Learning

I introduce learning by assuming that after the game starts, as long as no player concedes, each player receives an exogenous private signal that arrives according to a Poisson process. The Poisson processes are independent across players. The arrival rate is λ if $v = v_L$ and zero otherwise. Therefore, after receiving this signal, player i immediately believes with

[2] Compared to a complete information war of attrition game, the incomplete information setting substantially reduces the set of equilibria. In a complete information war of attrition, there always exists a degenerate equilibrium where player i concedes immediately while player $-i$ never concedes.

probability one that $v = v_L$. Absence of the signal will make the player more and more pessimistic about the probability that $v = v_L$. In this section, I will first solve a two-sided learning model where both players have access to the above learning technology and then solve a one-sided model where one of the two players is able to learn.

3.4.1 Two-Sided Learning

Compared to a model without learning, learning adds more uncertainty about each player's type. In particular, at any time $t > 0$, each normal type player may have different private beliefs about his payoff state depending on the learning outcomes. If the player has received at least one Poisson signal, he believes $v = v_L$ for sure. I call him a sure type player. If the player has not received any Poisson signal, his posterior belief about $v = v_L$ becomes $p_t = \frac{p_0 e^{-\lambda t}}{p_0 e^{-\lambda t} + 1 - p_0}$. I call him a learning type player. I will use γ_t^i (β_t^i) to denote the posterior belief that player i is a learning (sure) type at time t given he has not conceded by time t. Obviously, $\gamma_0^i = \gamma_0$ and $\beta_0^i = 0$.

A strategy for the learning type player 1 (2) is denoted as $X^1(t)$ ($X^2(t)$) where $X^i(t)$ denotes the probability that player i concedes to player $-i$ by time t (inclusive). A strategy for the sure type player 1 (2) is denoted as $X^1(t;\tau)$ ($X^2(t;\tau)$) where $\tau \leq t$ is the time when the first Poisson signal is received.[3] Both $X^i(0)$ and $X^i(\tau;\tau)$ are allowed to be strictly positive such that player i concedes to player $-i$ immediately. I use Y^i to denote the combination of sure players' strategies: $Y^i(t) = (X^i(t;\tau))_{\tau \leq t}$; and Z^i to denote the overall strategy of player i: $Z^i = (X^i(\cdot), Y^i(\cdot))$. Z^i determines $F^i(t)$, which is player $-i$'s expected probability that player i concedes to player $-i$ by time t. Given player 2's strategy Z^2 and F^2 induced by Z^2, a normal player 1's expected payoff by conceding at time t (if player 1 is still learning at time t) is given by:

[3] There is a continuum of sure type players that is indexed by the arrival time of the first Poisson signal. The sure type players are not required to use the same strategy at any time t. There might be a continuum of equilibria by assigning different sure type players different concession rates. However, all of the equilibria are outcome equivalent in terms of the expected concession rate.

$$
\begin{aligned}
U^1(t, Z^1, Z^2) = &\int_{s<t} (v_H - cs)(p_0 e^{-\lambda s} + 1 - p_0) dF^2(s) \\
&\quad + \int_{s<t} p_0 \lambda e^{-\lambda s}(1 - F^2(s))(W^1(Z^1, Z^2; s) - cs) ds \\
&\quad + (M - ct)(F^2(t) - F^2(t-)) + (p_t v_L - ct)(p_0 e^{-\lambda t} + 1 - p_0)(1 - F^2(t)).
\end{aligned}
$$

Here $F^2(t-) = \lim_{\tau \nearrow t} F^2(\tau)$ and $W^1(Z^1, Z^2; s)$ denote the expected discounted value for player 1 who becomes sure at time s under the strategy profile (Z^1, Z^2).

Given F^i and player i has not conceded by time t, I can use $F^i(s|t)$ to denote the truncated probability of conceding after time t. The expected payoff for player 1 who is sure at time τ and concedes at time t is given by:

$$
\begin{aligned}
W^1(t, Z^2; \tau) = &\int_{\tau \leq s < t} (v_H - c(s - \tau)) dF^2(s|\tau) + (M - c(t - \tau))(F^2(t|\tau) - F^2(t - |\tau)) \\
&\quad + (v_L - c(t - \tau))(1 - F^2(t|\tau)).
\end{aligned}
$$

Similarly, the expected payoff for player 1 who is still learning at time τ and concedes at time t (if player 1 is still learning at time t) is given by:

$$
\begin{aligned}
U^1(t, Z^1, Z^2; \tau) = &\int_{\tau \leq s < t} (v_H - c(s - \tau))(p_\tau e^{-\lambda(s-\tau)} + 1 - p_\tau) dF^2(s|\tau) \\
&\quad + \int_{\tau \leq s < t} p_\tau \lambda e^{-\lambda(s-\tau)}(1 - F^2(s|\tau))(W^1(Z^1, Z^2; s) - c(s - \tau)) ds \\
&+ (M - c(t - \tau))(F^2(t|\tau) - F^2(t - |\tau)) + (p_t v_L - c(t - \tau))(p_\tau e^{-\lambda(t-\tau)} + 1 - p_\tau)(1 - F^2(t|\tau)).
\end{aligned}
$$

Finally, define $U^1(Z^1, Z^2; \tau)$ ($W^1(Z^1, Z^2; \tau)$) to be the learning (sure) type player 1's expected discounted value by playing the profile (Z^1, Z^2) after τ. Formally, $U^1(Z^1, Z^2; \tau)$ and $W^1(Z^1, Z^2; \tau)$ can be written as:

$$
U^1(Z^1, Z^2; \tau) = \int_{t \in [\tau, \infty]} U^1(t, Z^1, Z^2; \tau) dX^1(t|\tau)
$$

and

$$W^1(Z^1, Z^2; \tau) = \int_{t \in [\tau, \infty]} W^1(t, Z^2; \tau) dX^1(t; \tau).$$

$U^2(Z^1, Z^2; \tau)$ and $W^2(Z^1, Z^2; \tau)$ can be defined similarly. A strategy profile (Z^1, Z^2) is a sequential equilibrium if both $U^i(\cdot, Z^{-i}; \tau)$ and $W^i(\cdot, Z^{-i}; \tau)$ are maximized at any time τ when nobody has conceded yet.

Any candidate sequential equilibrium shares the following key features of the equilibrium without learning:

(1) A rational player will not delay conceding once he knows that his opponent will never concede.

(2) F^i (the expected distribution by i's opponent) is continuous and strictly increasing for $0 < t \leq T^i$, where T^i is the terminal time at which a normal type player i will concede for sure.

(3) At time 0, at most one of the two players will concede with a positive probability.

The first property means that a normal player will not delay conceding once he knows that his opponent will never concede. The most important property is the second one, which means that $F^i(t)$ cannot have jumps or be constant in a time interval. Also the second property implies that expected values $U^1(t, F^2; \tau)$ and $W^1(t, F^2; \tau)$ are continuous.

The proofs of the above properties are similar to the proofs provided in Abreu and Gul (2000) and hence are omitted. For both the sure and learning type players, there are three possibilities: strictly prefer conceding, strictly prefer staying, or indifference. There are nine different combinations in total. The next lemma shows that only three of them can happen in any equilibrium.

Lemma 3.1. *In any sequential equilibrium, at any time t such that a normal player is still possible to concede, only one of the following three cases is possible:*

(1) the learning type is indifferent between conceding and staying and the sure type strictly prefers conceding;

(2) the sure type is indifferent between conceding and staying and the learning type strictly prefers staying;

(3) the sure type strictly prefers conceding but the learning type strictly prefers staying.

Proof. First, I will show that the sure type player can never strictly prefer staying in any candidate equilibrium. Suppose on the contrary that there exists a time interval (t_1, t_2) such that the sure type player 1 strictly prefers staying for any $t \in (t_1, t_2)$. Define $t^\star$ to be the supremum of t such that player 1 strictly prefers staying for (t_1, t). $t^\star$ must be finite since there is a strictly positive probability for player 2 to stay forever. This implies that there exists $\bar{\eta} > 0$ such that for all $\eta < \bar{\eta}$, $W^1(t, F^2; \tau) > v_L$ for any $t \in (t^\star, t^\star + \eta)$ and $\tau \in (t^\star - \eta, t^\star)$. From the expression of $W^1(t, Z^2; \tau)$, it must be:

$$W^1(t, Z^2; \tau) = \int_{\tau \le s < t} (v_H - c(s-\tau)) dF^2(s|\tau) + (v_L - c(t-\tau))(1 - F^2(t|\tau)) > v_L.$$

For the learning type player, if he chooses to concede at time t regardless of whether he receives a signal, the expected payoff can be written as:

$$\begin{aligned} U^1(t, Z^2; \tau) &= \int_{\tau \le s < t} (v_H - c(s-\tau)) dF^2(s|\tau) + (v_L - c(t-\tau))(1 - F^2(t|\tau)) \\ &= p_\tau W^1(t, Z^2; \tau) + (1-p_\tau)\left(\int_{\tau \le s < t} (v_H - c(s-\tau)) dF^2(s|\tau) - c(t-\tau)(1 - F^2(t|\tau))\right). \end{aligned} \tag{3.1}$$

The second term is strictly positive since $W^1(t, Z^2; \tau) > v_L$. Therefore, it must be the case that $U^1(t, Z^2; \tau) > p_\tau W^1(t, Z^2; \tau) > p_\tau v_L$. This implies that the learning type player also prefers staying in a neighborhood left of $t^\star$. Then F^1 must be flat in a neighborhood left of $t^\star$, which contradicts the second property.

Also, the learning type player cannot strictly prefer conceding at any time τ. Suppose not and the learning type player strictly prefers conceding for $t \in (\tau, \bar{t})$. Then the sure type player has to randomize for $t \in (\tau, \bar{t})$. The expected payoff for a learning type player who concedes at t hence is given by:

$$U^1(t, Z^2; \tau) = \int_{\tau \le s < t} (v_H - c(s - \tau)) dF^2(s|\tau) + (p_\tau v_L - c(t - \tau))(1 - F^2(t|\tau)).$$

Equation (3.1) immediately implies that if $U^1(t, Z^2; \tau) < p_\tau v_L$, then $W^1(t, Z^2; \tau) < v_L$, which contradicts the fact that the sure type is indifferent at time t.

The above analysis leaves only three possibilities on the equilibrium path, which are listed in the lemma. □

The above lemma has very intuitive interpretations. Since the sure type player is more optimistic about the private payoff state than the learning type, the sure type has a higher incentive to concede. As a result, if the learning type is indifferent between conceding and staying, the sure type must strictly prefer conceding; if the sure type is indifferent between conceding and staying, the learning type must strictly prefer staying. In the benchmark model without learning, the normal type must always be indifferent between conceding and staying. Here, it is possible that neither the learning type nor the sure type is indifferent.

If the sure or learning type player $-i$ is indifferent between conceding and staying at time t, then F^i must be differentiable at time t. In particular, the expected concession rate $f^i_t = \frac{dF^i(t)/dt}{1-F^i(t)}$ must be $\frac{c}{v_H - v_L}$ if the sure type player $-i$ is indifferent and be $\frac{c}{v_H - p_t v_L}$ if the learning type player $-i$ is indifferent. If the expected concession rate is between those two numbers, then the sure type player $-i$ strictly prefers conceding while the learning type player $-i$ strictly prefers staying. Therefore, the expected equilibrium concession rate must be between $\frac{c}{v_H - p_t v_L}$ and $\frac{c}{v_H - v_L}$. Finally, if the normal type player $-i$ does the above, then F^{-i} is also differentiable such that the expected concession rate is $\lambda \gamma_t p_t$.

Slow Learning Case

Based on the previous lemma, I am able to show that in any sequential equilibrium with learning, there will be some periods of time such that the equilibrium concession rate in those periods is lower than the equilibrium concession rate without learning.

Lemma 3.2. *Fix any sequential equilibrium with learning. There exists $T < \infty$ such that normal players concede with probability one by time T. Also there exists $\epsilon > 0$ such that for all $t \in (T - \epsilon, T]$, the expected equilibrium concession rate $f_t = \frac{dF(t)/dt}{1-F(t)}$ is*

$$\frac{c(1 - p_0 + p_0 e^{-\lambda t})}{v_H(1 - p_0 + p_0 e^{-\lambda t}) - p_0 e^{-\lambda t} v_L}.$$

.

Proof. Suppose T is infinite. Then, with a positive probability, the normal type player has to stay forever and get a payoff of $-\infty$. This cannot be optimal. Therefore, T must be finite. Also T cannot be zero. If not, then both normal type players concede with probability one. This contradicts the third property of the sequential equilibrium. For $T > 0$, suppose the statement is not true and there exists $t_1 < T$ such that the expected equilibrium concession rate is strictly larger than

$$\frac{c(1 - p_0 + p_0 e^{-\lambda t})}{v_H(1 - p_0 + p_0 e^{-\lambda t}) - p_0 e^{-\lambda t} v_L}$$

for all $t \in (t_1, T]$. Notice that it is impossible for the learning type to concede with probability one by time t_1. This implies that the posterior belief γ_{t_1} must be strictly positive. However, if the expected equilibrium concession rate is strictly larger than

$$\frac{c(1 - p_0 + p_0 e^{-\lambda t})}{v_H(1 - p_0 + p_0 e^{-\lambda t}) - p_0 e^{-\lambda t} v_L}$$

for all $t \in (t_1, T]$, the learning type strictly prefers to stay. As a result, $\gamma_T > 0$ as well. But since the normal type players stop waiting at T, there must be a jump in F at time T, which leads to a contradiction. □

The above lemma implies that with exogenous learning, there always exist some periods such that the expected equilibrium concession rate is $\frac{c}{v_H - p_t v_L}$. When the learning rate is low, the following theorem shows that this is always the case for the unique sequential equilibrium.

Theorem 3.1. *If $\lambda \gamma_0 p_0 \leq \frac{c}{v_H - p_0 v_L}$, there exists a unique sequential equilibrium such that:*

(1) each learning type player concedes with probability zero at time 0 and at a positive rate between time 0 and T. The sure type player concedes with probability one upon receiving the first Poisson signal;

(2) T satisfies:

$$\gamma_0 = 1 - e^{-\frac{cT}{v_H}} \left[\frac{v_H - p_0 v_L}{(v_H - v_L)p_0 e^{-\lambda T} + (1 - p_0)v_H} \right]^{-\frac{cv_L}{\lambda v_H (v_H - v_L)}}.$$

After time T, only the commitment type player stays;

(3) for each player at time $t \in [0, T]$, the expected concession rate $f_t = \frac{dF(t)/dt}{1-F(t)}$ is

$$\frac{c(1 - p_0 + p_0 e^{-\lambda t})}{v_H(1 - p_0 + p_0 e^{-\lambda t}) - p_0 e^{-\lambda t} v_L}.$$

Proof. First, notice that at the beginning of the game, if only the sure type concedes with probability one, the expected concession rate is no more than $\lambda\gamma_0 p_0$.[4] The assumption that $\lambda\gamma_0 p_0 \leq \frac{c}{v_H - p_0 v_L}$ implies that if only the sure type concedes, the expected concession rate is lower than the minimum requirement of the equilibrium concession rate, which is $\frac{c}{v_H - p_0 v_L}$. Therefore, in any candidate sequential equilibrium, it must be the case that the sure type concedes immediately and the learning type randomizes at the beginning of the game. Suppose the sure type continues to concede with probability one until time τ. Then at time τ, the posterior beliefs are such that $\beta_\tau = 0$ and $\gamma_\tau < \gamma_0$.

