UNIVERSITY OF CALIFORNIA

Los Angeles

Essays on Strategic Experimentation

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Economics

by

Kichool Park

2001

The dissertation of Kichool Park is approved.

Michael J. Brennan

Joseph M. Ostroy

William R. Zame

David K. Levine, Committee Chair

University of California, Los Angeles

2001

To My Parents

Table of Contents

Page

ACKNOWLEDGMENTS

I would like to thank my dissertation committee members for all of their helpful comments**.** Amongst them, David Levine, Joe Ostroy, and Bill Zame need special mentioning. Their influences on me at the early stage of developing as a theorist are tremendous. Not only from the classes they taught, but also from numerous conversations out of the classroom, I benefited enormously, and I am very grateful. I thank Hongbin Cai, John Riley, Peter Sorensen, and seminar participants at University of Arizona, Southern California Theory Meeting at Santa Barbara, The First World Congress of Game Theory at Bilbao, Theory Workshop of Korean Econometric Society at Seoul for their valuable comments on Chapter 2. Dirk Vergemann and Juuso Välimäki commented in detail on Chapter 3, which is gratefully acknowledged. I owe a particular debt to my advisor David Levine. Without his encouragement and advice for years, it would have been impossible for me to finish this work.

Most of all, my deepest thanks go to my wife, Meekyoung Kim, and our two daughters, Hyunjee and Jaewoo, for their forbearance of those years they have lived apart from their husband and dad.

VITA

July 29, 1968	Born, Sisan, Korea
1992	B.A., Law Seoul National University Seoul, Korea
1999	M.A., Economics University of California, Los Angeles
1999	C.Phil., Economics University of California, Los Angeles
1997-1999	Teaching Assistant Department of Economics University of California, Los Angeles
2000	Medford Fellowship Department of Economics University of California
2000-2001	Chancellor's Dissertation Year Fellowship University of California, Los Angeles

PUBLICATIONS AND PRESENTATIONS

Park, K. (March, 2000): Leader-Follower Model of Strategic Experimentation. Presented at Southern California Economic Theory Conference, Santa Barbara.

Park, K. (July, 2000): Two Player Continuous Time Bandit Problem. Presented at The First World Congress of Game Theory, Bilbao, Spain.

ABSTRACT OF THE DISSERTATION

Essays on Strategic Experimentation

by

Kichool Park

Doctor of Philosophy in Economics

University of California, Los Angeles, 2001

Professor David K. Levine, Chair

This dissertation includes three essays on strategic experimentation. In Chapter 2, we investigate a game of strategic experimentation that also appears in Bolton and Harris (1999). The setting is a two-player continuous-time two-armed bandit problem. We will show that there is no mixed strategy equilibrium and that there are only two asymmetric pure strategy equilibria, both of which are simple in nature. In so doing, it will be shown that this game is a kind of simple coordination game. We will also show that the asymmetric equilibria are robust for a wide range

of parameters against the perturbations in the noise structure. Based on this result, we argue that even in the symmetric game with a homogeneous noise structure, the focus of the analysis should be on the asymmetric equilibria, not on the symmetric equilibrium that was studied in Bolton and Harris (1999).

In Chapter 3, we will investigate the possibility that the inefficiency in Chapter 2 could be overcome by market competition. Two competing firms are now supplying the two options. It will be shown that results vary according to whether or not consumers are homogeneous. If all the consumers have the same abilities to evaluate the uncertain product, it is possible to achieve the efficient outcome. If the consumers have heterogeneous abilities, then it is impossible to achieve the efficient outcome.

In Chapter 4, we turn to discrete-time models. Unlike continuous-time set up, we can provide a general model of multi-player multi-armed bandit problems in discrete-time set up. Under the assumption of perfect observability, we will generalize Section 2.2 in Berry and Fristedt (1985) to *n* player case. Then, we will show that there exist pure Markov strategy equilibria.

Chapter 1

Introduction

In a standard undergraduate textbook of microeconomics, the only mention of uncertainty is found in the chapter covering expected utility theory. Consumers are assumed to know exactly how much utility they will get from the consumption of each good. No matter how subtle the differences among the goods we assume, a consumer will understand them perfectly. Producers perceive market demands exactly. We even take this a step forward in the theory of duopoly in which each firm is assumed to know the other firm's cost function as well.

In reality, however, almost nothing is known for sure. No firm has perfect knowledge of the market demand function. No consumer knows fully all the minute differences of differentiated products in the market. This lack of full knowledge is a fundamental aspect of human life.

What is perhaps more important, but often overlooked, is that we obtain additional information almost every time we make an economic decision. A consumer who buys an apple at a supermarket will get some information about the taste of the specific kind of apple he decides to purchase. The owner of a gas station will obtain information about the local market demand for gas as he posts the price of gas each day. This newly acquired information will then have effects on the decisions of the consumer and of the owner of the gas station in the future. If the taste of the apple was very bad, the consumer may decide not to purchase that type of apple in the future. Depending on the sales record today, the owner of the gas station may want to increase or decrease the gas price tomorrow. These decisions, of course, may bring additional information which will in turn affect future decisions.

In some situations, people may act primarily to get information. If a consumer finds a new strain of apple in the supermarket, and if he never knew that type of apple exists before he entered the supermarket, he may decide to purchase it primarily to gain information about the new variety. At the cost of his favorite apple from which he could expect to get certain amount of joy, he is going to gather information about the new kind of apple. In other situations, the trade off between the acquisition of information and the sacrifice of the instant reward may be less explicit and subtler. This trade off between information and instant reward is fundamental. Thus, in order to understand various economic phenomena better, it is crucial to understand how a rational person experiments to get information.

Multi-armed bandit theory studies this issue in a most abstract form: Suppose we ought to select sequentially one of k alternatives. If the payoffs of those k alternatives are known only probabilistically, then what would be an optimal strategy to maximize the sum of the payoffs?

Some immediate applications of bandit problems are simplest search problems. It is, however, not appropriate in general to attempt to apply the results of bandit problems to economic analysis directly. In multi-armed bandit problems which are one person decision problems, the possibility of the existence of other experimenters, which is a usual feature in economic problems, is ignored. If there are other people, their experimentation may also provide opportunities to get information. Except in the case of Robinson Crusoe on a deserted island, we can always get some information which is relevant to our decisions by observing others' decisions and the results of those decisions. If the owner of the gas station sees the owner of a nearby gas station post a higher price than his and observes that the number of cars stopping by that station for fuel still does not decrease and that the number of cars stopping by his station does not increase either, then he would increase his price the next day. Even Robinson Crusoe can learn from the experience of his cohabitant, Friday. Since experimentation is costly, people will take advantage of the opportunity to observe others and to learn from their experimentation so as to lower their own costs of experimentation at the expense of others.

Analysis of equilibria in an environment in which experimentation decisions are made strategically is new to economists, and there is not much research in this area. The very first, and perhaps most challenging, question is how to keep track of people's beliefs as they are updated with newly acquired information. The most general answer to this question, and hence, the full understanding of the issue of strategic experimentation, are not yet obtained. Nevertheless, we will keep collecting pebbles on the beach of this newly found ocean, and provide in this dissertation some that we have found.

Outlines

In Chapter 2, we investigate a game of strategic experimentation that also appears in Bolton and Harris (1999). The setting is a two-player continuous-time two-armed bandit problem. We will show that there is no mixed strategy equilibrium and that there are only two asymmetric pure strategy equilibria, both of which are simple in nature. In so doing, it will be shown that this game is a kind of simple coordination game. We will also show that the asymmetric equilibria are robust for a wide range of parameters against the perturbations in the noise structure. Based on this result, we argue that even in the symmetric game with a homogeneous noise structure, the focus of the analysis should be on the asymmetric equilibria, not on the symmetric equilibrium that was studied in Bolton and Harris (1999).

In Chapter 3, we will investigate the possibility that the inefficiency in Chapter 2 could be overcome by market competition. Two competing firms are now supplying the two options. It will be shown that results vary according to whether or not consumers are homogeneous. If all the consumers have the same abilities to evaluate the uncertain product, it is possible to achieve the efficient outcome. If the consumers have heterogeneous abilities, then it is impossible to achieve the efficient outcome.

In Chapter 4, we turn to discrete-time models. Unlike continuous-time set up, we can provide a general model of multi-player multi-armed bandit problems in discrete-time set up. Under the assumption of perfect observability, we will generalize Section 2.2 in Berry and Fristedt (1985) to n player case. Then, we will show that there exist pure Markov strategy equilibria.

Chapter 2

Robust Equilibria in Strategic Experimentation

2.1 Introduction

In a recent paper analyzing strategic experimentation, Bolton and Harris (1999) found a symmetric mixed strategy equilibrium and examined its properties. However, it can be shown that their equilibrium is extremely fragile with respect to perturbations in the noise structure. By contrast, the asymmetric equilibria they did not analyze are robust to these perturbations. After characterizing all equilibria in the two-player game, we will argue that the main focus of analysis ought to be on asymmetric equilibria.

The game we consider is a variation of the two-player continuous-time bandit problem. Each player should divide their time between two options; one of which is "safe" and the other "risky." If the safe option is chosen, the player will be rewarded at a known rate. If the risky one is chosen, the player who chooses it will obtain a payoff which is a sum of the reward from the unknown parameter and noise. Thus, the reward rate for the risky option is unknown, which could be higher or lower than that of the safe one, and it should be learned over time. We assume that all the previous actions of each player and the realizations of the noisy payoffs are known to each player at all times. Therefore, each player will get information about the unknown reward rate of the risky option not only from her own experimentation, but also from that of the other player.

The main difference between our model and that in Bolton and Harris (1999) is in our assumption that the noise from the risky option could have different variances across individuals. To be specific, without loss of generality, σ_2, the variance of the noise added to the payoff from the risky option to player 2, will be smaller than σ_1, the variance of the noise added to the payoff from the risky option to player 1. This perturbed game will serve to test the robustness of equilibria of the original game in Bolton and Harris (1999), where they assumed homogeneity of the noise structure.

This model is important to understand, since a lot of situations fit the description of "strategic experimentation." Examples include natural resource exploration, adoption of new institution, technologies, or products, research and development, and consumer search. None of these examples will be as simple as our model. Overall, however, we agree with Bolton and Harris (1999) that multi-agent two-armed bandit problem will be the backbone of multi-agent active learning theories. The depth of the current understanding of this problem, however, is not satisfactory. Thus, as a way of attaining a better understanding, we confine our focus to the 2 player case, and attempt to do full analysis. The basic results are as follows.