If the learning type player stops randomizing at time τ, it must be case that at time τ, $\lambda\gamma_\tau p_\tau \geq \frac{c}{v_H - p_\tau v_L}$. However, for any $t < \tau$, the law of motion for $\gamma_t p_t(v_H - p_t v_L)$ satisfies:

$$\frac{d\gamma_t p_t(v_H - p_t v_L)}{dt} = \dot{\gamma}_t p_t(v_H - p_t v_L) - \gamma_t \lambda p_t(1 - p_t)(v_H - 2p_t v_L).$$

The first term is negative since $\dot{\gamma}_t < 0$ and the second term is negative because $v_H \geq 2v_L$. As a result, $\gamma_t p_t(v_H - p_t v_L)$ is strictly decreasing over time. There cannot exist any τ such

[4]If the learning type player concedes with probability zero at time zero, $\lambda\gamma_0 p_0$ is exactly the expected concession rate. But if the learning type player concedes with a strictly positive probability at time zero, the expected concession rate is less than $\lambda\gamma_0 p_0$ since the posterior is less than γ_0.

that $\lambda\gamma_\tau p_\tau \geq \frac{c}{v_H - p_\tau v_L}$. Therefore, on the equilibrium path, the learning type is always randomizing between time 0 and T.

Denote x_t to be the equilibrium concession rate of the learning type. The indifference condition implies that:

$$(\gamma_t x_t + \gamma_t \lambda p_t + \lambda p_t) p_t v_L = -c + \gamma_t x_t v_H + \gamma_t \lambda p_t v_H + \lambda p_t v_L - \lambda p_t (1 - p_t) v_L.$$

Also γ_t is updated by Bayes rule:

$$\dot{\gamma}_t = -(\lambda p_t + x_t)\gamma_t(1 - \gamma_t).$$

As a result, it is straightforward to derive an ODE about γ_t and solve γ_t as:

$$\frac{1 - \gamma_0}{1 - \gamma_t} = e^{-\frac{ct}{v_H}} \left[\frac{v_H - p_0 v_L}{(v_H - v_L) p_0 e^{-\lambda t} + (1 - p_0) v_H} \right]^{-\frac{c v_L}{\lambda v_H (v_H - v_L)}}.$$

T is chosen such that $\gamma_T = 0$ and hence T satisfies:

$$\gamma_0 = 1 - e^{-\frac{cT}{v_H}} \left[\frac{v_H - p_0 v_L}{(v_H - v_L) p_0 e^{-\lambda T} + (1 - p_0) v_H} \right]^{-\frac{c v_L}{\lambda v_H (v_H - v_L)}}.$$

The above calculation also suggests that it is impossible to have a learning type player conceding with strictly positive probability at time zero. If player i does that, then to guarantee that both normal players stop conceding at the same T, it must be the case that the learning type player $-i$ also concedes with a strictly positive probability at time zero. This contradicts the third property of the candidate equilibrium. □

The expected equilibrium concession rate is changing over time, which is different from the model without learning. In particular, it is relatively easier to reach an agreement initially but it becomes more and more difficult over time. Since $p_t < p_0$ for all $t > 0$, it is trivial to observe:

Corollary 3.1. *If $\lambda\gamma_0 p_0 \leq \frac{c}{v_H - p_0 v_L}$, compared to a model without learning, the exogenous learning increases the expected time of delay.*

The result is quite surprising in the sense that on the equilibrium path, the sure player who receives a Poisson signal will concede immediately. Then it seems that the exogenous learning should facilitate agreement. However, this intuitive thinking ignores the strategic response of the rational players. Learning as a martingale process can make the rational player both more and less optimistic about his private payoff state. The more optimistic player is more willing to concede, while the less optimistic player becomes less willing to concede. If the learning rate is low, the expected equilibrium concession rate is to make the less optimistic rational player indifferent. This implies that the rational players will overreact to this exogenous learning process and cause a longer delay.

Intermediate Learning Case

Exogenous learning increases delay when the learning rate is low. However, if λ is sufficiently large, the strategy profile described above is no longer an equilibrium. This is because if it is still an equilibrium for the sure player to concede immediately, then the expected concession rate is very high when λ is large. As a result, the learning type player can never be indifferent. In this section, I will construct an equilibrium when the learning rate is intermediate.

Theorem 3.2. *Suppose $\lambda\gamma_0 p_0 \in (\frac{c}{v_H - p_0 v_L}, \frac{c}{v_H - v_L})$, and the unique sequential equilibrium has the following feature: there exists $T_1 < T_2$ such that for $t \in (0, T_1)$, each learning type player concedes with probability zero and the sure type player concedes with probability one upon receiving the first Poisson signal; for $t \in (T_1, T_2)$, the sure type player still concedes with probability one upon receiving the first Poisson signal and the learning type player concedes at a positive rate.*

Proof. If no player concedes with strictly positive probability at time 0, it must be the case that the learning type player strictly prefers staying while the sure type player strictly prefers conceding since $\lambda\gamma_0 p_0 \in (\frac{c}{v_H - p_0 v_L}, \frac{c}{v_H - v_L})$. As shown in the proof of the previous theorem, $\gamma_t p_t(v_H - p_t v_L)$ is strictly decreasing over time under the assumption $v_H \geq 2v_L$. It is trivial to notice that $\gamma_t p_t$ is also strictly decreasing over time. Therefore, there exists T_1 such that $\lambda\gamma_{T_1} p_{T_1} = \frac{c}{v_H - p_{T_1} v_L}$. For $t < T_1$, $\lambda\gamma_t p_t \in (\frac{c}{v_H - p_t v_L}, \frac{c}{v_H - v_L})$ and hence the learning type player strictly prefers staying while the sure type player strictly prefers conceding. For $t > T_1$, $\lambda\gamma_t p_t < \frac{c}{v_H - p_t v_L}$ and the equilibrium is characterized by the previous theorem.

The final thing to prove is that it cannot be the case that a normal player concedes with a strictly positive probability at time 0. Suppose on the contrary that is the case. Player 1 concedes with a positive probability at time 0. Then this implies that player 2's strategy is such that the learning type of player is indifferent. This can only happen if the sure type of player 2 is randomizing at time 0 since $\lambda\gamma_0 p_0 > \frac{c}{v_H - p_0 v_L}$. However, it is impossible for player 1 to find a strategy such that the sure type of player 2 is indifferent since $\lambda\gamma_0 p_0 < \frac{c}{v_H - v_L}$. This leads to a contradiction. □

When the learning rate is in the intermediate region, the sure type players will concede for sure once they receive the Poisson signal. The learning type players will strictly prefer staying initially and begin to concede after some period. It is hard to tell directly whether delay increases compared to a model without learning. For $t \in (0, T_1)$, it is possible that the expected concession rate is strictly larger than $\frac{c}{v_H - p_0 v_L}$ if λ is sufficiently large. However, for $t > T_1$, the expected concession rate is strictly lower than $\frac{c}{v_H - p_0 v_L}$.

Also in the intermediate learning case, the impact of the learning rate on delay is ambiguous. When the sure player strictly prefers conceding and the learning type player strictly prefers staying, a larger λ increases the expected concession rate $\lambda\gamma_t p_t$. However, when the learning type player is randomizing, a larger λ leads to a lower expected concession rate.

Fast Learning Case

If $\lambda\gamma_0 p_0 > \frac{c}{v_H - p_0 v_L}$, then the learning type players randomize at the beginning of the game. The unique sequential equilibrium in this fast learning case may have two different possibilities. In the first possible equilibrium, there exists $T_1 < T_2 < T_3$ such that for $t \in (0, T_1)$, each learning type player concedes with probability zero while the sure type player concedes with a strictly positive probability upon receiving the first Poisson signal and with a positive rate afterwards; for $t \in (T_1, T_2)$, each learning type player concedes with probability zero while the sure type player concedes with probability one upon receiving the first Poisson signal; for $t \in (T_2, T_3)$, each learning type player concedes with a positive rate while the sure type player still concedes with probability one upon receiving the first Poisson signal. In the second possible equilibrium,there exists $T_1 < T_2$ such that for $t \in (0, T_1)$, each learning type player concedes with probability zero while the sure type player concedes with a strictly positive probability upon receiving the first Poisson signal and with a positive rate afterwards; for $t \in (T_1, T_2)$, the learning type player concedes at a positive rate while the sure type player concedes with probability one upon receiving the first Poisson signal.

Notice that at time t such that $\lambda\gamma_t p_t = \frac{c}{v_H - p_t v_L}$, the sure type players cannot switch to strictly prefer conceding immediately. This is because there is a positive probability to be a sure type player at time t and the distribution F^i cannot have jumps. The sure type players will continue to randomize until the posterior belief to be a sure type player reaches zero. There are two possibilities at this point in time t'. In particular, it might be the case that $\lambda\gamma_{t'} p_{t'} < \frac{c}{v_H - p_{t'} v_L}$ and hence the equilibrium immediately jumps to the phase where the learning type player is randomizing.

In both types of equilibria, the expected concession rate initially is $\frac{c}{v_H - v_L}$, which is higher than the expected concession rate $\frac{c}{v_H - p_0 v_L}$ without learning. But eventually, the expected concession rate will drop to $\frac{c}{v_H - p_t v_L}$. The explicit expression for expected delay is hard to derive since the expected concession rate is changing over time. However, I can fully characterize the expected delay in the limiting case where $\lambda = \infty$. $\lambda = \infty$ corresponds to the immediate revelation case, where the normal type player i starts with two possible private types: either $v_i = v_L$ or $v_i = 0$. Each player i knows exactly what v_i is but his opponent does not know. The initial beliefs are such that $v_i = v_L$ with probability $\gamma_0 p_0$ and $v_i = 0$ with probability $\gamma_0(1 - p_0)$. The next result shows that in this limiting case, the expected delay is longer than in the case without learning.

Theorem 3.3. *Fix any pair $(\gamma_0, p_0) \in (0,1)^2$, if $\lambda = \infty$, conditional on at least one of the two players being normal, the longest delay is higher than the longest delay without learning and the expected delay is longer than the expected delay without learning.*

Proof. The prior beliefs are such that $v_i = v_L$ with probability $\gamma_0 p_0$ and $v_i = 0$ with probability $\gamma_0(1-p_0)$. The unique sequential equilibrium has the following feature: the normal type players with $v_i = v_L$ will randomize first and the $v_i = 0$ players will strictly prefer to stay. After some time T_1, the $v_i = v_L$ players concede with probability one and then the $v_i = 0$ players begin to randomize for T_2 length of time. The expected concession rate is $\frac{c}{v_H - v_L}$ before $T_1$and $\frac{c}{v_H}$ after T_1.

Notice that if no concession takes place before T_1, the posterior beliefs are such that with probability $\frac{\gamma_0(1-p_0)}{1-\gamma_0 p_0} > \gamma_0(1-p_0)$, each player is normal. The longest delay is:

$$T^{\infty} = T_1 + T_2 = -\frac{(v_H - v_L)\log(1-\gamma_0 p_0)}{c} - \frac{v_H \log \frac{1-\gamma_0}{1-\gamma_0 p_0}}{c}.$$

The longest delay without learning is $T = -\frac{(v_H - p_0 v_L)\log(1-\gamma_0)}{c}$. Obviously, $T^{\infty} > T$ since $\log(1-\gamma_0 p_0) > p_0 \log(1-\gamma_0)$.

For the expected delay, I have to consider two different cases. Conditional on one of the two players being normal, with probability $\frac{(2-\gamma_0 p_0)p_0}{2-\gamma_0}$, at least one of the two players has $v_i = v_L$; with probability $\frac{2\gamma_0(1-\gamma_0)(1-p_0)+\gamma_0^2(1-p_0)^2}{2\gamma_0-\gamma_0^2}$, neither of the two players have $v_i = v_L$ but at least one has $v_i = 0$.

The expected delay is given by:

$$\Omega^{\infty} = \frac{2(1-\gamma_0 p_0)\gamma_0 p_0}{2\gamma_0 - \gamma_0^2} \int_0^{T_1} td \frac{1-e^{-\frac{ct}{v_H - v_L}}}{\gamma_0 p_0} + \frac{\gamma_0^2 p_0^2}{2\gamma_0 - \gamma_0^2} \int_0^{T_1} td \left[1 - (1 - \frac{1-e^{-\frac{ct}{v_H - v_L}}}{\gamma_0 p_0})^2 \right]$$
$$+ \frac{2\gamma_0(1-\gamma_0)(1-p_0) + \gamma_0^2(1-p_0)^2}{2\gamma_0 - \gamma_0^2} T_1 + \frac{2\gamma_0(1-\gamma_0)(1-p_0)}{2\gamma_0 - \gamma_0^2} \int_0^{T_2} td \frac{(1-\gamma_0 p_0)(1-e^{-\frac{ct}{v_H}})}{\gamma_0(1-p_0)}$$
$$+ \frac{\gamma_0^2(1-p_0)^2}{2\gamma_0 - \gamma_0^2} \int_0^{T_2} td \left[1 - (1 - \frac{(1-\gamma_0 p_0)(1-e^{-\frac{ct}{v_H}})}{\gamma_0(1-p_0)})^2 \right].$$

The above expression can be simplified as:

$$\Omega^{\infty} = \frac{v_H - v_L}{c(2\gamma_0 - \gamma_0^2)} [\frac{1}{2}(1 - (1-\gamma_0 p_0)^2) + \log(1-\gamma_0 p_0)(1-\gamma_0)^2]$$
$$+ \frac{v_H}{c(2\gamma_0 - \gamma_0^2)} [\frac{1}{2}((1-\gamma_0 p_0)^2 - (1-\gamma_0)^2) + \log \frac{1-\gamma_0}{1-\gamma_0 p_0}(1-\gamma_0 p_0)^2]. \quad (3.2)$$

The expected delay without learning is:

$$\Omega = \frac{v_H - p_0 v_L}{c(2\gamma_0 - \gamma_0^2)} [\frac{1}{2}(1 - (1-\gamma_0)^2) + \log(1-\gamma_0)(1-\gamma_0)^2]. \quad (3.3)$$

It is straightforward to observe that for any fixed $\gamma_0 > 0$, Ω is linear in p_0 while Ω^{∞} is concave in p_0. Ω and Ω^{∞} coincide when p_0 is either 0 or 1. Therefore, $\Omega^{\infty} > \Omega$ for any pair $(\gamma_0, p_0) \in (0,1)^2$. □

Since $\frac{c}{v_H - p v_L}$ is convex in p: $p\frac{c}{v_H - v_L} + (1-p)\frac{c}{v_H} > \frac{c}{v_H - p v_L}$, it seems that the expected concession rate is higher if we can fully separate the $v_i = v_L$ and $v_i = 0$ players. Then intuitively, letting $\lambda = \infty$ will increase the expected concession rate and hence decrease delay. The intuition is wrong because it ignores another channel affecting delay. Since in equilibrium, the more optimistic player always concedes first, at the time when the $v_i = v_L$ players concede with probability one, the posterior belief that $v_i = 0$ increases from $\gamma_0(1-p_0)$ to $\frac{\gamma_0(1-p_0)}{1-\gamma_0 p_0}$. This increase in the posterior also leads to a longer delay. If the players think naively and do not update beliefs at time T_1 (i.e., at $t = T_1$, $\gamma_t = \gamma_0(1-p_0)$), the longest delay

$$\tilde{T}^{\infty} = T_1 + T_2 = -\frac{(v_H - v_L)\log(1-\gamma_0 p_0)}{c} - \frac{v_H \log(1-\gamma_0(1-p_0))}{c}$$

is lower than the longest delay without learning.

In summary, compared to a model without learning, there are three factors affecting delay in the limiting case of $\lambda = \infty$. First, the expected concession rate before T_1 is higher than the expected concession rate without learning, which leads to a shorter delay. Second, the expected concession rate after T_1 is lower than the expected concession rate without learning, which leads to a longer delay. Third, since the more optimistic ($v_i = v_L$) players concede first, the posterior belief that a player is less optimistic ($v_i = 0$) is increasing over time. The last effect implies that more weight has to be put on the lower expected concession rate, which also increases delay. The above analysis shows that the first effect dominates the second effect but is dominated by the combination of the second and third effects. Hence, the change of posterior beliefs is an important driving force leading to a longer delay.

For an arbitrary learning rate, it is hard to get an explicit solution for the longest delay and expected delay. But the idea of the above proof can be generalized to get a lower bound on expected delay. The next result shows that even this lower bound is longer than the expected delay without learning.

Theorem 3.4. *Fix any pair $(\gamma_0, p_0) \in (0,1)^2$ and any learning rate λ, conditional on at least one of the two players being normal, the longest delay with learning is higher than the longest delay without learning and the expected delay with learning is longer than the expected delay without learning.*

Proof. Suppose the learning type players begin to concede at time t. Before time t, only the sure type players concede. The probability of conceding before time t is x. Feasibility requires that $x \in [0, \gamma_0 p_0]$. This implies that at time t, the posterior beliefs are such that: $p_t = \frac{\gamma_0 p_0 - x}{\gamma_0 - x}$ and $\gamma_t = \frac{\gamma_0 - x}{1-x}$. Before time t, an upper bound for the concession rate is $\frac{c}{v_H - v_L}$; after time t, an upper bound for the concession rate is $\frac{c}{v_H - p_t v_L}$. Therefore, a lower bound on the longest delay is given by:

$$\hat{T} = -\frac{(v_H - v_L)\log(1-x)}{c} - \frac{(v_H - \frac{\gamma_0 p_0 - x}{\gamma_0 - x} v_L)\log\frac{1-\gamma_0}{1-x}}{c}$$
$$= \frac{-v_H \log(1-\gamma_0)}{c} + \frac{v_L \log(1-x)}{c} + \frac{\frac{\gamma_0 p_0 - x}{\gamma_0 - x} v_L \log\frac{1-\gamma_0}{1-x}}{c}. \quad (3.4)$$

The longest delay when there is no learning is $T = -\frac{(v_H - p_0 v_L)\log(1-\gamma_0)}{c}$. The difference $\hat{T} - T$ is proportional to $\Delta(x) = \gamma_0 \log(1-x) - x\log(1-\gamma_0)$. The first derivative of $\Delta(x)$ is $\frac{-\gamma_0}{1-x} - \log(1-\gamma_0)$, which is decreasing in x. It is trivial to observe that $\Delta'(0) > 0$ but $\Delta'(x)$ could be negative if x is sufficiently large. $\Delta(x)$ possibly first increases in x and then decreases in x. Since $\Delta(0) = 0$ and $\Delta(\gamma_0 p_0) > 0$, it must be the case that $\Delta(x) \geq 0$ for all $x \in [0, \gamma_0 p_0]$. Therefore, $\hat{T} \geq T$ for sure.

Similarly, a lower bound on the expected delay is given by:

$$\hat{\Omega}(x) = \frac{v_H - v_L}{c(2\gamma_0 - \gamma_0^2)}[\frac{1}{2}(1-(1-x)^2) + \log(1-x)(1-\gamma_0)^2]$$
$$+ \frac{v_H - \frac{\gamma_0 p_0 - x}{\gamma_0 - x} v_L}{c(2\gamma_0 - \gamma_0^2)}[\frac{1}{2}((1-x)^2 - (1-\gamma_0)^2) + \log\frac{1-\gamma_0}{1-x}(1-x)^2]. \quad (3.5)$$

Notice that the first derivative of $\hat{\Omega}$ is given by:

$$\hat{\Omega}'(x) = \frac{v_H - v_L}{c(2\gamma_0 - \gamma_0^2)}[1 - x - \frac{(1-\gamma_0)^2}{1-x}]$$
$$+ \frac{\gamma_0(1-p_0)v_L}{c(2\gamma_0 - \gamma_0^2)}[\frac{1}{2}((1-x)^2 - (1-\gamma_0)^2) + \log\frac{1-\gamma_0}{1-x}(1-x)^2]$$
$$+ \frac{v_H - \frac{\gamma_0 p_0 - x}{\gamma_0 - x} v_L}{c(2\gamma_0 - \gamma_0^2)} 2(1-x)\log\frac{1-x}{1-\gamma_0}. \quad (3.6)$$

On the RHS of the above equation, the second term is positive for sure. The first and third terms are positive because $x \leq \gamma_0 p_0$. Therefore, $\hat{\Omega}(x)$ is increasing in x. And the expected delay when there is no learning is

$$\Omega = \frac{v_H - p_0 v_L}{c(2\gamma_0 - \gamma_0^2)}[\frac{1}{2}(1-(1-\gamma_0)^2) + \log(1-\gamma_0)(1-\gamma_0)^2].$$

Obviously, $\hat{\Omega}(0) = \Omega$. Then it must be the case that $\hat{\Omega}(x) \geq \Omega$ for all $x \in [0, \gamma_0 p_0]$. $\square$

The basic idea of the above proof is that in any equilibrium, it is possible to divide the equilibrium into two phases. In the first phase, the learning type players strictly prefer staying; in the second phase, the learning type players is randomizing. The concession rate in the first phase may be as high as $\frac{c}{v_H - v_L}$. But compared to a model without learning, it is more difficult to reach an agreement in the second phase. Delay increases from two possible channels. One is the Bayesian updating process which increases the posterior belief of being a learning type; the other is the equilibrium concession rate becomes $\frac{c}{v_H - p_t v_L}$. The combination of these two effects in the second phase will completely offset the possible decrease in delay in the first phase.[5] As a result, the expected delay will always be increasing instead of decreasing.

3.4.2 One-Sided Learning

Another interesting situation is one which only player 1 is able to learn. Player 2 has no access to the exogenous learning process. Then, at any time t, player 1 has three possible types: a sure type who is sure that $v = v_L$, a learning type who is still unsure and a commitment type. I use γ_{1t} to denote the posterior belief that player 1 is a learning type, β_{1t} to denote the posterior belief that player 1 is a rational type, p_t to denote player 1's posterior belief that $v = v_L$ given he is a sure type at time t and finally γ_{2t} to denote the belief that player 2 is a normal type.

The next result shows that conditional on at least one of the two players being normal, the one-sided learning model has the same longest delay as the model without learning.

Theorem 3.5. *In the one-sided learning model, conditional on at least one of the two players being normal, the longest delay is always* $T = -\frac{(v_H - p_0 v_L)\log(1-\gamma_0)}{c}$.

[5] The change in posterior beliefs also plays an important role here. If the players update beliefs naively, then the lower bound on longest delay with learning is

$$\tilde{T} = -\frac{(v_H - v_L)\log(1-x)}{c} - \frac{(v_H - \frac{\gamma_0 p_0 - x}{\gamma_0 - x} v_L)\log(1-\gamma_0+x)}{c},$$

which could be less than the longest delay without learning for x close to $\gamma_0 p_0$.

Proof. To prove the theorem, I need to consider two separate cases $\lambda\gamma_0 p_0 \leq \frac{c}{v_H - p_0 v_L}$ and $\lambda\gamma_0 p_0 > \frac{c}{v_H - p_0 v_L}$. The unique sequential equilibrium when $\lambda\gamma_0 p_0 \leq \frac{c}{v_H - p_0 v_L}$ is characterized by the following proposition.

Proposition 3.2. *If $\lambda\gamma_0 p_0 \leq \frac{c}{v_H - p_0 v_L}$, there exists a unique sequential equilibrium in the one-sided learning model satisfying:*

(1) for player 1, the learning type concedes with probability zero at time 0 and at a positive rate between time 0 and T; the sure type player concedes with probability one upon receiving the first Poisson signal;

(2) the normal type player 2 concedes with strictly positive probability at time 0 and at a positive rate between time 0 and T;

(3) $T = -\frac{(v_H - p_0 v_L)\log(1-\gamma_0)}{c}$ and after time T, only the commitment type player stays;

(4) at time $t \in (0, T]$, player 1's expected concession rate is $\frac{c}{v_H - p_0 v_L}$ and player 2's expected concession rate is

$$\frac{c(1 - p_0 + p_0 e^{-\lambda t})}{v_H(1 - p_0 + p_0 e^{-\lambda t}) - p_0 e^{-\lambda t} v_L}.$$

Proof. The proof of the equilibrium properties is very similar to the proof in the two-sided learning model and is omitted. The assumption that $\lambda\gamma_0 p_0 \leq \frac{c}{v_H - p_0 v_L}$ guarantees that the learning type player 1 must randomize and the sure type player 1 must concede immediately in equilibrium. Therefore, denote x_{1t} (x_{2t}) to be the equilibrium concession rate of the learning type player 1 (normal type player 2). The indifference conditions imply:

$$(\gamma_{2t} x_{2t} + \lambda p_t) p_t v_L = -c + \gamma_{2t} x_{2t} v_H + \lambda p_t v_L - \lambda p_t (1 - p_t) v_L.$$

and

$$(\gamma_{1t} x_{1t} + \gamma_{1t} \lambda p_t) p_0 v_L = -c + \gamma_{1t} x_{1t} v_H + \gamma_{1t} \lambda p_t v_H.$$

γ_{1t} and γ_{2t} evolve as:

$$\dot{\gamma}_{1t} = -(\lambda p_t + x_{1t})\gamma_{1t}(1-\gamma_{1t}) \quad \text{and} \quad \dot{\gamma}_{2t} = -x_{2t}\gamma_{2t}(1-\gamma_{2t}).$$

Therefore, we have:

$$\dot{\gamma}_{1t} = -\frac{c}{v_H - p_0 v_L}(1-\gamma_{1t}) \quad \text{and} \quad \dot{\gamma}_{2t} = -\frac{c}{v_H - p_t v_L}(1-\gamma_{2t}).$$

Since $p_t = \frac{p_0 e^{-\lambda t}}{p_0 e^{-\lambda t} + 1 - p_0} < p_0$, the expected concession rate of player 2 is smaller than the expected concession rate of player 1. Also the learning type of player 1 and the normal type of player 2 have to stop conceding at the same time T. As a result, the normal type of player 2 has to concede with a strictly positive probability at time 0.

T is determined by the shortest time of concession $T = -\frac{(v_H - p_0 v_L)\log(1-\gamma_0)}{c}$ and γ_{2t} satisfies:

$$\frac{1-\gamma_0'}{1-\gamma_{2t}} = e^{-\frac{ct}{v_H}} \left[\frac{v_H - p_0 v_L}{(v_H - v_L)p_0 e^{-\lambda t} + (1-p_0)v_H}\right]^{-\frac{c v_L}{\lambda v_H (v_H - v_L)}}.$$

At time 0, the probability of concession by the normal type of player 2 is chosen such that:

$$\gamma_0' = 1 - e^{-\frac{cT}{v_H}} \left[\frac{v_H - p_0 v_L}{(v_H - v_L)p_0 e^{-\lambda T} + (1-p_0)v_H}\right]^{-\frac{c v_L}{\lambda v_H (v_H - v_L)}}.$$

□

If $\lambda \gamma_0 p_0 \leq \frac{c}{v_H - p_0 v_L}$, then initially the sure type of player 1 will randomize such that the expected concession rate is always $\frac{c}{v_H - p_0 v_L}$. Then at time 0, it is impossible for player 1 to concede with a positive probability. Next, I will show that there cannot exist two disjoint time intervals (t_0, t_1) and (t_2, t_3) such that the sure type is indifferent on both intervals and strictly prefers conceding for $t \in (t_1, t_2)$. If there exists such an equilibrium, at t_2, it must be the case that: $\lambda \gamma_{t_2} p_{t_2} \geq \frac{c}{v_H - p_0 v_L}$. Since $\gamma_{t_1} p_{t_1} > \gamma_{t_2} p_{t_2}$, $\lambda \gamma_{t_1} p_{t_1} > \frac{c}{v_H - p_0 v_L}$. Therefore, if a sure type player concedes with probability one at t_1, his normal opponent must stay for sure, which leads to a contradiction. Therefore, on the equilibrium path, the sure type will first randomize until T_1 and the learning type randomizes afterwards.

At $t < T_1$, denote γ_t to be the belief that player 1 is a learning type at time t and β_t to be the belief that a player is a sure type at time t. Suppose the existing sure type has a concession rate of x_t and a new sure type will concede with probability y_t. The indifference of player 2 means that:

$$(\beta_t x_t + \gamma_t \lambda p_t y_t) p_0 v_L = -c + (\beta_t x_t + \gamma_t \lambda p_t y_t) v_H.$$

The laws of motion for β_t and γ_t are such that:

$$\dot{\beta}_t = -x_t \beta_t (1 - \beta_t) - \gamma_t \lambda p_t y_t (1 - \beta_t) + \gamma_t \lambda p_t$$

and

$$\dot{\gamma}_t = -\gamma_t \lambda p_t + \gamma_t (\beta_t x_t + \gamma_t \lambda p_t y_t).$$

The above equations imply that:

$$\dot{\beta}_t + \dot{\gamma}_t = -\frac{c}{v_H - p_0 v_L}(1 - \beta_t - \gamma_t)$$

and hence

$$\beta_t + \gamma_t = 1 - (1 - \gamma_0) e^{\frac{c}{v_H - p_0 v_L}}.$$

Notice for $t \geq T_1$, $\beta_t = 0$ and the expected concession rate for the learning type of player is also $\frac{c}{v_H - p_0 v_L}$. Therefore, beginning from $\beta_0 + \gamma_0 = \gamma_0$, $\beta_t + \gamma_t$ satisfies:

$$\beta_t + \gamma_t = 1 - (1 - \gamma_0) e^{\frac{c}{v_H - p_0 v_L}}$$

for all $t \geq 0$. There is also no discontinuity in $\beta_t + \gamma_t$ for any $t > 0$. As a result, it must be the case that $T = -\frac{(v_H - p_0 v_L) \log(1 - \gamma_0)}{c}$. $\square$

The above result implies that allowing only one player to learn is better than allowing both players to learn in terms of delay regardless of what the initial parameters are. Compared to a model without learning, one-sided learning does not increase the longest time of waiting if either player 1 or player 2 is normal. In particular, the expected equilibrium concession rate of player 1 is exactly the same as the case without learning.

3.5 Endogenous Information Acquisition

This section briefly discusses the implications of the above results on endogenous information acquisition. In particular, following the setup in Bonatti and Hörner (2009), I assume that a player can achieve arrival rate λ with flow cost $c(\lambda)$ with $c(0) = 0$ and $c'(\cdot) > 0$. The information acquisition decision is made at every instant of time given there is no concession by time t. Formally, normal player i's information acquisition decision is denoted as $c^i : [0, \infty) \times \{0, 1\} \to [0, \bar{\lambda}]$. 0 means player i has not yet received any Poisson signal, and 1 means player i has received at least one signal. Obviously, if player i has received one signal at time τ, then $c_{it} = 0$ for all $t \geq \tau$. $\bar{\lambda}$ is the maximum achievable learning rate. Given the information acquisition strategy, the total cost of information acquisition from time 0 to t is given by $C^i(t) = \int_0^t e^{-rs} c^i_s dt$. Also define $\Lambda^i(t) = \int_0^t \lambda^i_s dt$.

Given player 2's strategy Z^2, a normal player 1's expected payoff by conceding at time t is given by:

$$\begin{aligned} U^1(t, Z^1, Z^2) = &\int_{s<t} (v_H - cs - C^1(s))(p_0 e^{-\Lambda^1(s)} + 1 - p_0) dF^2(s) \\ &+ \int_{s<t} (W^1(Z^1, Z^2; s) - cs - C^1(s)) p_0 \lambda^1_s e^{-\Lambda^1(s)} (1 - F^2(s)) ds \\ &+ (M - ct - C^1(t))(F^2(t) - F^2(t-)) \\ &+ (1 - F^2(t))(p_t v_L - ct - C^1(t))(p_0 e^{-\Lambda^1(t)} + 1 - p_0). \end{aligned}$$

Player 2's expected payoff can be defined similarly.