We first show that in addition to the symmetric mixed strategy equilibria Bolton and Harris (1999) investigated, the original game has two more equilibria, which are asymmetric pure strategy equilibria. In every perturbed game, however, there are only two asymmetric pure strategy equilibria. There is no mixed strategy equilibrium in perturbed game. The structures of the asymmetric pure strategy equilibria are very simple; either player 1 is experimenting more or player 2 is. To explain this in more detail, let p be the posterior probability that the risky option has a higher reward rate. Then, at one equilibrium, there will be two cutoff points $0 < c_1 < c_2 < 1$ such that player 1 chooses the risky option if and only if $p > c_1$ and player 2 chooses the risky one if and only if $p > c_2$. At the other equilibrium, there still are two cutoffs $0 < c_1' < c_2' < 1$, but now, player 1 and 2 choose the risky option if and only if $p > c_2'$ and $p > c_1'$, respectively.

Moreover, by showing that the symmetric mixed strategy equilibrium disappears as soon as we add heterogeneity into the structure of the noise, and that, by contrast, the asymmetric pure strategy equilibria are robust against this perturbations, we select the asymmetric pure strategy equilibria against the symmetric mixed strategy equilibrium in the original game. More precisely, in all pure strategy equilibria, the length of the interval of p on which only one of the players is experimenting is shown to be bounded away from zero as σ_2 converges to σ_1.

Some features of these asymmetric equilibria are worth emphasizing. As the experimentation of player 2 will provide more accurate information than that of player 1 does, social optimality requires that player 2 should choose the risky option in case only one of them experiments. There exists, however, an equilibrium at which the player with noisier signal chooses the risky option earlier, which is opposite to the efficient allocation. It will be shown that even the equilibrium where player 2 begins to choose the risky option first is inefficient due to the free riding incentives. At both asymmetric equilibria, players will cease to select the risky option for some beliefs even though social optimality requires there to be experimentation. Hence, we have too little experimentation for these beliefs, and thus too little social learning. The *ex ante* probability of adopting the better product is less than optimal.

Lastly, it will be shown that the perturbed game is a kind of coordination game. Depending on which type of equilibria they are playing, the payoff to each player will be determined accordingly. For example, if player 1 is doing more experimentation and player 2 is free riding, then the payoff to player 2 is greater than that of player 1.

This chapter is organized as follows. Section 2 describes the game. Section 3 summarizes the basic results that will be useful in later sections. All the results in Section 3 are straightforward modifications of those in Bolton and Harris (1999). The team problem is studied in Section 4. In Section 5, we show that the structure of the equilibrium strategy profile in our model is fairly simple. We show this in three steps. Firstly, it will be shown that there is no symmetric equilibrium. Second, we prove a characterization lemma for mixed strategy equilibria. Lastly, by showing that there is no mixed strategy equilibrium that satisfies this condition, we show that the only possible equilibrium is a pure strategy equilibrium. In Section 6, we provide one characterization theorem for pure strategy equilibria profiles, and with it, we show that asymmetric pure strategy equilibria are robust to perturbations in σ_2, given σ_1. In Section 7, it is shown that the game we analyze is simply a coordination game. The player engaged in less experimentation receives a higher payoff. The existence proof of asymmetric equilibria is given in Section 8.

2.2 The Model

There are 2 infinitely lived, risk-neutral players who will be denoted by $1, 2$. At each time period $[t, t+dt)$, each player has to decide how to allocate her time between two alternatives, one of which is safe and the other risky. Their choices are made simultaneously and independently. For player $i = 1, 2$, if she devotes a proportion of α_i of the current period $[t, t+dt)$ to the risky option, she will receive the total payoff

$$d\pi_i^0 = (1-\alpha_i)sdt + (1-\alpha_i)^{1/2}\sigma dZ_i^0(t)$$

from the safe option, and the total payoff

$$d\pi_i^1 = \alpha_i \mu dt + {\alpha_i}^{1/2}\sigma_i dZ_i^1(t)$$

from the risky option.

We assume that 1) s is known to both players, 2) neither of players knows μ, although the value of μ is fixed, 3) μ can be either h or l, where $0 < l < s < h$, 4) the $dZ_i^k(t)$ are the independent standard Brownian motions for $k \in \{0, 1\}$, $i = 1, 2$. Thus, by choosing the risky option, each player could get information about μ, although this information is subject to noise.

Since the $dZ_i^k(t)$ are standard Brownian motions, $dZ_i^k(t)$ will be distributed following normal distribution whose mean and variance are zero and dt, respectively. This implies that $d\pi_i^0$ is distributed normally with mean $(1-\alpha_i)sdt$ and variance $(1-\alpha_i)\sigma^2 dt$, and that $d\pi_i^1$ is distributed with mean $\alpha_i \mu dt$ and variance

$\alpha_i \sigma_i^2 dt$. Regarding the variances of the noise, we will assume that they will be different across players. This assumption implies that the signals received by the two players will differ in quality. Without loss of generality, we will assume that $0 < \sigma_2 < \sigma_1$. The signal for player 2 will be less noisy. We could interpret this assumption as each player having different technologies to evaluate the true value of the risky option. All the players have the same objective: to maximize the present discounted value of their payoff streams, namely $E[\int_0^\infty re^{-rt}(d\pi_i^0 + d\pi_i^1)(t)]$, where r is the common discount factor for the players. The meaning of this expectation will be clearer in the next section.

We will assume that each player's decision and her payoff in each period, $(\alpha_i(t), \pi_i^0(t), \pi_i^1(t))_{i=1,2}$, will be known to everyone at the beginning of the preceding period. Thus, $\{(\alpha_i(t),\ \pi_i^0(t),\ \pi_i^1(t))_{i=1,2}\}_{t<\tilde{t}}$ are common knowledge to the players at time $\tilde{t}$, for all $\tilde{t}$.

2.3 Basic Results[1]

Most of the results in this section are straightforward modifications of those in Bolton and Harris(1999).

[1] All the results in this section can be generalized in a straight forward way to the game of N players, where $0 < \sigma_N < .... < \sigma_1$.

2.3.1 The Filtering Problem

Following Bolton and Harris(1999), we will focus only on the perfect equilibria in stationary Markov strategies. Hence the properties of p, the posterior probability of μ being h, which is the natural state variable in our model, are important for our analysis.

Suppose player i devotes the proportion α_i of the current period $[t, t+dt)$ to the risky option, and let $p(t)$ be the prior probability that μ is h at time t, and $p(t+dt)$ be the posterior at the end of the current period. Let $dp(t) = p(t+dt) - p(t)$. We need to characterize the distribution of $dp(t)$. This problem is usually called a filtering problem.

Proposition 2.1 *Conditional on the information available to players at time t, the change in beliefs $dp(t)$ is distributed normally with mean 0 and variance $(\alpha_1/\sigma_1^2 + \alpha_2/\sigma_2^2)\Phi(p(t))dt$, where $\Phi(p) = [p(1-p)(h-l)]^2$.*

Proof . Since players derive the information about μ only from $d\pi_i^1(t)$, these payoffs are observationally equivalent to the signals $d\tilde{\pi}_i^1 = (\alpha_i)^{1/2}\tilde{\mu}_i + dZ_i^1(t)$, where $\tilde{\mu}_i = \mu/\sigma_i$. Note that $\tilde{\mu}_i$ takes the values $\tilde{l}_i = l/\sigma_i$ or $\tilde{h}_i = h/\sigma_i$ with probabilities $(1-p)$ and p. Therefore, applying Bayes' Rule,

$$p(t+dt) = \frac{p(t)F(\tilde{h}_1, \tilde{h}_2)}{p(t)F(\tilde{h}_1, \tilde{h}_2) + (1-p(t))F(\tilde{l}_1, \tilde{l}_2)},$$

where $F(\tilde{\mu}_1, \tilde{\mu}_2) = (2\pi dt)^{-1} \exp\{-(1/2dt) \sum_{i=1}^{2} (d\tilde{\pi}_i^1(t) - (\alpha_i)^{1/2} \tilde{\mu}_i dt)^2\}$ is the probability of observing the payoff profile $d\tilde{\pi}^1(t) = \prod_{i=1}^{2} d\tilde{\pi}_i^1(t)$ given $(\tilde{\mu}_1, \tilde{\mu}_2)$. Hence,

$$\begin{aligned} dp &= p(t+dt) - p(t) \\ &= \frac{p(1-p)(\tilde{F}(\tilde{h}_1, \tilde{h}_2) - \tilde{F}(\tilde{l}_1, \tilde{l}_2))}{p\tilde{F}(\tilde{h}_1, \tilde{h}_2) + (1-p)\tilde{F}(\tilde{l}_1, \tilde{l}_2)}, \end{aligned}$$

where $\tilde{F}(\tilde{\mu}_1, \tilde{\mu}_2) = \exp\{\sum_{i=1}^{2} (\alpha_i)^{1/2} \tilde{\mu}_i d\tilde{\pi}_i^1 - 1/2 \sum_{i=1}^{2} \alpha_i \tilde{\mu}_i^2 dt\}$, and we have suppressed the dependence of variables on t. Note that

$$\begin{aligned} \tilde{F}(\tilde{\mu}_1, \tilde{\mu}_2) &= 1 + (\sum_{i=1}^{2} (\alpha_i)^{1/2} \tilde{\mu}_i d\tilde{\pi}_i^1 - 1/2 \sum_{i=1}^{2} \alpha_i \tilde{\mu}_i^2 dt) \\ &\quad + \frac{1}{2} \{(\sum_{i=1}^{2} (\alpha_i)^{1/2} \tilde{\mu}_i d\tilde{\pi}_i^1 - 1/2 \sum_{i=1}^{2} \alpha_i \tilde{\mu}_i^2 dt)\}^2 \\ &= 1 + \sum_{i=1}^{2} (\alpha_i)^{1/2} \tilde{\mu}_i d\tilde{\pi}_i^1, \end{aligned}$$

where we have dropped the terms of order $dt^{3/2}$ and higher, and where we have used the fact that $(d\tilde{\pi}_i^1)^2 = dt$ and $d\tilde{\pi}_i^1 d\tilde{\pi}_j^1 = 0$ if $i \neq j$, respectively. Hence,

$$\begin{aligned} dp &= \frac{p(1-p)(h-l) \sum_{i=1}^{2} ((\alpha_i)^{1/2}/\sigma_i) d\tilde{\pi}_i^1}{1 + \sum_{i=1}^{2} (\alpha_i)^{1/2} \tilde{m}_i(p) d\tilde{\pi}_i^1} \\ &= p(1-p)(h-l)(\sum_{i=1}^{2} \frac{(\alpha_i)^{1/2}}{\sigma_i} d\tilde{\pi}_i^1)(1 - \sum_{i=1}^{2} (\alpha_i)^{1/2} \tilde{m}_i(p) d\tilde{\pi}_i^1) \\ &= p(1-p)(h-l)(\sum_{i=1}^{2} \frac{(\alpha_i)^{1/2}}{\sigma_i} d\tilde{\pi}_i^1 - \sum_{i=1}^{2} \alpha_i \tilde{m}_i(p) dt) \\ &= p(1-p)(h-l) \sum_{i=1}^{2} \frac{(\alpha_i)^{1/2}}{\sigma_i} d\tilde{Z}_i^1, \end{aligned}$$

where $\tilde{m}_i(p) = [(1-p)l + ph]/\sigma_i$, and $d\tilde{Z}_i^1 = d\tilde{\pi}_i^1 - (\alpha_i)^{1/2}((1-p)l + ph)dt$. We have used again the fact that $(d\tilde{\pi}_i^1)^2 = dt$ and $d\tilde{\pi}_i^1 d\tilde{\pi}_j^1 = 0$ if $i \neq j$, respectively, and neglected the terms of order $dt^{3/2}$and higher. Clearly $\tilde{Z}^1 = \prod_{i=1}^2 \tilde{Z}_i^1$ is a standard 2-dimensional Wiener process. Therefore, dp has mean 0 and variance $[p(1-p)(h-l)]^2(\sum_{i=1}^2 \alpha_i/\sigma_i^2)$. ■

This proposition explains the main reason we are using continuous-time model. In discrete-time model, it is difficult to describe the posterior beliefs in a tractable way.

From Proposition 2.1, we can see that $\Phi(0) = \Phi(1) = 0$. Therefore, once they become sure about μ, from that point on, there will be no further change in p, which is a common feature of Bayesian updating. Also, players' decision α_i's have weighted effects on the variance of $dp(t)$, and these weights are the inverses of σ_i^2. Hence, the more accurate player i's signal is, the greater will be the influence of the proportion of his time spent on the risky option on the variance of $dp(t)$.

2.3.2 Properties of The Value Function

Our primary interest in this paper is in Markov perfect equilibria. Therefore, we will not attempt to define the set of strategies for each player rigorously, and move directly to Markov strategies. A Markov strategy is a strategy dependent only on p, which is our natural state variable in this model.

Definition 2.2 **A Markov strategy** is a measurable function from $[0, 1]$ to

$[0,1]$.

Since $p(t)$ has a well-defined distribution, and since we are considering only Markov strategies which are functions of p, it is clear that $E[\int_0^\infty re^{-rt}(d\pi_i^0+d\pi_i^1)(t)]$ in the previous section is well-defined.

Let $\mathcal{M}_i$ be the set of Markov strategies for player i. Given any notation with subscript, such as a_i, following convention, we will denote a_{3-i} by a_{-i} for $i = 1, 2$.

Definition 2.3 A strategy profile $s = (s_1, s_2)$ is a **Nash equilibrium** if s_i is a best response for player i against s_{-i} for all $i = 1, 2$. A strategy profile $\xi = (\xi_1, \xi_2)$ is a **subgame perfect Markov equilibrium** if ξ is a Nash equilibrium and ξ_i is a Markov strategy for all $i = 1, 2$.

Henceforth, we will simply call a Markov strategy a strategy except where there is risk of confusion.

Let $m(p) = ph + (1-p)l$ be the myopic expected payoff from the risky option when μ is believed to be h with probability p. Then, player i's value function, when the other player's strategy is ξ_{-i}, will be the unique solution of the Hamilton-

Jacobi-Bellman equation[2]

$$\begin{aligned}u_i(p) &= \max_{0\le\alpha_i\le 1}\{(1-\alpha_i)s+\alpha_i m(p) \\ &\quad +\frac{1}{r}\left(\frac{\alpha_i}{\sigma_i^2}+\frac{\xi_{-i}(p)}{\sigma_{-i}^2}\right)\Phi(p)\frac{u_i''(p)}{2}\}.\end{aligned} \tag{2.1}$$

Thus, as in discrete-time dynamic programming equation, player i's value function will be sum of two parts. Firstly, $(1-\alpha_i)s+\alpha_i m(p)$ is the instant expected payoff of devoting α_i of her time to the risky option. The second part

$$\frac{1}{r}\left(\frac{\alpha_i}{\sigma_i^2}+\frac{\xi_{-i}(p)}{\sigma_{-i}^2}\right)\Phi(p)\frac{u_i''(p)}{2}$$

is the discounted expected value of the changes in u_i. As in the discrete-time case, given this dynamic programming formulation, a strategy ξ_i will be a best response to ξ_{-i} if and only if

$$\begin{aligned}\xi_i &\in \arg\max_{0\le\alpha_i\le 1}\{(1-\alpha_i)s+\alpha_i m(p) \\ &\quad +\frac{1}{r}\left(\frac{\alpha_i}{\sigma_i^2}+\frac{\xi_{-i}(p)}{\sigma_{-i}^2}\right)\Phi(p)\frac{u_i''(p)}{2}\}\end{aligned}$$

for all $p\in[0,1]$.

Since

$$\begin{aligned}u_i(p) &= \max_{0\le\alpha_i\le 1}\{(1-\alpha_i)s+\alpha_i m(p) \\ &\quad +\frac{1}{r}(\frac{\alpha_i}{\sigma_i^2}+\frac{\xi_{-i}(p)}{\sigma_{-i}^2})\Phi(p)\frac{u_i''(p)}{2}\} \\ &= \max_{0\le\alpha_i\le 1}\{s+\frac{1}{r}\frac{\xi_{-i}(p)}{\sigma_{-i}^2}\Phi(p)\frac{u_i''(p)}{2}\end{aligned}$$

[2] The Hamilton-Jacobi-Bellman equation is the dynamic programming equation in continuous-time model. As an introduction, see Dixit and Pindyck (1994).

$$+\alpha_i \left(\frac{1}{r\sigma_i^2} \Phi(p) \frac{u_i''(p)}{2} - (s - m(p)) \right) \},$$

a strategy ξ_i is a best response to ξ_{-i} if and only if

$$\xi_i(p) = \begin{cases} 0 & \text{if } \ (1/(r\sigma_i^2))\Phi(p)u_i''(p)/2 < s - m(p) \\ 1 & \text{if } \ (1/(r\sigma_i^2))\Phi(p)u_i''(p)/2 > s - m(p) \\ \in [0,1] & \text{if } \ (1/(r\sigma_i^2))\Phi(p)u_i''(p)/2 = s - m(p). \end{cases}$$

The above result has an immediate interpretation. $s - m(p)$ will be the opportunity cost of choosing the risky option for player i, while $(1/(r\sigma_i^2))\Phi(p)u_i''(p)/2$ measures the informational gain from her experimentation. Thus, player i will choose the risky option or the safe one depending on whether or not the informational gain outweighs the opportunity cost.

We will characterize properties of player i's value function, which will be useful for our analysis. Let $\underline{u}(p) = \max\{s,\, m(p)\}$, which is a myopic payoff, let $\overline{u}(p) = (1-p)s + ph$, which is the full-information ex-ante payoff, and let b be the myopic break-even point such that $m(b) = s$.

Proposition 2.4 *Suppose that player i plays a best response to the strategy profile ξ_{-i}. Let u_i be her value function. Then $\underline{u} \leq u_i \leq \overline{u}$ and $u_i'' \geq 0$ on $[0,1]$. In particular, $u_i''(b) > 0$.*

Proof . Since one possible strategy for her is to choose the safe option when $p \in [0, b]$ and the risky one when $p \in (b, 1]$, $\underline{u} \leq u_i$ is clearly true. Also since any strategy for her in incomplete information case will be also a strategy in complete information case, $u_i \leq \overline{u}$ is immediate. For the second part, note that the Bellman equation (2.1) holds if and only if

$$u_i \geq m + \frac{1}{r}(\frac{1}{\sigma_i^2} + \frac{\xi_{-i}}{\sigma_{-i}^2})\Phi\frac{u_i''}{2} \quad \text{and} \quad u_i \geq s + \frac{\xi_{-i}}{r\sigma_{-i}^2}\Phi\frac{u_i''}{2}$$

with at least one equality.

If the first inequality holds as an equality, then we have

$$\frac{1}{r}\Phi\frac{u_i''}{2} = \frac{u_i - m}{(1/\sigma_i^2 + \xi_{-i}/\sigma_{-i}^2)} \geq 0.$$

Similarly, if the second inequality holds as an equality and $\xi_{-i}/\sigma_{-i}^2 > 0$, then

$$\frac{1}{r}\Phi\frac{u_i''}{2} = \frac{u_i - s}{(\xi_{-i}/\sigma_{-i}^2)} \geq 0.$$

If the second inequality holds as an equality and $\xi_{-i}/\sigma_{-i}^2 = 0$, then $u_i = s$. Since $u_i \geq s$ on $[0,1]$, u_i attains a minimum in this case, and hence $u_i'' \geq 0$. Therefore, overall, $u_i'' \geq 0$ on $[0,1]$.

Lastly, suppose $u_i''(b) = 0$. Then, by (2.1) again, $u_i(b) = \underline{u}(b)$. With the first part of this Lemma, however, this implies that u_i has an upward kink at $p = b$, which is a contradiction to the fact that u is continuously differentiable for all $p \in [0,1]$. Hence $u_i''(b) > 0$. ■

It is easy to prove the following proposition using Proposition 2.4.

Proposition 2.5 *Player i's value function u_i is a non-decreasing, convex* *function.*

Proof . Since $u_i'' \geq 0$, convexity is trivial. Since $u_i(0) = s$, and since $u(p) \geq s$ for all $p \in [0, 1]$ from Proposition 2.4, it is also clear that u_i is non-decreasing. ■

If $p = b$, then each player will be myopically indifferent between the safe option and the risky one. By choosing the risky one, however, in addition to the instant payoffs, they could also get extra information about the risky option. Therefore, choosing the risky option is a dominant strategy for each player if $p \geq b$.

Proposition 2.6 *If, for player i, ξ_i is a best response to ξ_{-i} for some $\xi_{-i} \in M_{-i}$, then for all $p \in [b, 1]$, $\xi_i(p) = 1$.*

The following lemma which says that the payoff for player i will not decrease if the opponent will increase her experimentation can be proved using the fact that $u_i'' \geq 0$, and it will be useful when we prove the existence of equilibrium in Section 8.