The paper shows that when the maximum achievable learning rate is not high enough, the unique sequential equilibrium is such that no player acquires information on the equilibrium path. Then the unique sequential equilibrium is the same as the equilibrium in the no learning case.

Proposition 3.3. *If the maximum achievable learning rate $\bar{\lambda}$ satisfies:*

$$\bar{\lambda} \leq \frac{c}{\gamma_0 p_0 (v_H - p_0 v_L)},$$

then the unique sequential equilibrium is such that each player chooses $\lambda_t^i = 0$ almost everywhere.

Proof. Suppose the statement is not true. Then there exists a time interval $[t_1, t_2]$ such that at least of one the two normal type players begins to acquire information at time t_1 and then stops at t_2:

$$\int_{t_1}^{t_2} c(\lambda_s) ds > 0.$$

Since the player has not acquired any information before time t_1, $p_{t_1} = p_0$. The assumption $\bar{\lambda} < \frac{c}{\gamma_0 p_0 (v_H - p_0 v_L)}$ guarantees that for $t \in [t_1, t_2]$, the learning type has to randomize between conceding and staying. This implies that if the learning type player i concedes at t_2, the expected payoff at t_1 can be written as:

$$\begin{aligned} & p_0 v_L - \int_{s<t_2} C^i(s)(p_0 e^{-\Lambda^i(s)} + 1 - p_0) dF^{-i}(s|t_1) \\ & \quad - \int_{s<t_2} C^i(s) p_0 \lambda_s^i e^{-\Lambda^i(s)} (1 - F^{-i}(s|t_1)) ds - (1 - F^{-i}(t_2|t_1))(p_0 e^{-\Lambda^i(t_2)} + 1 - p_0) C^i(t_2). \end{aligned}$$

Obviously, the expected value of playing the war of attrition game at t_1 is always $p_0 v_L$ regardless of whether this player acquires information or not. Therefore, the learning type has no incentive to acquire information and the equilibrium arrival rate is zero almost everywhere in the endogenous learning model. □

3.6 Conclusion

Delay is a pervasive phenomenon in bargaining and voting environments. It is natural to ask whether there is any way to reduce delay since delay is usually costly. This paper develops a continuous-time incomplete information war of attrition model with private learning investigate whether delay will become shorter if there is an exogenous information flow that facilitates an agreement. It turns out that this exogenous private learning makes delay longer instead of shorter. Also, to minimize delay, it is better to allow one player to learn than to allow both to learn. The result that private learning may lead to a longer delay is quite robust to some changes in the model specifications. For example, similar results can be derived if the Poisson signal is such that it reveals $v_i = 0$ for sure, or exogenous learning is about the winning payoff v_H instead of the losing payoff. The key insight is that this private Bayesian learning is a martingale process and generates multiple normal types. Due to learning, it is always possible for a normal player to become less optimistic about the payoff state over time. In equilibrium, there must exist some periods such that the less optimistic players are randomizing. Compared to the benchmark model without learning, the concession rate in these periods will be smaller and the expected delay will be longer.

Appendix A

Appendices

A.1 Appendix to Chapter 1

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.[1] The monopolist's pricing decision is given by the mapping:

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

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

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 *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 A.1. *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 A.1. In proposition A.1, 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 (1.1) and (1.2) to update beliefs.

Definition A.1. *A Markovian strategy profile (P, α) is* strongly admissible *in the good news case if it satisfies:*[2]

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 $||\boldsymbol{\rho}' - \boldsymbol{\rho}|| \leq \epsilon$;*[3]

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

[3] 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.

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 A.1 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 (1.1) and (1.2) 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 an infinite number of changes on any time interval. This additional condition is needed to simplify the analysis of the equilibrium.

Definition A.1 is too strong in the sense that even cutoff strategies may not be strongly admissible.[4] 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

[4]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.

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 A.1.

Definition A.2. *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.

Observation A.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.[5]

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

B.1 Proof of Proposition 1.1

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

Claim A.1. *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:

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

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

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. $\square$

Notice in equation

$$rn\rho q(\rho)g + n\rho q(\rho)\lambda_H \Omega_1(\rho) = (r + n\rho q(\rho)\lambda_H)ns, \tag{A.1}$$

$\Omega_1(\cdot)$ 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 A.2. *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: $\rho_S^e > \rho_I^e$.*

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

$$rn\rho q(\rho)g + n\rho q(\rho)\lambda_H[g + (n-1)W(\rho)] = (r + n\rho q(\rho)\lambda_H)ns. \tag{A.2}$$

By contradiction, assume $\rho_S^e \leq \rho_I^e$ and $W(\rho_S^e) = s$ by definition. Equation (A.2) 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 $\rho_S^e \leq \rho_I^e$. Therefore, it must be true that $\rho_S^e > \rho_I^e$ and thus $W(\rho_S^e) > s$. $\square$

From claim A.2, ρ_S^e should satisfy equation (A.2) where $q(\rho_S^e)$ is given by equation (1.6). Given the priors, the efficient cutoffs $(\rho_S^e(\rho_0, q_0), q_S^e(\rho_0, q_0))$ can be solved jointly:

$$q_S^e = \frac{rs}{\rho_S^e[(r+\lambda_H)g + (n-1)\lambda_H W(\rho_S^e) - n\lambda_H s]}. \tag{A.3}$$

$$q_S^e = \frac{(1-\rho_0)^n q_0}{(1-\rho_0)^n q_0 + (1-\rho_S^e)^n(1-q_0)}. \tag{A.4}$$

Clearly, $W(\rho_S^e)$ is increasing in ρ_S^e and thus q_S^e is decreasing in ρ_S^e from equation (A.3). Equation (A.4) 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 (A.4) and (A.3) is unique. Equation (A.3) 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 $\rho_S^e > \rho_I^e$. $\square$

B.2 Proof of Proposition 1.2

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

$$rJ_I(\rho) = \max\left\{r(g-s), 2r(g\rho - s) + \rho\lambda_H(2(g-s) - J_I(\rho)) - \lambda_H\rho(1-\rho)J_I'(\rho)\right\}. \quad \text{(A.5)}$$

On the RHS of equation (A.5), $g-s$ 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:

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

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

$$rV_I(\rho) = r(g(1-\rho)+s) + \rho\lambda_H(s - V_I(\rho)) - \lambda_H\rho(1-\rho)V_I'(\rho) \quad \text{(A.6)}$$

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 (A.6) 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. $\square$

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

Lemma A.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 $\rho > \rho_S^\star$,*

$$\begin{aligned}\lim_{h\to 0}\frac{U_S(\rho)-\hat{U}(\rho;h)}{h} &= 2(r+\lambda_H\rho q)(U_S(\rho)-s)+\lambda_H\rho(1-\rho)U_S'(\rho)\\ &\quad + (r+\lambda_H\rho)g(1-\rho)q(\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)})^{r/\lambda_H}-\lambda_H g\rho(1-\rho)q\\ -\left[\frac{r+\lambda_H\rho_S^\star}{1-\rho_S^\star}(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H}-\lambda_H(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H}\right]&g(1-\rho)^2q(\frac{1-\rho}{\rho})^{r/\lambda_H}. \end{aligned} \tag{A.7}$$

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

$$\begin{aligned}\lim_{h\to 0}\frac{U_S(\rho)-\hat{U}(\rho;h)}{h} &= 2(r+\lambda_H\rho q)(U_S(\rho)-s)+\lambda_H\rho(1-\rho)U_S'(\rho)\\ &\quad -\frac{rg}{r+\lambda_H}\lambda_H\rho(1-\rho)q+\frac{r\lambda_H g}{r+\lambda_H}\frac{\rho_S^\star(1-\rho)^2q}{1-\rho_S^\star}(\frac{(1-\rho)\rho_S^\star}{\rho(1-\rho_S^\star)})^{r/\lambda_H};\end{aligned} \tag{A.8}$$

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

$$\begin{aligned}\lim_{h\to 0}\frac{U_S(\rho)-\hat{U}(\rho;h)}{h} &= 2(r+\lambda_H\rho q)(U_S(\rho)-s)+\lambda_H\rho(1-\rho)U_S'(\rho)\\ &\quad + (r+\lambda_H\rho)g(1-\rho)q(\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)})^{r/\lambda_H}-\lambda_H g\rho(1-\rho)q\\ -r\left[\frac{r+\lambda_H+\lambda_H\rho_I^\star}{(r+\lambda_H)(1-\rho_I^\star)}(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H}-\frac{\lambda_H}{r+\lambda_H}(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H}\right]&g(1-\rho)^2q(\frac{1-\rho}{\rho})^{r/\lambda_H}. \end{aligned} \tag{A.9}$$

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:

$$\lim_{h\to 0}\frac{U_S(\rho)-\hat{U}(\rho;h)}{h} = (r+\lambda_H\rho q(\rho))(U_S(\rho)-s)+\lim_{h\to 0}\frac{U_S(\rho)-U^D(\rho,\rho_h)}{h}. \tag{A.10}$$

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.[6] Therefore, there exists $\bar{h}'$ such that for all $h' \leq \bar{h}'$, we have:

$$\begin{aligned} U^D(\rho, \rho_h) = \mathbb{E} \int_{t=0}^{h'} & re^{-rt}(\rho_t q_t g - \tilde{P}_t)dt \\ & + \rho\tilde{q}_h(1 - e^{-\lambda_H h'})e^{-rh'}V_I(\rho_{h+h'}) + \rho_h\tilde{q}_h(1 - e^{-\lambda_H h'})e^{-rh'}s \\ & \quad + [1 - \rho\tilde{q}_h(1 - e^{-\lambda_H h'}) - \rho_h\tilde{q}_h(1 - e^{-\lambda_H h'})]e^{-rh'}U(\rho_{h'}, \rho_{h+h'}). \quad \text{(A.11)} \end{aligned}$$

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

$$\begin{aligned} U^{ND}(\rho, \rho_h) = \mathbb{E} \int_{t=0}^{h'} & re^{-rt}(\rho'_t q_t g - \tilde{P}_t)dt \\ & + \rho\tilde{q}_h(1 - e^{-\lambda_H h'})e^{-rh'}s + \rho_h\tilde{q}_h(1 - e^{-\lambda_H h'})e^{-rh'}V_I(\rho_{h'}) \\ & \quad + [1 - \rho\tilde{q}_h(1 - e^{-\lambda_H h'}) - \rho_h\tilde{q}_h(1 - e^{-\lambda_H h'})]e^{-rh'}U(\rho_{h+h'}, \rho_{h'}), \quad \text{(A.12)} \end{aligned}$$

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

[6]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.

for her to have "one-shot" deviations:

$$U^{ND}(\rho,\rho_h) \geq \tilde{U}(h') = \int_{t=0}^{h'} re^{-rt} s dt + \rho\tilde{q}_h(1-e^{-\lambda_H h'})e^{-rh'}s$$
$$+ [1-\rho\tilde{q}_h(1-e^{-\lambda_H h'})]e^{-rh'}U(\rho_h,\rho_{h'}). \quad \text{(A.13)}$$

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.

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:[7]

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

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

$$(r+\lambda_H\rho\tilde{q}_h)\tilde{U}(\rho;\rho_h) = (r+\lambda_H\rho\tilde{q}_h)s - \lambda_H\rho(1-\rho)\tilde{U}'(\rho;\rho_h) \quad \text{(A.14)}$$

where the expression for $\tilde{q}_h$ is provided by equation (1.5)

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

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

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

[7] The proof is similar to the proof of lemma A.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.

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

   Equation (A.14) 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:

$$C_h = \frac{U_S(\rho_h) - s}{(1-\rho_h)q_h(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H}}; \tag{A.15}$$

   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 (A.15) 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

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

   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 (A.12) from (A.11) yields:

$$\begin{aligned} Z(t) = \mathbb{E}\int_0^{h''} & re^{-r\tau}(\rho_\tau q_\tau g - \rho'_\tau q_\tau g)d\tau \\ & + e^{-rh''}(1-e^{-\lambda_H h''})\left\{\rho(t)\tilde{q}_h(t)[V_I(\rho_h(t+h'')) - s] + \rho_h(t)\tilde{q}_h(t)[s - V_I(\rho(t+h''))]\right\} \\ & + e^{-rh''}\left[1 - \rho(t)\tilde{q}_h(t)(1-e^{-\lambda_H h''}) - \rho_h(t)\tilde{q}_h(t)(1-e^{-\lambda_H h''})\right] Z(t+h''). \quad \text{(A.16)} \end{aligned}$$

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

$$(r + \lambda_H\rho(t)\tilde{q}_h(t) + \lambda_H\rho_h(t)\tilde{q}_h(t))Z(t) - \dot{Z}(t) = H(t) \quad \text{(A.17)}$$

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 (A.17) is

$$
\begin{aligned}
Z(t) &= (\rho(t)-\rho_h(t))\tilde{q}_h(t)g \\
&\quad - [(1-\rho_h(t))(\frac{1-\rho_h(t)}{\rho_h(t)})^{r/\lambda_H} - (1-\rho(t))(\frac{1-\rho(t)}{\rho(t)})^{r/\lambda_H}]\tilde{q}_h(t)g(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H} \\
&\qquad + Ce^{rt}(1-\rho(t))(1-\rho_h(t))\tilde{q}_h(t). \quad \text{(A.18)}
\end{aligned}
$$

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 $Z(\rho, \rho_h)$:

$$
\begin{aligned}
Z(\rho,\rho_h) &= (\rho-\rho_h)\tilde{q}_h g - [(1-\rho_h)(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (1-\rho)(\frac{1-\rho}{\rho})^{r/\lambda_H}]\tilde{q}_h g(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H} \\
&\qquad + D_h(1-\rho)(1-\rho_h)\tilde{q}_h[(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (\frac{1-\rho}{\rho})^{r/\lambda_H}]. \quad \text{(A.19)}
\end{aligned}
$$

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:

$$
D_h = -\frac{(e^{\lambda_H h}-1)g}{1-e^{-rh}}(\frac{\rho_S^\star}{1-\rho_S^\star})^{1+r/\lambda_H} + \frac{\left[1+(e^{\lambda_H h}-1)\rho_S^\star - e^{-rh}\right] g}{(1-\rho_S^\star)(1-e^{-rh})}(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H}. \quad \text{(A.20)}
$$

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

$$U^D(\rho,\rho_h) = s + (\rho-\rho_h)\tilde{q}_h g + \frac{(1-\rho)\tilde{q}_h(\frac{1-\rho}{\rho})^{r/\lambda_H}}{(1-\rho_h)q_h(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H}}(U_S(\rho_h)-s)$$
$$- [(1-\rho_h)(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (1-\rho)(\frac{1-\rho}{\rho})^{r/\lambda_H}]\tilde{q}_h g(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H}$$
$$+ D_h(1-\rho)(1-\rho_h)\tilde{q}_h[(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (\frac{1-\rho}{\rho})^{r/\lambda_H}], \quad \text{(A.21)}$$

where D_h is given by equation (A.20).