Proposition 2.7 *Let ξ_{-i} and $\hat{\xi}_{-i}$ be strategy profiles of the other player, and let u_i and $\hat{u}_i$ be the value functions of player i when she plays a best response to ξ_{-i} and $\hat{\xi}_{-i}$, respectively. If $\xi_{-i} \geq \hat{\xi}_{-i}$, then $u_i \geq \hat{u}_i$.*

Proof . The value function u_i satisfies the Bellman equation

$$u_i(p) = \max_{0 \le \alpha_i \le 1} \{(1-\alpha_i)s + \alpha_i m(p) + \frac{1}{r}(\frac{\alpha_i}{\sigma_i^2} + \frac{\xi_{-i}(p)}{\sigma_{-i}^2})\Phi(p)\frac{u_i''(p)}{2}\}.$$

Therefore,

$$u_i(p) \ge \max_{0 \le \alpha_i \le 1} \{(1-\alpha_i)s + \alpha_i m(p) + \frac{1}{r}(\frac{\alpha_i}{\sigma_i^2} + \frac{\hat{\xi}_{-i}(p)}{\sigma_{-i}^2})\Phi(p)\frac{u_i''(p)}{2}\}.$$

Comparing the inequality with the Bellman equation for $\hat{u}_i$

$$\hat{u}_i(p) = \max_{0 \le \alpha_i \le 1} \{(1-\alpha_i)s + \alpha_i m(p) + \frac{1}{r}(\frac{\alpha_i}{\sigma_i^2} + \frac{\hat{\xi}_{-i}(p)}{\sigma_{-i}^2})\Phi(p)\frac{\hat{u}_i''(p)}{2}\},$$

from the positivity of the Bellman operator, we conclude that $u_i \ge \hat{u}_i$. ■

2.4 The Team Problem

As a benchmark for our equilibrium analysis, we will investigate the team problem first. Since the two players will have signals of different quality, if social optimality requires only one of them to select the risky option, we would naturally conjecture that player 2 should select it. Indeed, at the socially optimal allocation, there will be two cutoffs $0 < c_2 < c_1 < 1$ such that player i will choose the risky option if and only if $p \in (c_i, 1]$. In this section, we will prove this result.

In the team problem, a social planner will maximize the average payoff of the two players. Hence,

Lemma 2.8 *The value function u_* for the team problem is the unique solution of the Bellman equation*

$$\begin{aligned} u_*(p) &= \max_{0\le\alpha_1,\alpha_2\le 1}\{\frac{1}{2}[(2-\sum_{i=1}^{2}\alpha_i)s+(\sum_{i=1}^{2}\alpha_i)m(p)] \\ &\quad +\frac{1}{r}(\sum_{i=1}^{2}\frac{\alpha_i}{\sigma_i^2})\Phi(p)\frac{u_*''(p)}{2}\}, \end{aligned}$$

and a strategy profile $\{\xi_i\}$ is an optimal policy for the team problem if and only if

$$\begin{aligned} \{\xi_i\} &\in \arg\max_{0\le\alpha_1,\alpha_2\le 1}\{\frac{1}{2}[(2-\sum_{i=1}^{2}\alpha_i)s+(\sum_{i=1}^{2}\alpha_i)m(p)] \\ &\quad +\frac{1}{r}(\sum_{i=1}^{2}\frac{\alpha_i}{\sigma_i^2})\Phi(p)\frac{u_*''(p)}{2}\}. \end{aligned}$$

Note that

$$\begin{aligned} u_*(p) &= \max_{0\le\alpha_1,\alpha_2\le 1}\{\frac{1}{2}[(2-\sum_{i=1}^{2}\alpha_i)s+(\sum_{i=1}^{2}\alpha_i)m(p)] \\ &\quad +\frac{1}{r}(\sum_{i=1}^{2}\frac{\alpha_i}{\sigma_i^2})\Phi(p)\frac{u_*''(p)}{2}\} \\ &= \max_{0\le\alpha_1,\alpha_2\le 1}\{s+\sum_{i=1}^{2}\alpha_i\left(\frac{1}{r\sigma_i^2}\Phi(p)\frac{u_*''(p)}{2}-\frac{s-m(p)}{2}\right)\}. \end{aligned}$$

Since $0<\sigma_2<\sigma_1$ by assumption, it is clear that if $\{\xi_i\}$ is optimal, then

$$\{\xi_i\}=\begin{cases} 0 \quad \text{for all } i=1,2 & \text{if } (1/(r\sigma_2^2))\Phi u_*''/2<(s-m)/2 \\ 1 \quad \text{for } i=2, \text{ and } \quad 0 \quad \text{for } i=1 & \text{if } (1/(r\sigma_2^2))\Phi u_*''/2>(s-m)/2 \\ & >(1/(r\sigma_1^2))\Phi u_*''/2 \\ 1 \quad \text{for all } i=1,2 & \text{if } \ (1/(r\sigma_1^2))\Phi u_*''/2>(s-m)/2, \end{cases}$$

where we suppress the dependence of Φ, u''_*, and m on p.

Hence the structure of the optimal policy of this problem is very simple. In general, there will be two cutoff points $0 < c_2 < c_1 < 1 = c_0$ so that if $p \in (c_1, 1]$, then both of the players should select the risky option, if $p \in (c_2, c_1]$, only player 2 ought to select the risky one, and if $p \in [0, c_2]$, then both should select the safe one. This structure of the optimal policy is fairly intuitive. We can interpret $(s-m(p))/2$ as the common opportunity cost for choosing the risky option. Therefore, all the players are facing the same conditions regarding cost. The differences in quality of signals, however, will give them different benefits from experimentation, which is $(1/(r\sigma_i^2))\Phi(p)u''_*(p)/2$, and therefore, they will have different break-even points. Social optimality requires that, unless they are too pessimistic or too optimistic, only player 2 experiment.

With this structure of the optimal policy in mind, we now know that for $p \in (c_{j+1}, c_j]$, the value function u_* is the solution of the second order differential equation

$$u_*(p) = \frac{1}{2}[js + (2-j)m(p)] + \frac{1}{r}(\sum_{i=j+1}^{2} \frac{1}{\sigma_i^2})\Phi(p)\frac{u''_*(p)}{2}$$

for $j = 0, 1$. Of course, if $p \in [0, c_2]$, then $u_* = s$.

It can be seen by direct calculation that

$$\begin{aligned} u_*(p) &= \frac{1}{2}[s + m(p)] \\ &\quad + \beta_1 p^{(\zeta_1+1)/2}(1-p)^{-(\zeta_1-1)/2} + \beta_2 p^{-(\zeta_1-1)/2}(1-p)^{(\zeta_1+1)/2}, \end{aligned}$$

where

$$\zeta_1 = (1 + 8r\sigma_2^2/(h-l)^2)^{1/2}$$

is the general solution of this differential equation for $p \in (c_2, c_1]$. For $p \in (c_1, 1]$, since u_* is bounded by h,

$$u_{*0}(p) = m(p) + \beta_0 p^{-(\zeta_0 - 1)/2}(1-p)^{(\zeta_0+1)/2}$$

will be the solution, where

$$\zeta_0 = \left(1 + \frac{8r}{(h-l)^2(1/\sigma_1^2 + 1/\sigma_2^2)}\right)^{1/2}.$$

To obtain the final solution, we need to specify 2 cutoff points c_1, c_2 in addition to the 3 parameters, $(\beta_0, \beta_1, \beta_2)$, and they will be determined by boundary conditions which include value matching conditions and smooth pasting conditions.[3]

Let $u_*(p)$ over $(c_{j+1}, c_j]$ be $u_{*j}(p)$ for $j = 0, 1$. Then the value matching conditions will be

$$u_{*0}(c_1) = u_{*1}(c_1), \text{ and } u_{*1}(c_2) = s.$$

The smooth pasting conditions for the value functions, u_{*j}'s, are

$$u'_{*0}(c_1) = u'_{*1}(c_1), \text{ and } u'_{*1}(c_2) = 0.$$

Since we have five parameters to determine, we need one more boundary condition, and the following theorem provides it.

---

[3] For an introductory explanation for these conditions, see Dixit and Pindyck (1994), Ch.4.

Lemma 2.9 *At the optimal solution of the team problem,*

$$u''_{*1}(c_1) = u''_{*0}(c_1).$$

Proof . In the following proof, we will extend the domains of u_{*0} and u_{*1} to $(0,1)$, if necessary. Suppose $u''_{*1}(c_1) > u''_{*0}(c_1)$. Since

$$\frac{s - m(p)}{2} \leq \frac{1}{r\sigma_1^2}\Phi(p)u''_{*0}(p)/2$$

for all $p \in (c_1, 1]$, and since u''_{*0} and u''_{*1} are continuous, there exists $\varepsilon > 0$ such that

$$\frac{s - m(p)}{2} < \frac{1}{r\sigma_1^2}\Phi(p)u''_{*1}(p)/2$$

for all $p \in (c_1 - \varepsilon, c_1 + \varepsilon)$, which is contradictory to the optimality of $\xi_1 = 0$ for $p \in [0, c_1]$.

Suppose $u''_{*1}(c_1) < u''_{*0}(c_1)$. Optimality and continuity imply that

$$\frac{1}{r\sigma_1^2}\Phi(c_1)\frac{u''_{*1}(c_1)}{2} \leq \frac{s - m(c_1)}{2} \leq \frac{1}{r\sigma_1^2}\Phi(c_1)\frac{u''_{*0}(c_1)}{2}.$$

We first show that it is impossible to have the second inequality above hold with equality. If it holds with equality, then

$$\begin{aligned} u_*(p) &\rightarrow m(c_1) + \frac{1}{r}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)\Phi(c_1)\frac{u''_{*0}(c_1)}{2} \\ &= \frac{s + m(c_1)}{2} + \frac{1}{r\sigma_2^2}\Phi(c_1)\frac{u''_{*0}(c_1)}{2} \end{aligned}$$

as $p \searrow c_1$. Also

$$u_*(p) \to \frac{s + m(c_1)}{2} + \frac{1}{r\sigma_2^2}\Phi(c_1)\frac{u''_{*1}(c_1)}{2}$$

as $p \nearrow c_1$. Since $u''_{*1}(c_1) < u''_{*0}(c_1)$, however, u_* can't be continuous at c_1. Thus,

$$\frac{1}{r\sigma_1^2}\Phi(c_1)\frac{u''_{*1}(c_1)}{2} \leq \frac{s - m(c_1)}{2} < \frac{1}{r\sigma_1^2}\Phi(c_1)\frac{u''_{*0}(c_1)}{2}.$$

Since u''_{*0} is continuous, this implies that there exists $\varepsilon > 0$ such that

$$\frac{s - m(p)}{2} < \frac{1}{r\sigma_1^2}\Phi(p)\frac{u''_{*0}(p)}{2}$$

for all $p \in (c_1 - \varepsilon, c_1)$. Also, 1) u_{*1} and u_{*0} are convex, 2) $u_{*0}(c_1) = u_{*1}(c_1)$, and $u'_{*0}(c_1) = u'_{*1}(c_1)$, 3) $u''_{*1}(c_1) < u''_{*0}(c_1)$ altogether imply that there exists $\varepsilon' > 0$ such that $u_{*0}(p) > u_{*1}(p)$ for all $p \in (c_1 - \varepsilon', c_1)$. This, however, contradicts the optimality for player 1 of ceasing to choose the risky option if $p \leq c_1$. ■

It is notable that $u''_{*1}(c_1) = u''_{*0}(c_1)$ implies that, when the social planner allocates player 1's time, she ought to be indifferent between the safe option and the risky one at $p = c_1$. This could also be shown as following by invoking a stopping time, which will show more clearly that this boundary condition comes from the fact that $p(t)$ follows Brownian motion. We will show that the social planner can't strictly prefer player 1 choosing the safe option at $p = c_2$. The argument for the other case is similar.

Suppose so. Let τ_1 be a stopping time which is induced from ξ_i, adapted to an obvious filtration, and indicates the first time player 1 will choose the risky option assuming that $p(0) \in (c_1, c_2)$. Let t_1 be the first time such that $p(t_1) = c_1$. Because of the path properties of Brownian motion, for any $\varepsilon > 0$, $\Pr(p(t) > c_1$ for some $t_1 < t < t_1 + \varepsilon | p(t_1) = c_1) = 1$. Therefore, $\Pr(\tau_1 < t_2 + \varepsilon | \tau > t_2) = 1$ for all $\varepsilon > 0$. This implies, however, that $\Pr(\tau \leq t_2 | \tau > t_2) = 1$, which is a contradiction.

From this, it follows immediately that all the parameters are uniquely determined.

Theorem 2.10 *At the optimal solution of the team problem, there will be two cutoff points* $0 < c_2 < c_1 < 1$ *so that*

$$\xi_1(p) = \begin{cases} 0 & \text{if } p \in [0, c_1] \\ 1 & \text{if } p \in (c_1, 1] \end{cases}, \text{ and } \xi_2(p) = \begin{cases} 0 & \text{if } p \in [0, c_2] \\ 1 & \text{if } p \in (c_1, 1] \end{cases}.$$

The two cutoff points c_1 *and* c_2 *will be unique.*

2.5 The Leader-Follower Model of Strategic Experimentation

We will call the 2 player game in our analysis Leader-Follower Model of Strategic Experimentation (LFMSE). Sometimes, to emphasize the importance of the parameters σ_1 and σ_2, we say LFMSE with parameters (σ_1, σ_2). It will be shown in three steps that the equilibria of LFMSE will be of very simple type.

In the first step, we will show that there is no symmetric equilibrium. In the second step, it will be shown that if a strategy profile $\xi = (\xi_1, \xi_2)$ is an equilibrium of LFMSE, then there are $0 < c_1 < c_2 < 1$ such that both players will choose the safe one for $p \in [0, c_1]$, and both will choose the risky one for $p \in (c_2, 1]$. If $p \in (c_1, c_2]$, then there will be two possible cases. Either both will use mixed strategies or one of the players will choose the safe option, and the other chooses the risky one. In the last step, by showing that it is impossible for both players to use mixed strategies at the same time, we conclude that there are only asymmetric equilibria, the structure of which is extremely simple.

2.5.1 Non-Existence of a Symmetric Equilibrium

We will now show that there is no symmetric equilibrium in the LFMSE. We begin with the following lemma, which follows immediately from the results summarized in Section 3.

Lemma 2.11 *A strategy profile (ξ_1, ξ_2) is a subgame perfect Markov equilibrium in LFMSE if and only if for all $p \in [0, 1]$,*

$$\xi_1(p) \in \arg\max_{0 \leq \alpha_1 \leq 1} \{(1 - \alpha_1)s + \alpha_1 m(p) + \frac{1}{r}\left(\frac{\alpha_1}{\sigma_1^2} + \frac{\xi_2(p)}{\sigma_2^2}\right)\Phi(p)\frac{u_1''(p)}{2}\}$$

and

$$\xi_2(p) \in \arg\max_{0 \leq \alpha_2 \leq 1} \{(1 - \alpha_2)s + \alpha_2 m(p) + \frac{1}{r}\left(\frac{\alpha_2}{\sigma_2^2} + \frac{\xi_1(p)}{\sigma_1^2}\right)\Phi(p)\frac{u_2''(p)}{2}\},$$

where u_1 and u_2 are the unique solutions of the following Bellman equations, respectively.

$$u_1(p) = \max_{0 \leq \alpha_1 \leq 1} \{(1-\alpha_1)s + \alpha_1 m(p) + \frac{1}{r}\left(\frac{\alpha_1}{\sigma_1^2} + \frac{\xi_2(p)}{\sigma_2^2}\right)\Phi(p)\frac{u_1''(p)}{2}\} \tag{2.2}$$

and

$$u_2(p) = \max_{0 \leq \alpha_2 \leq 1} \{(1-\alpha_2)s + \alpha_2 m(p) + \frac{1}{r}\left(\frac{\alpha_2}{\sigma_2^2} + \frac{\xi_1(p)}{\sigma_1^2}\right)\Phi(p)\frac{u_2''(p)}{2}\}. \tag{2.3}$$

It is easy to see why there can't be any symmetric equilibrium if $\sigma_1 \neq \sigma_2$. If there were a symmetric equilibrium, since both players would behave in the same way, their value functions should be the same. If, however, the value functions were the same, then the informational gains from experimentation, which is dependent on σ_i, would be different across players. Therefore, the cutoff point at which the informational gain begins to outweigh the common opportunity cost of experimentation would be different. Thus, they couldn't use the same strategies.

Theorem 2.12 *There is no symmetric equilibrium in LFMSE.*

Proof . Suppose a strategy profile (ξ, ξ) is a symmetric equilibrium. Let u_1 be the unique solution of the Bellman equation

$$u_1 = \max_{0 \leq \alpha_1 \leq 1} \{(1-\alpha_1)s + \alpha_1 m(p) + \frac{1}{r}\left(\frac{\alpha_1}{\sigma_1^2} + \frac{\xi}{\sigma_2^2}\right)\Phi(p)\frac{u_1''(p)}{2}\},$$

and u_2 be the unique solution of the Bellman equation

$$u_2 = \max_{0 \le \alpha_2 \le 1} \{(1-\alpha_2)s + \alpha_2 m(p) + \frac{1}{r}\left(\frac{\alpha_2}{\sigma_2^2} + \frac{\xi}{\sigma_1^2}\right)\Phi(p)\frac{u_2''(p)}{2}\}.$$

Since (ξ, ξ) is an equilibrium,

$$u_1 = (1-\xi)s + \xi m(p) + \frac{1}{r}\left(\frac{\xi}{\sigma_1^2} + \frac{\xi}{\sigma_2^2}\right)\Phi(p)\frac{u_1''(p)}{2}, \tag{2.4}$$

and

$$u_2 = (1-\xi)s + \xi m(p) + \frac{1}{r}\left(\frac{\xi}{\sigma_1^2} + \frac{\xi}{\sigma_2^2}\right)\Phi(p)\frac{u_2''(p)}{2}. \tag{2.5}$$

From the value matching conditions including $u_1(1) = u_2(1) = h$ and $u_1(0) = u_2(0) = l$, and from the smooth pasting conditions, the above two differential equations will have the same boundary conditions so that $u_1 = u_2 = u$.

Suppose that there exists an open interval (c_1, c_2) such that $0 < \xi(p) < 1$ for all $p \in (c_1, c_2)$. Then, $c_2 \le b$, because of Proposition 2.6. Therefore, for $p \in (c_1, c_2)$, $s - m(p) > 0$. Since (ξ, ξ) is an equilibrium, $0 < \xi(p) < 1$ for $p \in (c_1, c_2)$ implies that

$$\frac{1}{r\sigma_1^2}\Phi(p)\frac{u''(p)}{2} = s - m(p) = \frac{1}{r\sigma_2^2}\Phi(p)\frac{u''(p)}{2} > 0,$$

which is impossible. Therefore, there is no open interval (c_1, c_2) such that $0 < \xi(p) < 1$ for all $p \in (c_1, c_2)$.

Now let $c_1 = \sup\{c : \xi(p) = 0 \text{ for all } p \in [0, c]\}$. Thus, $u(p) = s$ for all $p \in [0, c_1]$. Note again that by Proposition 2.6, $c_1 < b$. Hence, there exists $c_2 \in (c_1, b)$

such that $\xi(p) = 1$ for all $p \in (c_1, c_2)$. Therefore, for all $p \in (c_1, c_2)$,

$$u(p) = m(p) + \frac{1}{r}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)\Phi(p)\frac{u''(p)}{2}.$$

That is,

$$\frac{1}{r}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)\Phi(p)\frac{u''(p)}{2} = u - m(p). \tag{2.6}$$

Also, since (ξ, ξ) is an equilibrium, for $p \in (c_1, c_2)$,

$$\frac{1}{r}\frac{1}{\sigma_1^2}\Phi(p)\frac{u''(p)}{2} \geq s - m(p) \quad \text{and} \quad \frac{1}{r}\frac{1}{\sigma_2^2}\Phi(p)\frac{u''(p)}{2} \geq s - m(p). \tag{2.7}$$

From (2.6) and (2.7), for $p \in (c_1, c_2)$, we have

$$\begin{aligned} \frac{1}{r}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)\Phi(p)\frac{u''(p)}{2} &= u(p) - m(p) \\ &\geq 2(s - m(p)), \end{aligned}$$

which is a contradiction since $u(p)$ and $m(p)$ are continuous at c_1, $u(c_1) = s$, and $s - m(c_1) > 0$. ■

2.5.2 Simple Pure Strategies and Simple Mixed Strategies

We will show in this section that if there is an equilibrium, its structure is fairly simple. If ξ is an equilibrium, then there are always three intervals of p; on the first interval starting from zero, neither of the two players are experimenting, on the second one adjacent to the first interval, either only one of them is experimenting or both of them are experimenting at the same time, and on the last interval ending at 1, both of them are experimenting. According to the status over the second interval, we will call an equilibrium simply pure or simply mixed.