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 (A.17):

$$Z(t) = \frac{rg}{r+\lambda_H}(\rho(t)-\rho_h(t))\tilde{q}_h(t) + Ce^{rt}(1-\rho(t))(1-\rho_h(t))\tilde{q}_h(t). \quad \text{(A.22)}$$

Repeating the above procedure yields

$$Z_3(\rho,\rho_h) = \frac{rg}{r+\lambda_H}(\rho-\rho_h)\tilde{q}_h + D_{h3}(1-\rho)(1-\rho_h)\tilde{q}_h[(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (\frac{1-\rho}{\rho})^{r/\lambda_H}], \quad \text{(A.23)}$$

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 $\rho_h(t) \le \rho_I^\star$, 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:

$$Z_2(\rho,\rho_h) = \frac{rg}{r+\lambda_H}(\rho-\rho_h)\tilde{q}_h - \frac{\lambda_H g}{r+\lambda_H}\rho_h(1-\rho)\tilde{q}_h + \rho_h(1-\rho)\tilde{q}_h g[\frac{(1-\rho)\rho_I^\star}{\rho(1-\rho_I^\star)}]^{r/\lambda_H}$$
$$+ D_{h2}(1-\rho)(1-\rho_h)\tilde{q}_h(\frac{1-\rho}{\rho})^{r/\lambda_H}. \quad \text{(A.24)}$$

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

$$\begin{aligned} Z_1(\rho,\rho_h) = (\rho-\rho_h)\tilde{q}_h g - [(1-\rho_h)(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (1-\rho)(\frac{1-\rho}{\rho})^{r/\lambda_H}]\tilde{q}_h g(\frac{\rho_I^\star}{1-\rho_I^\star})^{r/\lambda_H} \\ + D_{h1}(1-\rho)(1-\rho_h)\tilde{q}_h[(\frac{1-\rho_h}{\rho_h})^{r/\lambda_H} - (\frac{1-\rho}{\rho})^{r/\lambda_H}]. \quad \text{(A.25)} \end{aligned}$$

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

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

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 (A.10) and we get the equations stated in lemma A.1. □

B.4 "Binding" Incentive Constraint

Lemma A.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 A.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 1.3

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

Lemma A.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 (A.21).

Rearranging terms yields

$$\hat{U}(\rho; h) - s = e^{-rh}[1 - \rho q(1 - e^{-\lambda_H h})](U^D(\rho, \rho_h) - s). \tag{A.26}$$

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

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

and

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

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:

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

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

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

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 A.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 A.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 A.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:

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

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 $U_S(\rho)$.[8] 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 A.3 and lemma A.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 1.4

Proof. In a niche market, $U_S(\rho_S^\star) = s$ and equation (1.21) 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 (1.26) yields

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

Then boundary conditions

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

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

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 (1.29) 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).

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 1.1, 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,

$$rg + \lambda_H(V_I(\rho) + J_I(\rho)) - \lambda_H s < (r + \lambda_H)g + \lambda_H W(\rho) - 2\lambda_H s. \tag{A.27}$$

Second, it cannot be the case that $\rho_S^\star \leq \rho_I^e$ 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

$$\rho_S^\star \times q(\rho_S^\star) = \rho_I^e = \frac{rs}{rg + \lambda_H(g - s)}. \tag{A.28}$$

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

Since $W(\cdot)$ is a strictly increasing function for $\rho > \rho_I^e$, inequality (A.27) implies that $\rho_S^\star > \rho_S^e$. $\square$

B.7 Proof of Proposition 1.5

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

$$rU_S(q) = r(gq - P_S(q)) + nq\lambda_H(s - U_S(q)) - n\lambda_H q(1-q)U_S'(q). \tag{A.29}$$

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

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

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 $q \geq q_S^\star$. 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:

$$rJ_S(q) = nrP_S(q)dt + nq\lambda_H(n(g-s) - J_S(q)) - n\lambda_H q(1-q)J_S'(q). \tag{A.30}$$

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

$$U_S(q_S^\star) = s, \quad J_S(q_S^\star) = 0 \quad \text{and} \quad J_S'(q_S^\star) = 0. \tag{A.31}$$

Boundary conditions (A.31) 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 (A.29), the equilibrium price is $P_S(q) = gq - s$. Substituting the price expression into equation (A.30) 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. $\square$

B.8 Proof of Theorem 1.1

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 A.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

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

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

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

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

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

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

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

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

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

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

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

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

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

and it is trivial to check that

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

for all $1 \leq k \leq n-2$. □

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

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

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 $\rho > \rho_I^e$ 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: $\rho_S^\star > \rho_S^e$. □

C Proofs of Results from Section 4

C.1 Proof of Proposition 1.7

Proof. Notice the derivative of

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

is $\frac{r+\lambda_H n\rho q}{\lambda_H \rho(1-\rho)}$. From observation A.1, a general solution to differential equation (1.32) 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 A.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 $\rho \leq \rho_I^e$,

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

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 $\rho_S^e \leq \rho_I^e$, 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 $q < 1$. $\square$

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

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

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 $\rho \leq \rho_S^e(q_0)$. $\square$

C.2 Proof of Theorem 1.2

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 $\rho \leq \rho_I^e$ and hence the equilibrium cutoff in the social learning phase must be larger than ρ_I^e. For $\rho > \rho_I^e$,

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

Solving the above differential equation yields

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

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_I^e. This implies: $D = C - h(\rho_I^e)ns$, where C is the constant given in the proof of proposition 1.7. From integration by parts,

$$\begin{aligned}
&\int_{\rho_I^e}^{\rho} h(x)\frac{rn(A - xq(x)B - s)}{\lambda_H x(1-x)}dx \\
&\qquad = \int_{\rho_I^e}^{\rho} h(x)\frac{rn(A - xq(x)B) + \lambda_H n^2 xq(x)s}{\lambda_H x(1-x)}dx - ns(h(\rho) - h(\rho_I^e)).
\end{aligned}$$

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. $\square$

A.2 Appendix to Chapter 2

Proof of Lemma 2.2

Proof. The worker $p \in (0,1)$ always has the choice that stays in one firm y forever. Then the value is $\frac{\mu_y(p)-rV_y}{r+\delta}$. But obviously, this is not an optimal choice (Suppose not, then all of the workers will stay in one type of firms and the market is not cleared). So we have that the equilibrium value function $W_y(p)$ must satisfy: $W_y(p) > \frac{\mu_y(p)-rV_y}{r+\delta}$. This immediately implies:

$$\Sigma_y(p){W_y}''(p) = (r+\delta)W_y(p) - (\mu^i(p) - rV^i) > 0.$$

So the equilibrium value functions W_y convex for $p \in (0,1)$. □

Proof of Lemma 2.3

Proof. Suppose workers with $p \in [0, \underline{p})$ are employed by type y firm. This implies that $W_y(p) = \frac{\mu_y(p)-rV_y}{r+\delta} + k_{y2}p^{\alpha_y}(1-p)^{1-\alpha_y}$ since 0 is included in the domain. It is easy to see that ${W_y}'(0) = \frac{\mu_{Hy}-\mu_{Ly}}{r+\delta} > 0$ and since W_y is strictly convex, $W'_y(p) > 0$ for all $p \in [0, \underline{p})$. At $\underline{p}$, worker will transfer to type $-y$ firm but smooth pasting condition implies $W'_{-y}(\underline{p}) = W'_y(\underline{p}) > 0$. Strict convexity implies $W'_{y'}(p) > 0$ so on and so forth. Therefore, we must have the equilibrium value functions W_y are strictly increasing. □

Proof of Claim 2.2

Proof. We will actually prove a more general claim, i.e., that the result holds for any combination (s_H, s_L), including $s_H < s_L$. This makes the proof also applicable to the case of $\sigma_H \neq \sigma_L$. Under strict supermodularity, for any combination of (s_H, s_L), it is impossible to have $p_1 < p_2$ and equilibrium value functions W_H (for $p \in [p_1, p_2]$), W_{L1} (for $p < p_1$), W_{L2} (for $p > p_2$) such that:

$$W_H(p_1) = W_{L1}(p_1) \quad \text{and} \quad W''_H(p_1) = W''_{L1}(p_1)$$

$$W_H(p_2) = W_{L2}(p_2) \quad \text{and} \quad W''_H(p_2) = W''_{L2}(p_2)$$

are satisfied simultaneously.

Suppose on the contrary the equations described above hold simultaneously. Then from Equation (2.3), we should get:

$$w_H(p_1) + \Sigma_H(p_1)W''_H(p_1) = w_L(p_1) + \Sigma_L(p_1)W''_{L1}(p_1)$$

and

$$w_H(p_2) + \Sigma_H(p_2)W''_H(p_2) = w_L(p_2) + \Sigma_L(p_2)W''_{L2}(p_2)$$

since

$$W_H(p_2) = W_{L2}(p_2) \quad \text{and} \quad W_H(p_1) = W_{L1}(p_1).$$

Notice that

$$W''_H(p_2) = W''_{L2}(p_2) \quad \text{and} \quad W''_H(p_1) = W''_{L1}(p_1),$$

by Lemma 2.5 and hence:

$$\frac{\Sigma_H(p_1) - \Sigma_L(p_1)}{\Sigma_H(p_1)}(r+\delta)W_H(p_1) = w_L(p_1) - \frac{\Sigma_L(p_1)}{\Sigma_H(p_1)}w_H(p_1) \tag{A.32}$$

and

$$\frac{\Sigma_H(p_2) - \Sigma_L(p_2)}{\Sigma_H(p_2)}(r+\delta)W_H(p_2) = w_L(p_2) - \frac{\Sigma_L(p_2)}{\Sigma_H(p_2)}w_H(p_2). \tag{A.33}$$

By definition,

$$\frac{\Sigma_H(p_1) - \Sigma_L(p_1)}{\Sigma_H(p_1)} = \frac{\Sigma_H(p_2) - \Sigma_L(p_2)}{\Sigma_H(p_2)} = \frac{s_H^2 - s_L^2}{s_H^2}.$$

First, if $s_H^2 = s_L^2$, Equations (A.32) and (A.33) imply that: $w_H(p_1) - w_L(p_1) = w_H(p_2) - w_L(p_2) = 0$ which cannot hold simultaneously for $p_1 \neq p_2$ since $w_H(\cdot)$ and $w_L(\cdot)$ are linear functions with different slopes Δ_H and Δ_L.

Second, if $s_H^2 > s_L^2$, then Equations (A.32) and (A.33) could be simplified as:

$$\frac{s_H^2 - s_L^2}{s_H^2}(r+\delta)(W_H(p_2) - W_H(p_1)) = w_L(p_2) - w_L(p_1) - \frac{\Sigma_L(p_2)}{\Sigma_H(p_2)}(w_H(p_2) - w_H(p_1)).$$

Under strict supermodularity, the LHS of the above equation is strictly larger than $\frac{s_H^2-s_L^2}{s_H^2}(r+\delta)W_H'(p_1)(p_2-p_1)$ by the convexity of the value function. And

$$\frac{s_H^2-s_L^2}{s_H^2}(r+\delta)W_H'(p_1)(p_2-p_1) \geq \frac{s_H^2-s_L^2}{s_H^2}\Delta_L(p_2-p_1)$$

by Lemma 2.4. Meanwhile, the RHS of the above equation is strictly smaller than

$$\Delta_L(p_2-p_1) - \frac{\Sigma_L(p_2)}{\Sigma_H(p_2)}\Delta_H(p_2-p_1)) = \frac{s_H^2-s_L^2}{s_H^2}\Delta_L(p_2-p_1)$$

which contradicts the fact that LHS is the same as RHS. The impossibility in $s_H^2 < s_L^2$ case could be proved similarly and is thus omitted. By contradiction, we immediately know the claim at the beginning of the proof is correct.

For the strict submodularity case, it suffices to relabel 'H' by 'L' and 'L' by 'H'. The claim is obviously correct given we have already proved the strict supermodularity result. □

Proof of Lemma 2.6

Proof. We will actually prove a more general Lemma, i.e., that the result holds for any combination (s_H, s_L), including $s_H < s_L$. This makes the proof also applicable to the case of $\sigma_H \neq \sigma_L$. First of all, we want to show all of the one-shot deviations are ruled out by our no-deviation condition as $dt \to 0$.

Under strict supermodularity, PAM is the only candidate equilibrium allocation by Theorem 2.1. The value functions thus are given by:

$$W_L(p) = \frac{w_L(p)}{r+\delta} + k_L p^{\alpha_L}(1-p)^{1-\alpha_L}$$

and

$$W_H(p) = \frac{w_H(p)}{r+\delta} + k_H p^{1-\alpha_H}(1-p)^{\alpha_H}.$$

Let

$$\mathcal{G}_L(p) = k_L p^{\alpha_L}(1-p)^{1-\alpha_L}(\frac{\alpha_L - p}{p(1-p)}) > 0$$

and

$$\mathcal{G}_H(p) = k_H p^{1-\alpha_H}(1-p)^{\alpha_H}(\frac{1-\alpha_H-p}{p(1-p)}) < 0$$

be the first derivatives for the non-linear parts of the value functions. Smooth pasting at $\underline{p}$ implies:

$$\frac{\Delta_L}{r+\delta} + \mathcal{G}_L(\underline{p}) = \frac{\Delta_H}{r+\delta} + \mathcal{G}_H(\underline{p}).$$

From the proof of Lemma 2.5, it suffices to show that inequality (2.11) holds for $p < \underline{p}$ and inequality (2.9) holds for $p > \underline{p}$.

For $p < \underline{p}$, define:

$$Z_L(p) = w_H(p) - w_L(p) + \frac{s_H^2 - s_L^2}{s_L^2}((r+\delta)W_L(p) - w_L(p)). \tag{A.34}$$

Obviously, we have $\lim_{p \nearrow \underline{p}} Z_L(p) = 0$ from Lemma 2.5. If we can show that $Z_L(p)$ is increasing in p as p increases from 0 to $\underline{p}$, then we are done since $Z_L(p) < Z_L(\underline{p}) = 0$. Notice that

$$Z_L'(p) = \Delta_H - \frac{s_H^2}{s_L^2}\Delta_L + \frac{s_H^2 - s_L^2}{s_L^2}(r+\delta)W_L'(p)$$

and $W_L'(p)$ lies between $\frac{\Delta_L}{r+\delta}$ and $\frac{\Delta_L}{r+\delta} + \mathcal{G}_L(\underline{p})$ for $p \in [0, \underline{p}]$.[9]

If $s_H^2 \geq s_L^2$, then

$$Z_L'(p) \geq \Delta_H - \frac{s_H^2}{s_L^2}\Delta_L + \frac{s_H^2 - s_L^2}{s_L^2}(r+\delta)\frac{\Delta_L}{r+\delta} = \Delta_H - \Delta_L > 0;$$

if $s_H^2 < s_L^2$, then

$$\begin{aligned}
Z_L'(p) &\geq \Delta_H - \frac{s_H^2}{s_L^2}\Delta_L + \frac{s_H^2 - s_L^2}{s_L^2}(r+\delta)[\frac{\Delta_L}{r+\delta} + \mathcal{G}_L(\underline{p})] \\
&= \Delta_H - \frac{s_H^2}{s_L^2}\Delta_L + \frac{s_H^2 - s_L^2}{s_L^2}(r+\delta)[\frac{\Delta_H}{r+\delta} + \mathcal{G}_H(\underline{p})] \\
&= \frac{s_H^2}{s_L^2}(\Delta_H - \Delta_L) + \frac{s_H^2 - s_L^2}{s_L^2}(r+\delta)\mathcal{G}_H(\underline{p}) > 0.
\end{aligned}$$

[9]This comes from the fact that $W_L(\cdot)$ is a strictly convex function.

Therefore, we conclude that $Z'_L(p) > 0$ for both $s_H \geq s_L$ and $s_H < s_L$ cases, which implies that $Z_L(p) < 0$ for all $p < \underline{p}$ and hence there is no profitable one-shot deviation as dt is sufficiently small.

For $p > \underline{p}$, similarly define:

$$Z_H(p) = w_L(p) - w_H(p) + [\Sigma_L(p) - \Sigma_H(p)]W''_H(p). \tag{A.35}$$

Under PAM equilibrium, we have $Z_H(\underline{p}+) = 0$ from Lemma 2.5. Notice that

$$\begin{aligned} Z_H(p) &= w_L(p) - w_H(p) + [\Sigma_L(p) - \Sigma_H(p)]W''_H(p) \\ &= w_L(p) - w_H(p) + \frac{s_L^2 - s_H^2}{s_H^2}((r+\delta)W_H(p) - w_H(p)), \end{aligned}$$

with $W'_H(p)$ lies between $\frac{\Delta_H}{r+\delta} + \mathcal{G}_H(\underline{p})$ and $\frac{\Delta_H}{r+\delta}$ for $p \in [\underline{p}, 1]$. Similar to the proof for $p < \underline{p}$ case, if $s_L^2 > s_H^2$

$$Z'_H(p) \leq \Delta_L - \Delta_H < 0;$$

and if $s_L^2 \leq s_H^2$

$$Z'_H(p) \leq \Delta_L - \frac{s_L^2}{s_H^2}\Delta_H + \frac{s_L^2 - s_H^2}{s_H^2}(r+\delta)(\frac{\Delta_L}{r+\delta} + \mathcal{G}_L(\underline{p})) < 0.$$

Therefore, $Z'_H(p) < 0$ for both $s_H \geq s_L$ and $s_H < s_L$ cases and hence $Z_H(p) < 0$ for all $p > \underline{p}$.

Second, since there is no one-shot deviation for any p, obviously there will be no any other deviation for any p. Consider any deviation starting at p. Then the above result says it is better not to deviate for at least dt time. Suppose after dt, we achieve a new p'. Similarly, there should be no profitable deviation for at least dt' time. Keep using the same logic and we can conclude that any deviation is not profitable. □

Derivation of the Boundary Conditions

Here, we just investigate the boundary conditions for the first case: $\underline{p} < p_0$. The derivation is similar for the second case.

In a stationary equilibrium, both the total measure $\int_0^1 f_y(p,t)dp$ and the expectations $\int_0^1 pf_y(p,t)dp$ are constant over time. Hence, it must be the case that $\int_0^1 \frac{\partial f_y(p,t)}{\partial t}dp = 0$ and $\int_0^1 p\frac{\partial f_y(p,t)}{\partial t}dp = 0$

From

$$\frac{\partial f_y(p,t)}{\partial t} = \frac{d^2}{dp^2}[\Sigma_y(p)f_y(p,t)] - \delta f_y(p,t),$$

we should have:

$$\int_0^{\underline{p}} \{\frac{d^2}{dp^2}[\Sigma_L(p)f_L(p)] - \delta f_L(p)\}dp = 0$$

and

$$\int_{\underline{p}}^{p_0} \{\frac{d^2}{dp^2}[\Sigma_H(p)f_H(p)] - \delta f_H(p)\}dp + \int_{p_0}^1 \{\frac{d^2}{dp^2}[\Sigma_H(p)f_H(p)] - \delta f_H(p)\}dp = 0.$$

The above two equations give us:

$$\frac{d}{dp}[\Sigma_L(p)f_L(p)]|_{\underline{p}-} = \delta(1-\pi)$$

and

$$\Sigma_H(p_0)[f_H'(p_0-) - f_H'(p_0+)] = \frac{d}{dp}[\Sigma_H(p)f_H(p)]|_{\underline{p}+} + \delta\pi$$

since the market clearing conditions imply:

$$\int_0^{\underline{p}} f_L(p)dp = 1-\pi$$

$$\int_{\underline{p}}^1 f_H(p)dp = \pi$$

and there is continuity at p_0:

$$f_H(p_0-) = f_H(p_0+).$$

Meanwhile, notice that inflow at p_0 must be the same as δ, which implies that $\Sigma_H(p_0)[f_H'(p_0-) - f_H'(p_0+)] = \delta$. This immediately gives us the flow equation at $\underline{p}$:

$$\frac{d}{dp}[\Sigma_L(p)f_L(p)]|_{\underline{p}-} = \frac{d}{dp}[\Sigma_H(p)f_H(p)]|_{\underline{p}+}.$$

Now apply similar logic and we can get:

$$\int_0^{\underline{p}} \left\{ p\frac{d^2}{dp^2}[\Sigma_L(p)f_L(p)] - p\delta f_L(p) \right\} dp + \int_{\underline{p}}^1 \left\{ p\frac{d^2}{dp^2}[\Sigma_H(p)f_H(p)] - p\delta f_H(p) \right\} dp = 0.$$

Notice that

$$\int_0^{\underline{p}} p\delta f_L(p)dp + \int_{\underline{p}}^1 p\delta f_H(p)dp = \delta p_0$$

by the martingale property. Meanwhile, we still have: $\Sigma_H(p_0)[f'_H(p_0-) - f'_H(p_0+)] = \delta$. Hence,after some tedious algebra, we can get:

$$\left\{ p\frac{d}{dp}[\Sigma_L(p)f_L(p)] + \Sigma_L(p)f_L(p) \right\} |_{\underline{p}-} = \left\{ p\frac{d}{dp}[\Sigma_H(p)f_H(p)] + \Sigma_H(p)f_H(p) \right\} |_{\underline{p}+}$$

which gives us the boundary condition at $\underline{p}$:

$$\Sigma_H(\underline{p}+)f_H(\underline{p}+) = \Sigma_L(\underline{p}-)f_L(\underline{p}-).$$

Proof of Proposition 2.1

Proof. First, we can express $f_{H0}, f_{H1}, f_{H2}, f_{L0}$ as functions of $\underline{p}$. Equations (2.25) and (2.27) imply:

$$f_{L0} = \frac{1-\pi}{\int_0^{\underline{p}} p^{\gamma_{L1}}(1-p)^{\gamma_{L2}}dp}.$$

and

$$f_{H2} = f_{H0}(\frac{p_0}{1-p_0})^{\gamma_{H1}-\gamma_{H2}} + f_{H1}$$

From Equations (2.23) and (2.26), f_{H0} and f_{H1} as could be written as:

$$f_{H0} = \frac{\eta_H + \eta_L}{2\eta_H}\frac{s_L^2}{s_H^2}(\frac{\underline{p}}{1-\underline{p}})^{\eta_L - \eta_H} f_{L0}$$

and

$$f_{H1} = -\frac{\eta_L - \eta_H}{2\eta_H}\frac{s_L^2}{s_H^2}(\frac{\underline{p}}{1-\underline{p}})^{\eta_L + \eta_H} f_{L0}.$$

Here,

$$\eta_L = \sqrt{\frac{1}{4} + \frac{2\delta}{s_L^2}} > \eta_H = \sqrt{\frac{1}{4} + \frac{2\delta}{s_H^2}} > 1/2.$$

Next, we want to show that both f_{H0} and f_{H1} are decreasing in $\underline{p}$.

Rewrite f_{H0} as:

$$f_{H0} = \frac{\eta_H + \eta_L}{2\eta_H}\frac{s_L^2}{s_H^2}(\frac{\underline{p}}{1-\underline{p}})^{\eta_L - \eta_H}\frac{1-\pi}{\int_0^{\underline{p}} p^{\gamma_{L1}}(1-p)^{\gamma_{L2}}dp}.$$

and it suffices to show that $(\frac{\underline{p}}{1-\underline{p}})^{\eta_L-\eta_H}\frac{1-\pi}{\int_0^{\underline{p}} p^{\gamma_{L1}}(1-p)^{\gamma_{L2}}dp}$ is decreasing in $\underline{p}$. Notice that

$$(\frac{\underline{p}}{1-\underline{p}})^{\eta_L-\eta_H} = \int_0^{\underline{p}}[(\frac{p}{1-p})^{\eta_L-\eta_H}]'dp = \int_0^{\underline{p}}(\eta_L - \eta_H)(\frac{p}{1-p})^{\eta_L-\eta_H-1}(\frac{1}{1-p})^2dp.$$

Let $G_1(p) = p^{\gamma_{L1}}(1-p)^{\gamma_{L2}}$ and $G_2(p) = (\frac{p}{1-p})^{\eta_L-\eta_H-1}(\frac{1}{1-p})^2$ such that:

$$\frac{G_1(p)}{G_2(p)} = p^{-\frac{1}{2}+\eta_H}(1-p)^{-\frac{1}{2}-\eta_H}$$

is increasing in p. Therefore, we could derive:

$$(\frac{\underline{p}}{1-\underline{p}})^{\eta_L-\eta_H}\frac{1-\pi}{\int_0^{\underline{p}} p^{\gamma_{L1}}(1-p)^{\gamma_{L2}}dp}$$

is decreasing in $\underline{p}$[10] and hence f_{H0} is decreasing in $\underline{p}$ as well.

Similarly, we can rewrite f_{H1} as:

$$f_{H1} = -\frac{\eta_L - \eta_H}{2\eta_H}\frac{s_L^2}{s_H^2}(\frac{\underline{p}}{1-\underline{p}})^{\eta_L + \eta_H}\frac{1-\pi}{\int_0^{\underline{p}} p^{\gamma_{L1}}(1-p)^{\gamma_{L2}}dp}.$$

Similarly,

$$(\frac{\underline{p}}{1-\underline{p}})^{\eta_L+\eta_H} = \int_0^{\underline{p}}(\eta_L + \eta_H)(\frac{p}{1-p})^{\eta_L+\eta_H-1}(\frac{1}{1-p})^2dp.$$

Let $G_3(p) = (\frac{p}{1-p})^{\eta_L+\eta_H-1}(\frac{1}{1-p})^2$ and we have:

$$\frac{G_1(p)}{G_3(p)} = p^{-\frac{1}{2}-\eta_H}(1-p)^{-\frac{1}{2}+\eta_H}$$

[10] Actually, we are using the result that if $\frac{G_2(p)}{G_1(p)}$ is decreasing in p, then $\frac{\int_0^{\underline{p}} G_2(p)dp}{\int_0^{\underline{p}} G_1(p)dp}$ will also be decreasing in $\underline{p}$. This is true because by the definition of Riemann integral, $\int_0^{\underline{p}} G_1(p)dp$ and $\int_0^{\underline{p}} G_2(p)dp$ could be written as the limit of Riemann sum. The ratio of two Riemann sums is always decreasing in $\underline{p}$ since $\frac{G_2(p)}{G_1(p)}$ is decreasing in p.

is decreasing in p. Therefore, it must be the case that

$$-(\frac{\underline{p}}{1-\underline{p}})^{\eta_L+\eta_H}\frac{1-\pi}{\int_0^{\underline{p}} p^{\gamma_{L1}}(1-p)^{\gamma_{L2}}dp}$$

is decreasing in $\underline{p}$ and hence f_{H1} is also decreasing in $\underline{p}$.

Finally, it is immediate that

$$f_{H2} = f_{H0}(\frac{p_0}{1-p_0})^{\gamma_{H1}-\gamma_{H2}} + f_{H1}$$

is also decreasing in $\underline{p}$. Therefore, we can expressing f_{H0}, f_{H1} and f_{H2} as $\xi_0(\underline{p})$, $\xi_1(\underline{p})$ and $\xi_2(\underline{p})$ respectively such that $\xi_0{}' < 0$, $\xi_1{}' < 0$ and $\xi_2{}' < 0$.

Hence, the market clearing condition (2.24) implies:

$$H(\underline{p}) = \int_{\underline{p}}^{p_0}[\xi_0(\underline{p})p^{\gamma_{H1}}(1-p)^{\gamma_{H2}} + \xi_1(\underline{p})p^{\gamma_{H2}}(1-p)^{\gamma_{H1}}]dp + \int_{p_0}^{1}\xi_2(\underline{p})p^{\gamma_{H2}}(1-p)^{\gamma_{H1}}dp = \pi.$$

It is easy to check that $H' < 0$ since $\xi_0{}' < 0$, $\xi_1{}' < 0$ and $\xi_2{}' < 0$. There exists $\underline{p} \in (0, p_0)$ such that $H(\underline{p}) = \pi$ if and only if $\lim_{p\to 0} H(p) > \pi$ and $\lim_{p\to p_0} H(p) < \pi$.

As $\underline{p} \to 0$, $f_{H0} = \xi_0(\underline{p}) \to \infty$ and $f_{H1} = \xi_1(\underline{p}) \to 0$, which imply:

$$\lim_{p\to 0} H(p) \to \infty > \pi.$$

Meanwhile, when $\underline{p} \to p_0$, it is obvious that $H(\underline{p}) \to \int_{p_0}^1 f_{H2}p^{\gamma_{H2}}(1-p)^{\gamma_{H1}}dp$. Notice that

$$f_{H2} = f_{H0}(\frac{p_0}{1-p_0})^{\gamma_{H1}-\gamma_{H2}} + f_{H1} \to \frac{s_L^2}{s_H^2}(\frac{p_0}{1-p_0})^{\eta_L+\eta_H}\frac{1-\pi}{\int_0^{p_0} p^{\gamma_{L1}}(1-p)^{\gamma_{L2}}dp}$$

as $\underline{p} \to p_0$.

As a result, $\lim_{p\to p_0} H(p) < \pi$ if and only if:

$$\frac{s_L^2}{s_H^2}(\frac{p_0}{1-p_0})^{\eta_L+\eta_H}\frac{1-\pi}{\int_0^{p_0} p^{\gamma_{L1}}(1-p)^{\gamma_{L2}}dp}\int_{p_0}^1 p^{\gamma_{H2}}(1-p)^{\gamma_{H1}}dp < \pi,$$

which establishes Equation 2.28 in the proposition. Moreover, since $H(\cdot)$ is strictly decreasing, the solution to $H(p) = \pi$ must be at most one. This completes our proof of Proposition 2.1. $\square$

Proof of Corollary 2.1

Proof. To make the proof, we have to redefine the $H(\cdot)$ function in the proof of Proposition 2.1 as $H(p;\pi,p_0)$ with equilibrium cutoff $\underline{p}$ satisfying $H(\underline{p};\pi,p_0)=\pi$. It is obviously to verify that H is linear in $(1-\pi)$. So as π increases, $\pi/(1-\pi)$ increases and we have to decrease $\underline{p}$ to balance the equation. On the other hand,

$$
\begin{aligned}
\frac{\partial H}{\partial p_0} &= \xi_0(\underline{p})p_0^{\gamma_1^H}(1-p_0)^{\gamma_2^H}+\xi_1(\underline{p})p_0^{\gamma_2^H}(1-p_0)^{\gamma_1^H}-\xi_2(\underline{p})p_0^{\gamma_2^H}(1-p_0)^{\gamma_1^H} \\
&+ \int_{p_0}^1 \frac{\partial \xi_2(\underline{p})}{\partial p_0} p^{\gamma_2^H}(1-p)^{\gamma_1^H}\,dp.
\end{aligned}
$$

It is easy to verify that the first line on the RHS is zero while the second line is strictly positive. Hence $H(\underline{p};\pi,p_0)$ is increasing in p_0 and we have to increase $\underline{p}$ to keep the equation as p_0 increases.

The proof for the comparative statics for $\underline{p}>p_0$ case is similar and hence is omitted. □

Proof of Proposition 2.2

Proof. First, from equation (2.35), we have:

$$
f_{H0}=\frac{\pi}{\int_{\underline{p}}^1 p^{\gamma_{H2}}(1-p)^{\gamma_{H1}}dp}.
$$

Second, Equations (2.34) and (2.37) imply:

$$
f_{L1}=\frac{\eta_L-\eta_H}{2\eta_L}\frac{s_H^2}{s_L^2}\left(\frac{\underline{p}}{1-\underline{p}}\right)^{-\eta_L-\eta_H}f_{H0}
$$

and

$$
f_{L2}=\frac{\eta_L+\eta_H}{2\eta_L}\frac{s_H^2}{s_L^2}\left(\frac{\underline{p}}{1-\underline{p}}\right)^{\eta_L-\eta_H}f_{H0}.
$$

Here,

$$
\eta_L=\sqrt{\frac{1}{4}+\frac{2\delta}{s_L^2}}>\eta_H=\sqrt{\frac{1}{4}+\frac{2\delta}{s_H^2}}>1/2.
$$

It is easy to verify that f_{H0}, f_{L1}, f_{L2} are increasing in $\underline{p}$ and hence $f_{L0} = f_{L1} + f_{L2}(\frac{p_0}{1-p_0})^{-2\eta_L}$ is also increasing in $\underline{p}$ by Equation (2.38).