Definition 2.13 A strategy profile $\xi = (\xi_1, \xi_2)$ is **a simple pure strategy (SPS)** if there exist $0 < c_1, c_2 < 1$ such that $\xi_i = 0$ for $p \in [0, c_i]$, and $\xi_i = 1$ for $p \in (c_i, 1]$. A strategy profile $\xi = (\xi_1, \xi_2)$ is **a simple mixed strategy (SMS)** if there exist $0 < c_1 < c_2 < 1$ such that for $p \in [0, c_1]$, $\xi_1 = \xi_2 = 0$, for $p \in (c_1, c_2)$, $0 < \xi_1, \xi_2 < 1$, and for $p \in [c_2, 1]$, $\xi_1 = \xi_2 = 1$.

Suppose that player 1 is definitely choosing either the safe or the risky option on an open interval $(\underline{c}, \overline{c})$. Player 2's best response on this interval will be the solution to the appropriate optimal stopping problem restricted on this interval. Hence, it will be characterized by a cutoff rule. The following lemma verifies this intuition. It shows that it is impossible for only one player to strictly mix in equilibrium on any open interval.

Lemma 2.14 *If a strategy profile $\xi = (\xi_1, \xi_2)$ is an equilibrium in LFMSE, then there is no open interval $(\underline{c}, \overline{c})$ such that for all $p \in (\underline{c}, \overline{c})$, either $\xi_i(p) = 0$ and $0 < \xi_{-i}(p) < 1$ or $\xi_i(p) = 1$ and $0 < \xi_{-i}(p) < 1$, for $i = 1, 2$.*

Proof . Suppose there exists an open interval $(\underline{c}, \overline{c})$ such that for all $p \in (\underline{c}, \overline{c})$,

$\xi_1(p) = 0$ and $0 < \xi_2(p) < 1$. Then for $p \in (\underline{c}, \overline{c})$,

$$\frac{1}{r\sigma_2^2}\Phi\frac{u_2''}{2} = s - m.$$

Thus, for all $p \in (\underline{c}, \overline{c})$,

$$\begin{aligned} u_2 &= (1-\xi_2)s + \xi_2 m + \frac{1}{r}\frac{\xi_2}{\sigma_2^2}\Phi\frac{u_2''}{2} \\ &= s. \end{aligned}$$

Therefore, $u_2'' = 0$ on this interval. From Proposition 2.6, we know that $\overline{c} \leq b$, and thus for $p \in (\underline{c}, \overline{c})$,

$$\frac{1}{r\sigma_2^2}\Phi(p)\frac{u_2''(p)}{2} = s - m(p) > 0,$$

which is a contradiction. The proof for the case when $\xi_2(p) = 0$ and $0 < \xi_1(p) < 1$ is similar.

Now suppose that there exists an open interval $(\underline{c}, \overline{c})$ such that, for all $p \in (\underline{c}, \overline{c})$, $\xi_1(p) = 1$ and $0 < \xi_2(p) < 1$. Then, again for all $p \in (\underline{c}, \overline{c})$,

$$\frac{1}{r\sigma_2^2}\Phi\frac{u_2''}{2} = s - m,$$

and this implies that

$$\begin{aligned} u_2 &= (1-\xi_2)s + \xi_2 m + \frac{1}{r}\left(\frac{1}{\sigma_1^2} + \frac{\xi_2}{\sigma_2^2}\right)\Phi\frac{u_2''}{2} \\ &= s + (\frac{\sigma_2}{\sigma_1})^2(s-m), \end{aligned}$$

which is a strictly decreasing function. This is a contradiction, since u_2 is increasing by Proposition 2.5. The proof when $\xi_2(p) = 1$ and $0 < \xi_1(p) < 1$ is similar. ■

With the above lemma, we could drastically reduce the set of strategy profiles that could be equilibria.

Theorem 2.15 *If a strategy profile $\xi = (\xi_1, \xi_2)$ is an equilibrium, then it is either an SPS or an SMS.*

Proof . We will first show that there is no equilibrium $\xi = (\xi_1, \xi_2)$ such that

$$\xi_1 = \begin{cases} 0 & \text{if } p \in [0, c_1) \text{ or } p \in (c_2, c_3] \\ 1 & \text{if } p \in (c_1, c_2) \text{ or } p \in (c_3, 1], \end{cases}$$

and

$$\xi_2 = \begin{cases} 0 & \text{if } p \in [0, c_2) \\ 1 & \text{if } p \in (c_2, 1], \end{cases}$$

for $0 < c_1 < c_2 < c_3 < 1$. We will prove this by comparing the number of boundary conditions and the number of parameters in the differential equation system. If ξ is an equilibrium, we will have 10 coefficients and 3 cutoff points to determine. As boundary conditions, we will have six value matching conditions at cutoff points, and five smooth pasting conditions; one at each $p = c_1, c_2$, and c_3 for player 1, and one at each $p = c_2$ and c_3 for player 2. Moreover, we will have three boundary conditions about the second order derivatives of the value functions, which could be derived from an argument similar to the discussion following the proof of Lemma 2.9. In summary, we have 13 parameters whereas we have 14 boundary conditions. It can be shown similarly that, in general, it is impossible for the two players to experiment in turn at equilibrium. Now, let $\underline{c}_1 = \sup\{c : \xi_1 = 0 \text{ for all } p \in [0, c)\}$, $\underline{c}_2 = \sup\{c : \xi_2 = 0 \text{ for all } p \in [0, c)\}$, $\overline{c}_1 = \inf\{c : \xi_1 = 1 \text{ for all } p \in (c, 1]\}$ and $\overline{c}_2 = \inf\{c : \xi_2 = 1 \text{ for all } p \in (c, 1]\}$.

Then, with the result demonstrated at the beginning of this proof, by Theorem 2.12 and Lemma 2.14, we have to consider only the following four possibilities.

If $\underline{c}_1 = \overline{c}_2$, then by Lemma 2.14, $\xi_1 = \xi_2 = 0$ for $0 \leq p < \underline{c}_1$, and $\xi_1 = \xi_2 = 1$ for $\underline{c}_1 < p \leq 1$. This, however, is shown to be impossible in the proof of Theorem 2.12.

If $\overline{c}_2 < \underline{c}_1$, then again by Lemma 2.14, it is clearly an SPS. Also, if $\underline{c}_1 < \overline{c}_2$, and $\underline{c}_1 = \underline{c}_2$, then again by Lemma 2.14, it is an SMS. If $\underline{c}_1 < \underline{c}_2$, it is clearly an SPS. ■

Suppose that player 1 and 2 are strictly mixing on an open interval $(\underline{c}, \overline{c})$. Then, their private benefits from experimentation should be the same, given the common opportunity cost of experimentation. That is,

$$\frac{1}{r\sigma_1^2}\Phi\frac{u_1''}{2} = s - m = \frac{1}{r\sigma_2^2}\Phi\frac{u_2''}{2}.$$

Therefore, $u_1''(p) = (\sigma_1/\sigma_2)^2 u_2''(p)$, which implies that the curvature of u_1 is strictly greater than u_2. Since their values at $\underline{c}$ are the same, it implies that u_1 should be greater than u_2. For $u_1 = u_2 = h$ at $p = 1$, however, if $u_1 > u_2$ for some open interval, then u_1 would be required to have smaller curvature. This is the main reason why we can't have an SMS equilibrium if $\sigma_1 \neq \sigma_2$.

The following lemma, which is a modification of Theorem 9 in Bolton and Harris (1999), will be useful in the proof of non-existence of SMS equilibria. Since its proof is similar to that of Theorem 2.21, it will be omitted.

Lemma 2.16 Let a strategy profile $\xi = (\xi_1, \xi_2)$ be an SMS, let u_i be player i's value function when she plays a best response against ξ_{-i}, and let

$$\beta_1 = \left(\frac{u_2 - s}{s - m}\right) \text{ and } \beta_2 = \left(\frac{u_1 - s}{s - m}\right).$$

If ξ is an SMS equilibrium, then

$$\xi_1 = \begin{cases} (\sigma_1/\sigma_2)^2 \beta_1 & \text{if } \beta_1 \leq (\sigma_2/\sigma_1)^2 \text{ and } p < b \\ 1 & \text{otherwise,} \end{cases}$$

and

$$\xi_2 = \begin{cases} (\sigma_2/\sigma_1)^2 \beta_2 & \text{if } \beta_2 \leq (\sigma_1/\sigma_2)^2 \text{ and } p < b \\ 1 & \text{otherwise.} \end{cases}$$

Theorem 2.17 *If a strategy profile $\xi = (\xi_1, \xi_2)$ is an equilibrium in the LFMSE, then it is an SPS equilibrium.*

Proof . With Theorem 2.15, it suffices to show that there is no SMS equilibrium in the LFMSE. Suppose that $\xi = (\xi_1, \xi_2)$ is an SMS equilibrium in the LFMSE, and that $0 < \xi_1(p), \xi_2(p) < 1$ if and only if $p \in (\underline{c}, \overline{c})$. Note that $\overline{c} \leq b$ by Proposition

2.6.

Now $0 < \xi_1(p), \xi_2(p) < 1$ for $p \in (\underline{c}, \overline{c})$ implies that

$$u_1''(p) = (\frac{\sigma_1}{\sigma_2})^2 u_2''(p) > 0. \tag{2.8}$$

Also $(1/(r\sigma_2^2))\Phi u_2''/2 = s - m$ for $p \in (\underline{c}, \overline{c})$ implies that the value function for player 2, u_2, will be the solution of the following second order differential equation:

$$\begin{aligned} u_2 &= (1-\xi_2)s + \xi_2 m + \frac{1}{r}(\frac{\xi_1}{\sigma_1^2} + \frac{\xi_2}{\sigma_2^2})\Phi\frac{u_2''}{2} \\ &= s + \frac{1}{r}\frac{\xi_1}{\sigma_1^2}\Phi\frac{u_2''}{2}. \end{aligned} \tag{2.9}$$

Therefore, for $p \in [\underline{c}, \overline{c}]$,

$$u_2(p) = s + \beta_1 f(p) + \beta_2 g(p), \tag{2.10}$$

where $f(p)$ and $g(p)$ are two independent solutions for the differential equation

$$u_2(p) = \frac{1}{r}\frac{\xi_1(p)}{\sigma_1^2}\Phi(p)\frac{u_2''(p)}{2}.$$

From (2.8) with the boundary conditions for the value functions including

$$u_1(\underline{c}) = u_2(\underline{c}) = s, \text{ and } u_1'(\underline{c}) = u_2'(\underline{c}) = 0,$$

we obtain

$$u_1(p) = s + (\frac{\sigma_1}{\sigma_2})^2(\beta_1 f(p) + \beta_2 g(p)) \tag{2.11}$$

for $p \in [\underline{c}, \overline{c}]$.