Hence, we can express f_{L0}, f_{L1}, f_{L2} as $\xi_0(\underline{p})$, $\xi_1(\underline{p})$ and $\xi_2(\underline{p})$ respectively such that $\xi_0{}' > 0$, $\xi_1{}' > 0$ and $\xi_2{}' > 0$.

Finally, the market clearing condition (2.36) implies:

$$H(\underline{p}) = \int_0^{p_0} \xi_0(\underline{p}) p^{\gamma_{L1}} (1-p)^{\gamma_{L2}} dp + \int_{p_0}^{\overline{p}} [\xi_1(\underline{p}) p^{\gamma_{L1}} (1-p)^{\gamma_{L2}} + \xi_2(\underline{p}) p^{\gamma_{L2}} (1-p)^{\gamma_{L1}}] dp = 1 - \pi.$$

Obviously, $H(\cdot)$ is strictly increasing, which guarantees the solution is unique if it exists and $\lim_{p \to p_0} H(p) \leq 1 - \pi$ will give us Equation (2.39) in Proposition 2.2. □

Proof of Lemma 2.8

Proof. By substituting $\mu_H(p)$ and $\mu_L(p)$, the total expected surplus for allocation 1 could be written as:

$$S_1 = \int_{\Omega_H} (\Delta_H p + \mu_{LH}) f_H(p) dp + \int_{\Omega_L} (\Delta_L p + \mu_{LL}) f_L(p) dp.$$

From market clearing and martingale property conditions, we can furthermore rewrite S_1 as:

$$S_1 = (\Delta_H - \Delta_L) \int_{\Omega_H} p f_H(p) dp + \Delta_L p_0 + \pi \mu_{LH} + (1-\pi) \mu_{LL}.$$

And similarly,

$$S_2 = (\Delta_H - \Delta_L) \int_{\tilde{\Omega}_H} p f_H(p) dp + \Delta_L p_0 + \pi \mu_{LH} + (1-\pi) \mu_{LL}.$$

Therefore, $S_1 > S_2$ if and only if

$$\int_{\Omega_H} p f_H(p) dp > \int_{\tilde{\Omega}_H} p \tilde{f}_H(p) dp$$

or alternatively, $\int_{\Omega_L} p f_H(p) dp < \int_{\tilde{\Omega}_L} p \tilde{f}_L(p) dp$. □

Proof of Theorem 2.4

Proof. We establish the proof of Theorem 2.4 under supermodularity. The same logic goes through for submodularity. The proof is constructed in the following three steps: 1. for $N = 3$ we show that the planner can increase output when changing the cutoffs; 2. for $N = 3$ no allocation dominates PAM; 3. For any N, the allocation with $N - 2$ cutoffs dominates that with N cutoffs.

1. For $N = 3$, output increases from changing the cutoffs

Consider any allocation with three cutoffs $0 < \underline{p}_3 < \underline{p}_2 < \underline{p}_1 < 1$ such that workers with $p \in (\underline{p}_1, 1]$ and $p \in (\underline{p}_3, \underline{p}_2)$ are allocated to the high type firms while workers with $p \in [0, \underline{p}_3)$ and $p \in (\underline{p}_2, \underline{p}_1)$ are allocated to the low type firms. Furthermore, denote the ergodic density function for this allocation to be f_y and for p close to 0, let the density function be $f_L(p) = \tilde{f}_{L0} p^{\gamma_L}(1-p)^{1-\gamma_L}$ while the ergodic density function for p close to 1 is denoted by $f_H(p) = \tilde{f}_{H0} p^{1-\gamma_H}(1-p)^{\gamma_H}$ where $\tilde{f}_{L0}$ and $\tilde{f}_{H0}$ are constants. Correspondingly, denote the ergodic density under the PAM allocation to be f_y^* with the unique cutoff $\underline{p}$.

1. Suppose the planner changes the allocation by moving the interval to the left: $(\underline{p}_2, \underline{p}_1) \to (\underline{p}'_2, \underline{p}'_1)$ where $(\underline{p}'_2, \underline{p}'_1) = (\underline{p}_2 - \epsilon_2, \underline{p}_1 - \epsilon_1)$. Choose ϵ_1, ϵ_2 such that market clearing is satisfied:
$$\int_{\underline{p}'_1}^{\underline{p}_1} f_H(p)dp = \int_{\underline{p}'_2}^{\underline{p}_2} f_H(p)dp.$$
2. Given the new cutoffs, the Kolmogorov forward equation will pin down a new density $\hat{f}_L$ in the interval $(\underline{p}'_2, \underline{p}'_1)$. Globally, we need to satisfy market clearing and the martingale property conditions. The market clearing condition for the H types is satisfied by the construction. For the L type firms it requires that:
$$\int_{\underline{p}'_2}^{\underline{p}'_1} \hat{f}_L(p)dp = \int_{\underline{p}_2}^{\underline{p}_1} f_L(p)dp.$$

The martingale property condition requires that $\mathbb{E}_{\Omega'_H} p + \mathbb{E}_{\Omega'_L} p = p_0$ or:

$$\int_0^{p_3} p f_L(p)dp + \int_{p_3}^{p'_2} p f_H(p)dp + \int_{p'_2}^{p'_1} p \hat{f}_L(p)dp + \int_{p'_1}^{1} p f_H(p)dp = p_0.$$

Above are a system of two linear equations about the distributional parameters for $\hat{f}_L$ and $\hat{f}_L$ could be solved as a result.[11]

3. Then comparing the original allocation to the new one, we get

$$\mathbb{E}_{\Omega'_H} p - \mathbb{E}_{\Omega_H} p = \int_{\underline{p}'_1}^{\underline{p}_1} p f_H(p)dp - \int_{\underline{p}'_2}^{\underline{p}_2} p f_H(p)dp > 0$$

since by construction

$$\int_{\underline{p}'_1}^{\underline{p}_1} f_H(p)dp = \int_{\underline{p}'_2}^{\underline{p}_2} f_H(p)dp$$

and the interval $[\underline{p}'_2, \underline{p}'_1]$ is strictly to the left of $[\underline{p}_2, \underline{p}_1]$. From Lemma 2.8, $\mathbb{E}_{\Omega'_H} p > \mathbb{E}_{\Omega_H} p$ implies the planner prefers allocation Ω' over Ω.

4. Similarly, we can consider another transform which is to move the interval to the right: $(\underline{p}_3, \underline{p}_2) \to (\underline{p}'_3, \underline{p}'_2)$ where $(\underline{p}'_3, \underline{p}'_2) = (\underline{p}_3 + \epsilon_2, \underline{p}_2 + \epsilon_1)$. This can also lead to output increases. Keep on doing such transformations and eventually, we can have both the distance and the measure between $\underline{p}'_3$ and $\underline{p}'_1$ arbitrarily small while the new $(\underline{p}'_1, \underline{p}'_2, \underline{p}'_3)$ allocation strictly dominates the original $(\underline{p}_1, \underline{p}_2, \underline{p}_3)$ allocation.

2. For $N = 3$, no allocation dominates PAM

1. We now show by contradiction that no allocation dominates PAM for $N = 3$. Suppose on the contrary that there exists an allocation with cutoffs $\tilde{p}_1$, $\tilde{p}_2$ and $\tilde{p}_3$ which dominates the PAM allocation. Then by Lemma 2.8, we should have:

$$\int_{\tilde{p}_1}^{1} p f_H(p)dp + \int_{\tilde{p}_3}^{\tilde{p}_2} p f_H(p)dp > \int_{\underline{p}}^{1} p f_H^*(p)dp \tag{A.36}$$

[11] Things are slightly different if we have $p_0 \in (p'_2, p'_1)$. Then we have four new distribution coefficients but we also have two more equations: $\hat{f}_L(p_0-) = \hat{f}_L(p_0+)$ and $\Sigma_L(p_0)(\hat{f}'_L(p_0-) - \hat{f}'_L(p_0+)) = \delta$. We can use this system of four linear equations to pin down the four parameters.

and

$$\int_{\tilde{p}_2}^{\tilde{p}_1} pf_L(p)dp + \int_0^{\tilde{p}_3} pf_L(p)dp < \int_0^{\underline{p}} pf_L^*(p)dp. \tag{A.37}$$

From Step 1, we can first fix $\tilde{p}_3$ and make $\tilde{p}_2'$ move towards $\tilde{p}_3$, which is efficiency improving. $\tilde{p}_1$ could be extended to the left until it reaches $\hat{p}_1$: $\int_{\hat{p}_1}^1 f_H(p)dp = \pi$. Since $\int_{\tilde{p}_1'}^1 f_H(p)dp < \pi$, it must be the case that $\hat{p}_1 < \tilde{p}_1'$. If $\tilde{p}_2'$ is sufficiently close to $\tilde{p}_3$, we will have $\tilde{p}_2' < \hat{p}_1$. By hypothesis:

$$\int_{\hat{p}_1}^1 pf_H(p)dp > \int_{\tilde{p}_1'}^1 pf_H(p)dp + \int_{\tilde{p}_3}^{\tilde{p}_2'} pf_H(p)dp > \int_{\underline{p}}^1 pf_H^*(p)dp.$$

On the other hand, it is also efficiency improving by fixing $\tilde{p}_1$ and making $\tilde{p}_2'$ move towards $\tilde{p}_1$. Similarly define $\hat{p}_3$ as: $\int_0^{\hat{p}_3} f_L(p)dp = (1-\pi)$ such that $\hat{p}_3 > \tilde{p}_3'$. By hypothesis,

$$\int_0^{\hat{p}_3} pf_L(p)dp < \int_0^{\underline{p}} pf_L^*(p)dp.$$

since we can make $\tilde{p}_2'$ sufficiently close to $\tilde{p}_1$.

2. The next step of the proof requires Lemma A.7 below. The Lemma implies that we should have $\tilde{p}_3' < \hat{p}_3 < \underline{p} < \hat{p}_1 < \tilde{p}_1'$ to guarantee that

$$\int_{\hat{p}_1}^1 pf_H(p)dp > \int_{\underline{p}}^1 pf_H^*(p)dp \text{ and } \int_0^{\hat{p}_3} pf_L(p)dp < \int_0^{\underline{p}} pf_L^*(p)dp.$$

Therefore, inequalities (A.36) and (A.37) only hold when $\tilde{p}_1' - \tilde{p}_3' > \hat{p}_1 - \hat{p}_3 > 0$ which contradicts that fact that we can make the distance between $\tilde{p}_1'$ and $\tilde{p}_3'$ arbitrarily small while still keeping the inequalities (A.36) and (A.37). Hence, no allocation with $N = 3$ cutoffs could be better than the PAM allocation in terms of aggregate surplus.

3. For N cutoffs, the allocation is dominated by any allocation with $N-2$ cutoffs.

Consider three adjacent cutoffs $\underline{p}_{n-1}, > \underline{p}_n > \underline{p}_{n+1}$ such that workers with $p \in (\underline{p}_{n-1}, \underline{p}_{n-2})$ and $p \in (\underline{p}_{n+1}, \underline{p}_n)$ are allocated to high type firms; workers with $p \in (\underline{p}_n, \underline{p}_{n-1})$ and $p \in (\underline{p}_{n+2}, \underline{p}_{n+1})$ are allocated to low type firms. Suppose the density functions are such that the market clears and the expectation of p's is p_0. Then we just need to choose κ such that

$$\int_{\underline{p}_{n-1}-\kappa}^{\underline{p}_{n-1}} f_H(p)dp = \int_{\underline{p}_{n+1}}^{\underline{p}_n} f_H(p)dp.$$

Now $\underline{p}_{n-1}$, $\underline{p}_n$ and $\underline{p}_{n+1}$ converge to $\underline{p}_{n-1} - \kappa$ but p_{n+2} is kept to be the same. The market clearing condition requires that

$$\int_{\underline{p}_{n+2}}^{\underline{p}_{n-1}-\kappa} \tilde{f}_L(p)dp = \int_{\underline{p}_n}^{\underline{p}_{n-1}} f_L(p)dp + \int_{\underline{p}_{n+2}}^{\underline{p}_{n+1}} f_L(p)dp.$$

Meanwhile, the martingale property condition requires that:

$$\int_{\underline{p}_1}^{1} pf_H(p)dp + \cdots + \int_{\underline{p}_{n-1}-\kappa}^{\underline{p}_{n-2}} pf_H(p)dp + \int_{\underline{p}_{n+2}}^{\underline{p}_{n-1}-\kappa} p\tilde{f}_L(p)dp + \cdots + \int_{0}^{\underline{p}_N} pf_L(p)dp = p_0.$$

Similar to Step 1, we have a system of two linear equations about two distributional coefficients and density $\tilde{f}_L$ could be solved. As before,

$$\mathbb{E}_{\Omega_H} p = \int_{\Omega_H} pf_H(p)dp$$

must become higher and this allocation with $N-2$ cutoffs will generate a higher aggregate payoff.

Finally, by the standard induction argument, we can conclude that the PAM allocation with one cutoff dominates any allocation with $N \geq 3$ cutoffs in aggregate surplus. $\square$

Lemma A.7

Lemma A.7. *Let $\hat{p}_1$ be such that $\int_{\hat{p}_1}^1 f_H(p)dp = \pi$, where $f_H(p)$ satisfies the Kolmogorov forward equation, then $\int_{\hat{p}_1}^1 pf_H(p)dp$ is increasing in $\hat{p}_1$. Let $\hat{p}_3$ be such that $\int_0^{\hat{p}_3} f_L(p)dp = (1-\pi)$, where $f_L(p)$ satisfies the Kolmogorov forward equation, then $\int_0^{\hat{p}_3} pf_L(p)dp$ is also increasing in $\hat{p}_3$.*

Proof. We just prove the case that $\hat{p}_1 > p_0$. The other cases are similar. Let $f_H(p) = C_H(1-p)^{\gamma_{H1}} p^{\gamma_{H2}}$ where

$$\gamma_{H1} = -\frac{3}{2} + \eta_H \quad \text{and} \quad \gamma_{H2} = -\frac{3}{2} - \eta_H.$$

From Kolmogorov forward equation,

$$\int_{\hat{p}_1}^{1} f_H(p)dp = \frac{1}{\delta}\int_{\hat{p}_1}^{1} \frac{d^2}{dp^2}[\Sigma_H(p)f_H(p)] = \pi$$

or

$$\frac{\eta_H + \hat{p}_1 - \frac{1}{2}}{\hat{p}_1(1-\hat{p}_1)}\Sigma_H(\hat{p}_1)f_H(\hat{p}_1) = \delta\pi.$$

Notice that

$$\int_{\hat{p}_1}^{1} pf_H(p)dp = \frac{1}{\delta}\int_{\hat{p}_1}^{1} p\frac{d^2}{dp^2}[\Sigma_H(p)f_H(p)]dp$$

and could be simplified as:

$$\pi\hat{p}_1 + \frac{\pi\hat{p}_1(1-\hat{p}_1)}{\eta_H + \hat{p}_1 - \frac{1}{2}} = \frac{\pi\hat{p}_1(\eta_H + \frac{1}{2})}{\eta_H + \hat{p}_1 - \frac{1}{2}}$$

which is increasing in $\hat{p}_1$ since

$$\eta_H = \sqrt{\frac{1}{4} + \frac{2\delta}{s_y^2}} > \frac{1}{2}.$$

□

On the Job Human Capital Accumulation

Under the assumption of $\underline{p}^u = \underline{p}^e = \underline{p}$, the value functions could be written as:

$$\begin{aligned}
W_y^u(p) &= \frac{\mu_y(p) - rV_y}{r+\delta+\lambda} + k_{y1}^u p^{1-\alpha_y^u}(1-p)^{\alpha_y^u} + k_{y2}^u p^{\alpha_y^u}(1-p)^{1-\alpha_y^u} \\
&- \frac{\lambda\frac{(s_y^u)^2}{(s_y^e)^2}}{(r+\delta+\lambda)[(\lambda+\delta+r) - \frac{(s_y^u)^2}{(s_y^e)^2}(r+\delta)]}[\mu_y(p) + \xi(p) - rV_y] \\
&+ \frac{\lambda}{(\lambda+\delta+r) - \frac{(s_y^u)^2}{(s_y^e)^2}(r+\delta)}W_y^e(p)
\end{aligned}$$

and

$$W_y^e(p) = \frac{\mu_y(p) + \xi(p) - rV_y}{r+\delta} + k_{y1}^e p^{1-\alpha_y^e}(1-p)^{\alpha_y^e} + k_{y2}^e p^{\alpha_y^e}(1-p)^{1-\alpha_y^e}$$

where

$$\alpha_y^u = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{2(r+\delta+\lambda)}{(s_y^u)^2}} \geq 1$$
$$\alpha_y^e = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{2(r+\delta)}{(s_y^e)^2}} \geq 1.$$

Boundary conditions

$$W_L^e(\underline{p}) = W_H^e(\underline{p}), \quad W_L^{e\prime}(\underline{p}) = W_H^{e\prime}(\underline{p}), \quad W_L^{e\prime\prime}(\underline{p}) = W_H^{e\prime\prime}(\underline{p})$$

would imply (by normalizing $V_L = 0$ as usual):

$$r\tilde{V}_H^e = (\mu_{LH} - \mu_{LL}) + \frac{\alpha_H^e(\alpha_L^e - 1)(\Delta_H - \Delta_L)\underline{p}}{\alpha_H^e(\alpha_L^e - 1) - (1-\underline{p})(\alpha_L^e - \alpha_H^e)}.$$

And from

$$W_L^u(\underline{p}) = W_H^u(\underline{p}), \quad W_L^{u\prime}(\underline{p}) = W_H^{u\prime}(\underline{p}), \quad W_L^{u\prime\prime}(\underline{p}) = W_H^{u\prime\prime}(\underline{p}),$$

another equilibrium payoff $\tilde{V}_H^u$ could be derived as:

$$\begin{aligned} r\tilde{V}_H^u &= (\mu_{LH} - \frac{A_L}{B_L}\frac{B_H}{A_H}\mu_{LL}) - \frac{B_H}{A_H}\frac{\lambda\xi_L}{r+\delta+\lambda}(\frac{1-A_H}{B_H} - \frac{1-A_L}{B_L}) \\ &+ \frac{B_H}{A_H}\frac{\alpha_H^u(\alpha_L^u - 1)(D_H - D_L)\underline{p}}{\alpha_H^u(\alpha_L^u - 1) - (1-\underline{p})(\alpha_L^u - \alpha_H^u)}, \end{aligned}$$

where

$$D_H = \frac{A_H}{B_H}\Delta_H - \frac{1-A_H}{B_H}\frac{\lambda\Delta_\xi}{r+\delta+\lambda}$$
$$D_L = \frac{A_L}{B_L}\Delta_L - \frac{1-A_L}{B_L}\frac{\lambda\Delta_\xi}{r+\delta+\lambda}$$
$$A_H = 1 - \frac{(s_H^u)^2}{(s_H^e)^2} \quad B_H = (\lambda+\delta+r) - \frac{(s_H^u)^2}{(s_H^e)^2}(r+\delta)$$
$$A_L = 1 - \frac{(s_L^u)^2}{(s_L^e)^2} \quad B_L = (\lambda+\delta+r) - \frac{(s_L^u)^2}{(s_L^e)^2}(r+\delta).$$

Proof of Proposition 2.3

Proof. Supermodularity is equivalent to $\Delta_H > \Delta_L$, and $\xi_H \simeq \xi_L$ is equivalent to $\Delta_\xi = \xi_H - \xi_L \to 0$. The proof can be divided into three parts. As a sufficient condition,

1.
$$(\mu_{LH} - \frac{A_L}{B_L}\frac{B_H}{A_H}\mu_{LL}) - \frac{B_H}{A_H}\frac{\lambda\xi_L}{r+\delta+\lambda}(\frac{1-A_H}{B_H} - \frac{1-A_L}{B_L}) < (\mu_{LH} - \mu_{LL})$$

2.
$$\frac{B_H}{A_H}(D_H - D_L) < \Delta_H - \Delta_L$$

and

3.
$$\frac{\alpha_H^u(\alpha_L^u - 1)\underline{p}}{\alpha_H^u(\alpha_L^u - 1) - (1-\underline{p})(\alpha_L^u - \alpha_H^u)} < \frac{\alpha_H^e(\alpha_L^e - 1)\underline{p}}{\alpha_H^e(\alpha_L^e - 1) - (1-\underline{p})(\alpha_L^e - \alpha_H^e)}$$

should be satisfied simultaneously.

First of all, notice that $\frac{(s_H^u)^2}{(s_H^e)^2} > \frac{(s_L^u)^2}{(s_L^e)^2}$ since $\Delta_H > \Delta_L$. As a result, $\frac{A_H}{B_H} < \frac{A_L}{B_L}$ and $\frac{1-A_H}{B_H} > \frac{1-A_L}{B_L}$. The first inequality holds since $\mu_{LH} - \frac{A_L}{B_L}\frac{B_H}{A_H}\mu_{LL} < \mu_{LH} - \mu_{LL}$ and $\frac{A_L}{B_L}\frac{B_H}{A_H}\mu_{LL}) - \frac{B_H}{A_H}\frac{\lambda\xi_L}{r+\delta+\lambda}(\frac{1-A_H}{B_H} - \frac{1-A_L}{B_L}) > 0$. The second inequality could be proved similarly.

For the last inequality, we just need to compare:

$$\alpha_H^u(\alpha_L^u - 1)[\alpha_H^e(\alpha_L^e - 1) - (1-\underline{p})(\alpha_L^e - \alpha_H^e)]$$

and

$$\alpha_H^e(\alpha_L^e - 1)[\alpha_H^u(\alpha_L^u - 1) - (1-\underline{p})(\alpha_L^u - \alpha_H^u)].$$

To prove 3, it suffices to show

$$\alpha_H^u(\alpha_L^u - 1)(\alpha_L^e - \alpha_H^e) > alpha_H^e(\alpha_L^e - 1)(\alpha_L^u - \alpha_H^u).$$

The direct proof is not easy. But notice from the expressions of α's:

$$(\alpha_L^e - \alpha_H^e)(\alpha_L^e + \alpha_H^e - 1) = 2(r+\delta)[\frac{\sigma^2}{(\Delta_L + \Delta_\xi)^2} - \frac{\sigma^2}{(\Delta_H + \Delta_\xi)^2}]$$

and

$$(\alpha_L^u - \alpha_H^u)(\alpha_L^u + \alpha_H^u - 1) = 2(r+\delta+\lambda)[\frac{\sigma^2}{\Delta_L^2} - \frac{\sigma^2}{\Delta_H^2}].$$

Hence, when $\Delta_\xi = 0$,

$$\frac{\alpha_L^e - \alpha_H^e}{\alpha_L^u - \alpha_H^u} = \frac{r+\delta}{r+\delta\lambda} \frac{\alpha_L^u + \alpha_H^u - 1}{\alpha_L^e + \alpha_H^e - 1}.$$

The original inequality is transformed to compare:

$$(r+\delta)\alpha_H^u(\alpha_L^u - 1)(\alpha_L^u + \alpha_H^u - 1)$$

and

$$(r+\delta+\lambda)\alpha_H^e(\alpha_L^e - 1)(\alpha_L^e + \alpha_H^e - 1).$$

Meanwhile, we have:

$$(r+\delta)\alpha_H^u(\alpha_L^u - 1)\alpha_L^u = (r+\delta)\alpha_H^u \frac{2(r+\delta+\lambda)}{\Delta_L^2}$$
$$> (r+\delta+\lambda)\alpha_H^e(\alpha_L^e - 1)\alpha_L^e = (r+\delta+\lambda)\alpha_H^e \frac{2(r+\delta)}{\Delta_L^2}$$

and

$$(r+\delta)\alpha_H^u(\alpha_L^u - 1)(\alpha_H^u - 1) = (r+\delta)(\alpha_L^u - 1)\frac{2(r+\delta+\lambda)}{\Delta_H^2}$$
$$> (r+\delta+\lambda)\alpha_H^e(\alpha_L^e - 1)(\alpha_H^e - 1) = (r+\delta+\lambda)(\alpha_L^e - 1)\frac{2(r+\delta)}{\Delta_H^2}$$

since $\alpha_y^u > \alpha_y^e$. This implies:

$$\alpha_H^u(\alpha_L^u - 1)(\alpha_L^e - \alpha_H^e) > \alpha_H^e(\alpha_L^e - 1)(\alpha_L^u - \alpha_H^u)$$

and therefore,

$$\frac{\alpha_H^u(\alpha_L^u - 1)\underline{p}}{\alpha_H^u(\alpha_L^u - 1) - (1-\underline{p})(\alpha_L^u - \alpha_H^u)} < \frac{\alpha_H^e(\alpha_L^e - 1)\underline{p}}{\alpha_H^e(\alpha_L^e - 1) - (1-\underline{p})(\alpha_L^e - \alpha_H^e)}.$$

Notice from the above proof, 3 holds only when Δ_ξ is small and will not hold as Δ_ξ becomes sufficiently large.

Finally, we can conclude that $\tilde{V}_H^u < \tilde{V}_H^e$ when $\xi_H \simeq \xi_L$, and as a result $\underline{p}^e < \underline{p}^u$. $\square$

No-deviation condition for the non-Bayesian learning example

Under the non-Bayesian learning case, suppose it is optimal for a p worker to choose firm y, the value function for this worker should be such that (from Hamilton-Jacobi-Bellman equation):

$$(r+\delta)W_y(p) = w_y(p) + \lambda_y p W_y'(p).$$

Suppose there is a cutoff $\underline{p}$ such that workers with $p > \underline{p}$ are matched with H firms and vice versa.

Then the absence of deviation implies that a $p > \underline{p}$ worker has no incentive to deviate, rematch with a L firm and switch back after dt time:

$$W_H(p) > \tilde{W}_L(p) = \mathbb{E}\left\{\int_t^{t+dt} e^{-(r+\delta)(s-t)} w_L(p_s) ds + e^{-(r+\delta)dt} W(p_{t+dt})\right\}.$$

For dt sufficiently small, p_{t+dt} is still close to p such that it is optimal for a p_{t+dt} worker to choose firm H as well. It is immediate to see that:

$$\lim_{dt\to 0} \frac{W_H(p) - \tilde{W}_L(p)}{dt} = w_H(p) - w_L(p) + (\lambda_H - \lambda_L) p W_H'(p),$$

and hence no deviation implies that:

$$w_H(p) - w_L(p) + (\lambda_H - \lambda_L) p W_H'(p) > 0$$

for all $p > \underline{p}$. Let $p \to \underline{p}+$ and we have by applying the value matching condition:

$$w_H(\underline{p}+) - w_L(\underline{p}-) + (\lambda_H - \lambda_L)\underline{p} W_H'(\underline{p}+) = \lambda_L \underline{p}(W_L'(\underline{p}-) - W_H'(\underline{p}+)) \geq 0$$

or equivalently $W_L'(\underline{p}-) \geq W_H'(\underline{p}+)$. On the other hand, a $p < \underline{p}$ worker also has no incentive to deviate, rematch with a H firm and switch back after dt time. Similarly, no deviation implies that:

$$w_L(p) - w_H(p) + (\lambda_L - \lambda_H)pW_L'(p) > 0$$

for all $p < \underline{p}$. Let $p \to \underline{p}-$ and it could be shown:

$$w_L(\underline{p}-) - w_H(\underline{p}+) + (\lambda_L - \lambda_H)\underline{p}W_L'(\underline{p}-) = \lambda_H\underline{p}(W_H'(\underline{p}+) - W_L'(\underline{p}-)) \geq 0$$

or equivalently $W_H'(\underline{p}+) \geq W_L'(\underline{p}-)$. Therefore, at $\underline{p}$, it must be the case that $W_H'(\underline{p}) = W_L'(\underline{p})$ and no-deviation condition coincides with the smooth-pasting condition.

Bibliography

Abreu, D., and F. Gul (2000): “Bargaining and Reputation,” *Econometrica*, 68(1), 85–117.

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

Admati, A., and M. Perry (1987): “Strategic Delay in Bargaining,” *Review of Economic Studies*, 54(3), 345–364.

Anderson, A., and L. Smith (2000): “Dynamic Matching and Evolving Reputations,” *Review of Economic Studies*, 77(1), 3–29.

Avery, C., and P. Zemsky (1994): “Option Values and Bargaining Delays,” *Games and Economic Behavior*, 7, 139–153.

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

Baron, D., and J. Ferejohn (1989): “Bargaining in Legislatures,” *American Political Science Review*, 83, 1181–1206.

Becker, G. (1973): “A Theory of Marriage: Part I,” *Journal of Political Economy*, 81, 813–846.

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.

Chatterjee, K., and L. Samuelson (1987): "Bargaining with Two-sided Incomplete Information: An Infinite Horizon Model with Alternating Offers," *Review of Economic Studies*, 54(2), 175–192.

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.

Cohen, A., and E. Solan (2009): "Bandit problems with Levy payoff processes," .

Cox, D., and H. Miller (1965): *The theory of stochastic processes*. Chapman and Hall.

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

Cripps, M., J. Ely, G. Mailath, and L. Samuelson (2008): "Common Learning," *Econometrica*, 76(4), 909–933.

Daley, B., and B. Green (2008): "Waiting for News in the Dynamic Market for Lemons," mimeo.

Damiano, E., H. Li, and W. Suen (2010a): "Delay in Strategic Information Aggregation," mimeo.

——— (2010b): "Optimal Deadlines for Agreements," mimeo.

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

Dumas, B. (1991): "Super contact and related optimality conditions," *Journal of Economic Dynamics and Control*, 15, 675–685.

Eeckhout, J., and X. Weng (2010): “The Value of Continued Learning,” mimeo.

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.

Faingold, E., and Y. Sannikov (2007): “Reputation Effects and Equilibrium Degeneracy in Continuous-Time Games,” Yale mimeo.

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

Groes, F., P. Kircher, and I. Manovskii (2009): “The U-Shapes of Occupational Mobility,” .

Harris, M., and B. Holmström (1982): “A Theory of Wage Dynamics,” *Review of Economic Studies*, 49(3), 315–333.

Hendricks, K., A. Weiss, and C. Wilson (1988): “The War of Attrition in Continuous Time with Complete Information,” *International Economic Review*, 29(4), 663–680.

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

Jovanovic, B. (1979): “Job Matching and the Theory of Turnover,” *Journal of Political Economy*, 87, 972–990.

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.

Kennan, J., and R. Wilson (1993): "Bargaining with Private Information," *Journal of Economic Literature*, 31(1), 45–104.

Kim, K., and F. Xu (2011): "Delay and Duplication of Information Acquisition in a War of Attrition," mimeo.

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

Merlo, A., and C. Wilson (1995): "A Stochastic Model of Sequential Bargaining with Complete Information," *Econometrica*, 63, 371–399.

Moscarini, G. (2005): "Job Matching and the Wage Distribution," *Econometrica*, 73(2), 481–516.

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`.

Papageorgiou, T. (2009): "Learning Your Comparative Advantages," .

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

Rubinstein, A. (1982): "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50(1), 97–110.

Sannikov, Y. (2007): "Games with Imperfectly Observable Actions in Continuous Time," *Econometrica*, 75(5), 1285–1329.

——— (2008): "A Continuous-Time Version of the Principal-Agent Problem," *Review of Economic Studies*, 75(3), 957–984.

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.

Yildiz, M. (2004): "Waiting to Pursuade," *Quarterly Journal of Economics*, 119(1), 223–248.