Since $\xi_1(\overline{c}) = \xi_2(\overline{c}) = 1$, from Lemma 2.16,

$$(\frac{\sigma_2}{\sigma_1})^2(\frac{u_1 - s}{s - m}) = (\frac{\sigma_1}{\sigma_2})^2(\frac{u_2 - s}{s - m}) \tag{2.12}$$

at $p = \overline{c}$. From (3.10), (3.11), and (3.12),

$$\beta_1 f(\overline{c}) + \beta_2 g(\overline{c}) = 0.$$

Also, the value matching condition $u_1(\underline{c}) = u_2(\underline{c}) = s$ implies that

$$\beta_1 f(\underline{c}) + \beta_2 g(\underline{c}) = 0,$$

which is a contradiction since u_i is strictly increasing on $(\underline{c}, \overline{c})$ due to both Proposition 2.5 and $u_i'' > 0$ for $p \in (\underline{c}, \overline{c})$. ■

We have proved that there is no SMS equilibrium in the LFMSE. Consequently, the symmetric equilibrium in Bolton and Harris (1999), which is an SMS, disappears as soon as we add heterogeneity into the structure of the noise.

2.5.3 Uniqueness

For the uniqueness of each type of SPS equilibrium, given the differential equation systems for the value functions on each interval, we should have an adequate number of boundary conditions. In this section, we will provide two extra boundary conditions in addition to the usual value matching conditions and the smooth pasting conditions. Uniqueness of each type of equilibrium will follow immediately from this.

Lemma 2.18 *Suppose that*

$$\xi_i(p) = \begin{cases} 0 & \text{if } p \in [0, c_1] \\ 1 & \text{if } p \in (c_1, 1] \end{cases}$$

and

$$\xi_{-i}(p) = \begin{cases} 0 & \text{if } p \in [0, c_2] \\ 1 & \text{if } p \in (c_2, 1] \end{cases}$$

is an equilibrium of LFMSE, where $0 < c_2 < c_1 < 1 = c_0$. *Let* $u_i^j(p)$ *be the player* i*'s value function associated with* $\xi = (\xi_1, \xi_2)$ *on* $(c_{j+1}, c_j]$, $j = 0, 1$. *Then,*

$$(u_i^0)''(c_1) = (u_i^1)''(c_1), \text{ and } (u_{-i}^0)'(c_1) = (u_{-i}^1)'(c_1).$$

Proof . Since the proof for $(u_i^0)''(c_1) = (u_i^1)''(c_1)$ is similar to that of Theorem 2.9, we will prove only $(u_{-i}^0)'(c_1) = (u_{-i}^1)'(c_1)$. Since u_{-i}^j is convex and increasing, it is impossible to have $(u_{-i}^0)'(c_1) < (u_{-i}^1)'(c_1)$. Suppose $(u_{-i}^0)'(c_1) > (u_{-i}^1)'(c_1)$. Then, there exists $\varepsilon > 0$ such that $u_{-i}^0(p) < u_{-i}^1(p)$ for all $p \in (c_1 - \varepsilon, c_1)$. If $\xi_i(p)$ were 1 for $p \in (c_1 - \varepsilon, c_1)$, then $u_{-i}^0(p)$ would still be the value function for player i. Since $\xi_i = 0$ for $p \in (c_1 - \varepsilon, c_1)$, by Proposition 2.7, $u_{-i}^0(p) \geq u_{-i}^1(p)$, which is a contradiction. ■

With the previous lemma, we see that the number of boundary conditions is equal to that of parameters in our differential equation systems. Hence, it should be clear that each type of equilibrium is unique.

Theorem 2.19 *Suppose that*

$$\xi_i(p) = \begin{cases} 0 & \text{if } p \in [0, c_1] \\ 1 & \text{if } p \in (c_1, 1] \end{cases}$$

and

$$\xi_{-i}(p) = \begin{cases} 0 & \text{if } p \in [0, c_2] \\ 1 & \text{if } p \in (c_2, 1] \end{cases},$$

where $0 < c_2 < c_1 < 1 = c_0$*, is an equilibrium of LFMSE. Then, there is no other equilibrium* $\tilde{\xi} = (\tilde{\xi}_1, \tilde{\xi}_2)$ *such that* $\tilde{\xi}_i \leq \tilde{\xi}_{-i}$ *for all* $p \in [0, 1]$.

In the remainder of this paper, we will use a leader and a follower to distinguish each player at SPS equilibria in the following sense.

Definition 2.20 A strategy profile $\xi = (\xi_1, \xi_2)$ is an **SPS with player** i **as a leader and player** j **as a follower**, $i = 1$ or 2, $j = 3 - i$, if ξ is an SPS and if there exist $0 < c_1 < c_2 < 1$ such that $\xi_i = 1$ and $\xi_j = 0$ for $p \in (c_1, c_2]$.

2.6 Robustness of SPS Equilibria

The main result of this section shows that, for a broad range of parameters, the length of the interval of p on which only one player is doing experimentation at an SPS equilibrium will be bounded away from zero. Bolton and Harris (1999) focus only on the symmetric equilibrium with the assumption of homogeneity. That is, with the assumption that $\sigma_1 = \sigma_2$, they do not investigate the properties of an asymmetric equilibrium. The results of this paper, however, show that the unique symmetric equilibrium in Bolton and Harris (1999) is not robust to perturbations in σ_i's when there are two players. Moreover, when σ_2 is close to σ_1, the length of the interval of p on which only one player is doing experimentation will be shown to be bounded away from zero. This suggests the existence of an SPS equilibrium which is asymmetric in the setting of Bolton and Harris (1999). In fact, the existence proof for an SPS equilibrium in Section 2.8 will be valid for the homogeneous case, too. That is, when we have homogeneous noise structures, we have three equilibria; two asymmetric pure strategy equilibria, and one symmetric mixed strategy equilibrium. Only asymmetric equilibria, however, are robust to perturbations. Thus, we argue that, when we have two players, focusing on the symmetric equilibrium as in Bolton and Harris (1999) has little justification, and that the focus of the future research be on the asymmetric equilibria even in the symmetric game.

Following characterization theorem for an SPS equilibrium is crucial for the main theorem.

Theorem 2.21 *Suppose that a strategy profile $\xi = (\xi_1, \xi_2)$ is an SPS with player i as a leader and player j as a follower, $i = 1$ or 2, $j = 3 - i$. Let u_i be player i's* *value function when she plays a best response to ξ_{-i}, and let $\beta_i = (u_i - s)/(s - m)$ for $i = 1,2$. Then, ξ is an SPS equilibrium if and only if*

$$\xi_i = \begin{cases} 0 & \text{if } \beta_i \leq 0 \text{ and } p < b \\ 1 & \text{otherwise,} \end{cases}$$

and

$$\xi_j = \begin{cases} 0 & \text{if } \beta_j \leq (\sigma_j/\sigma_i)^2 \text{ and } p < b \\ 1 & \text{otherwise.} \end{cases}$$

For the proof of Theorem 2.21, we need a technical lemma.

Lemma 2.22 *Suppose that a strategy profile $\xi = (\xi_1, \xi_2)$ is an SPS equilibrium, and that $\xi_1(p) = 0$ if $0 \leq p \leq c_1$, $\xi_1(p) = 1$ if $c_1 < p \leq 1$, $\xi_2(p) = 0$ if $0 \leq p \leq c_2$, and $\xi_1(p) = 1$ if $c_2 < p \leq 1$, where $0 < c_1 \neq c_2 < 1$. Let u_i be the value function of player i when she plays a best response to ξ_{-i}. Then, for all $i = 1, 2$, for all $p > c_i$,*

$$\frac{1}{r\sigma_i^2}\Phi\frac{u_i''}{2} \geq s - m,$$

where equality holds for at most one $\tilde{p}$ such that $\tilde{p} \in (\max\{c_1, c_2\}, b]$.

Proof . We will prove this for the case $c_1 < c_2$. The proof for the case $c_1 > c_2$ is similar. Since $s - m(p) < 0$ for $p > b$, it is clear that if there exists a p such that the above inequality holds as an equality, it should be less than or equal to b.

That $\xi_1(p) = 1$ for $c_1 < p \leq 1$ implies

$$\frac{1}{r\sigma_1^2}\Phi(p)\frac{u_1''(p)}{2} \geq s - m(p)$$

for $p \in (c_1, 1]$. Now suppose that

$$\frac{1}{r\sigma_1^2}\Phi(\hat{p})\frac{u_1''(\hat{p})}{2} = s - m(\hat{p})$$

for some $\hat{p} \in (c_1, c_2)$. Then, since $\xi_2(p) = 0$ for $p \in (c_1, c_2)$,

$$\begin{aligned} u_1(\hat{p}) &= m(\hat{p}) + \frac{1}{r\sigma_1^2}\Phi(\hat{p})\frac{u_1''(\hat{p})}{2} \\ &= s. \end{aligned}$$

As u_1 is increasing by Proposition 2.5, $u_1(\hat{p}) = s$ implies that $u_1 = s$ for all $p \leq \hat{p}$. Then, $u_1'' = 0$ for all $p \leq \hat{p}$, which is a contradiction to the assumption that

$$\frac{1}{r\sigma_1^2}\Phi(\hat{p})\frac{u_1''(\hat{p})}{2} = s - m(\hat{p}) > 0.$$

Suppose

$$\frac{1}{r\sigma_1^2}\Phi(\tilde{p})\frac{u_1''(\tilde{p})}{2} = s - m(\tilde{p})$$

for some $\tilde{p} \in (c_2, b]$. Then, we will show that there exists at most one such $\tilde{p}$ in $(c_2, b]$.

From $\xi_2(p) = 1$ for $p \in (c_2, b]$,

$$\begin{aligned} u_1(\tilde{p}) &= m(\tilde{p}) + \frac{1}{r}(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2})\Phi(\tilde{p})\frac{u_1''(\tilde{p})}{2} \\ &= s + (\frac{\sigma_1}{\sigma_2})^2(s - m(\tilde{p})). \end{aligned}$$

As $s + (\sigma_1/\sigma_2)^2(s - m(p))$ is a strictly decreasing function of p, that u_1 is increasing clearly implies that there exists at most one such $\tilde{p}$. The proof for player 2 is similar. ■

Proof of the Main Theorem . We will prove this for the case when player 1 is a leader and player 2 is a follower. The proof for the case when player 2 is a leader and player 1 is a follower is similar.

$\Longrightarrow$

A strategy profile $\xi = (\xi_1, \xi_2)$ is an equilibrium if and only if

$$\xi_1 \in \arg\max_{0 \le \alpha_1 \le 1} \{(1-\alpha_1)s + \alpha_1 m(p) + \frac{1}{r}\left(\frac{\alpha_1}{\sigma_1^2} + \frac{\xi_2(p)}{\sigma_2^2}\right)\Phi(p)\frac{u_1''(p)}{2}\}$$

and

$$\xi_2 \in \arg\max_{0 \le \alpha_2 \le 1} \{(1-\alpha_2)s + \alpha_2 m(p) + \frac{1}{r}\left(\frac{\alpha_2}{\sigma_2^2} + \frac{\xi_1(p)}{\sigma_1^2}\right)\Phi(p)\frac{u_2''(p)}{2}\},$$

where u_1 and u_2 are the unique solutions of the following Bellman equations, respectively:

$$u_1 = \max_{0 \le \alpha_1 \le 1} \{(1-\alpha_1)s + \alpha_1 m(p) + \frac{1}{r}\left(\frac{\alpha_1}{\sigma_1^2} + \frac{\xi_2(p)}{\sigma_2^2}\right)\Phi(p)\frac{u_1''(p)}{2}\} \tag{2.13}$$

and

$$u_2 = \max_{0 \le \alpha_2 \le 1} \{(1-\alpha_2)s + \alpha_2 m(p) + \frac{1}{r}\left(\frac{\alpha_2}{\sigma_2^2} + \frac{\xi_1(p)}{\sigma_1^2}\right)\Phi(p)\frac{u_2''(p)}{2}\}. \tag{2.14}$$

Let

$$v_i(p) = \frac{1}{r\sigma_i^2}\Phi(p)\frac{u_i''(p)}{2}$$

for $i = 1, 2$. Note that $v_i \geq 0$ by Proposition 2.5.

Suppose that $\xi_1 \neq \xi_2$ if and only if $p \in (c_1, c_2]$.

Suppose $p < b$. Then, there are three possible cases.

If $0 \leq p \leq c_1$, then $\xi_1 = \xi_2 = 0$. Hence, from (2.13) and (2.14), $u_1 = u_2 = s$. $\therefore \beta_1 = \beta_2 = 0$.

If $c_1 < p \leq c_2$, then $\xi_1 = 1$, $\xi_2 = 0$. That $\xi_2 = 0$ implies that $v_2 \leq s - m$. Note that $v_1 > s - m$ from Lemma 2.22. Now, from (2.13), $u_1 = m + v_1$. Thus,

$$\beta_1 = \frac{u_1 - s}{s - m} = -1 + \frac{v_1}{s - m} > 0.$$

Also from (2.14),

$$u_2 = s + (\sigma_2/\sigma_1)^2 v_2.$$

Hence,

$$\begin{aligned} \beta_2 &= \frac{u_2 - s}{s - m} \\ &= (\frac{\sigma_2}{\sigma_1})^2 \frac{v_2}{s - m} \leq (\frac{\sigma_2}{\sigma_1})^2. \end{aligned}$$

If $c_2 < p < b$, then $\xi_1 = \xi_2 = 1$. That $\xi_1 = \xi_2 = 1$ implies that $v_1, v_2 \geq s - m$. Therefore, from (2.13),

$$\begin{aligned} u_1 - s &= m + v_1 + \frac{1}{r\sigma_2^2} \Phi \frac{u_1''}{2} - s \\ &\geq \frac{1}{r\sigma_2^2} \Phi \frac{u_1''}{2}. \end{aligned} \tag{2.15}$$

By Lemma 2.22, strict inequality holds in (2.15) for all but at most one $p \in (c_2, b)$. Since $u_1 - s$ is increasing and $u_1'' \geq 0$, then it is clearly true that $u_1 - s > 0$ for all

$p \in (c_2, b)$. Hence, $\beta_1 > 0$.

From (2.14),

$$u_2 = m + (1/(r\sigma_1^2))\Phi u_2''/2 + v_2.$$

Therefore,

$$\begin{aligned} \beta_2 &= \frac{u_2 - s}{s - m} \\ &= -1 + \left(1 + (\frac{\sigma_2}{\sigma_1})^2\right)(\frac{v_2}{s - m}) \\ &\geq -1 + \left(1 + (\frac{\sigma_2}{\sigma_1})^2\right) = (\frac{\sigma_2}{\sigma_1})^2. \end{aligned} \tag{2.16}$$

By Lemma 2.22 again, in (2.16), strict inequality holds for all but at most one $p \in (c_2, b)$. It is easy to see that $(u_2 - s)/(s - m)$ is strictly decreasing. Hence, it is clearly true that $\beta_2 > (\sigma_2/\sigma_1)^2$ for all $p \in (c_2, b)$.

Suppose $p \geq b$. Then, $\xi_1 = \xi_2 = 1$ by Proposition 2.6, and hence there remains nothing to show.

$\Longleftarrow$

Suppose $p < b$. There are three possible cases.

If $v_1, v_2 \leq s - m$, then we will show that $(\xi_1, \xi_2) = (0, 0)$. From (2.14),

$$u_2 = s + \xi_1(\sigma_2/\sigma_1)^2 v_2.$$

Hence,

$$(\frac{u_2 - s}{s - m}) = (\frac{\sigma_2}{\sigma_1})^2 \xi_1 (\frac{v_2}{s - m})$$

$$
\begin{aligned}
&\leq (\frac{\sigma_2}{\sigma_1})^2 \xi_1 \\
&\leq (\frac{\sigma_2}{\sigma_1})^2.
\end{aligned}
$$

Therefore, $\xi_2 = 0$. With this result, from (2.13),

$$u_1 = s + \xi_2 (\sigma_1/\sigma_2)^2 v_1 = s.$$

That is, $\xi_1 = 0$.

If $v_1 > s - m \geq v_2$, then we will show that $(\xi_1, \xi_2) = (1, 0)$. From (2.14),

$$u_2 = s + \xi_1 (\sigma_2/\sigma_1)^2 v_2.$$

Hence, as in the previous case, $\xi_2 = 0$. With this result, from (2.13),

$$
\begin{aligned}
u_1 &= m + v_1 + \xi_2 (\frac{\sigma_1}{\sigma_2})^2 v_1 \\
&= m + v_1.
\end{aligned}
$$

$\therefore$

$$\frac{u_1 - s}{s - m} = -1 + \frac{v_1}{s - m} > 0.$$

Hence, $\xi_1 = 1$.

If $v_1, v_2 > s - m$, then we will show that $(\xi_1, \xi_2) = (1, 1)$. From (2.13),

$$u_1 = m + v_1 + \xi_2 (\sigma_1/\sigma_2)^2 v_1.$$

Thus,

$$\frac{u_1 - s}{s - m} = -1 + \left((\frac{\sigma_1}{\sigma_2})^2 \xi_2 + 1 \right) (\frac{v_1}{s - m})$$

$$> \quad (\frac{\sigma_1}{\sigma_2})^2 \xi_2 \geq 0.$$

$\therefore \xi_1 = 1$. With this result, from (2.14),

$$u_2 = m + v_2 + \xi_1 (\sigma_2/\sigma_1)^2 v_2.$$

Hence,

$$\begin{aligned} \frac{u_2 - s}{s - m} &= -1 + \left((\frac{\sigma_2}{\sigma_1})^2 \xi_1 + 1 \right) (\frac{v_2}{s-m}) \\ &> (\frac{\sigma_2}{\sigma_1})^2 \xi_1 = (\frac{\sigma_2}{\sigma_1})^2. \end{aligned}$$

Therefore, $\xi_2 = 1$.

If $p \geq b$, then again there remains nothing to show due to Proposition 2.6. ■

Given an SPS $\xi = (\xi_1, \xi_2)$, suppose that $\xi_1(p) \neq \xi_2(p)$ if and only if $p \in (\underline{c}, \overline{c}]$. Then, let $\eta : \mathcal{S}^2 \to R$ be a function defined as

$$\eta(\xi) = \overline{c} - \underline{c},$$

where $\mathcal{S}$ is the set of SPS. Hence, η is a function which maps each SPS to the length of the interval on which only one player is experimenting.

Theorem 2.23 *Given LFMSE with parameters* $(\sigma_1, \sigma_2), 0 < \sigma_2 < \sigma_1$, *let*

$$l(\sigma_1, \sigma_2) = \inf_{\xi} \eta(\xi),$$

where the inf *is taken over the set of* ξ*'s such that* ξ *is an SPS equilibrium in*

LFMSE with parameters (σ_1, σ_2)*. If* $\sigma_1 < (h-l)/r^{1/2}$*, then*

$$\lim_{\sigma_2 \to \sigma_1} \inf\, l(\sigma_1, \sigma_2) > 0.$$

Proof . Suppose that a strategy profile $\xi = (\xi_1, \xi_2)$ is an SPS equilibrium in the LFMSE with parameters (σ_1, σ_2), and that $\xi_1(p) \neq \xi_2(p)$ if and only if $p \in (\underline{c}, \overline{c}]$. Let u_i be the value function of player i when she plays a best response to ξ_{-i}. Let c_1^* be the cutoff point for player 1 when player 2 is using $\xi_2 = 0$, and let c_2^* be the cutoff point for player 2 when player 1 is using $\xi_1 = 0$. That is, against $\xi_1 = 0$, it is optimal for player 2 to choose the safe option when $p \in [0, c_2^*]$, and the risky one when $p \in (c_2^*, 1]$. Also, against $\xi_2 = 0$, it is optimal for player 1 to choose the safe option when $p \in [0, c_1^*]$, and the risky one when $p \in (c_1^*, 1]$. Finally let u_1^* and u_2^* be the value functions of player 1 and 2 in this case, respectively. Then, by Proposition 2.7, $u_1^* \leq u_1$, and $u_2^* \leq u_2$. Suppose that $\xi = (\xi_1, \xi_2)$ is an SPS equilibrium with player 2 as a leader and player 1 as a follower. Then, $u_2^* \leq u_2$ implies that $\underline{c} \leq c_2^*$. Note that c_2^* is determined by the value matching condition and smooth pasting condition as in the analysis of the Team Problem, and the result is

$$c_2^* = \frac{(\zeta - 1)(s - l)}{(\zeta - 1)(s - l) + (\zeta + 1)(h - s)},$$

where

$$\zeta = \left(1 + \frac{8r\sigma_2^2}{(h-l)^2}\right)^{1/2}.$$

The following argument is based on simple geometry. Therefore, drawing a diagram while reading the proof will be helpful.

By Theorem 2.21, at $p = \overline{c}$

$$u_1 = s + (\frac{\sigma_1}{\sigma_2})^2(s - m).$$

Since u_1 is an increasing convex function and $u_1 \leq \overline{u} = ph + (1-p)s$, $\overline{c}$ can't be less than α such that

$$\overline{u}(\alpha) = s + (\sigma_1/\sigma_2)^2(s - m(\alpha)).$$

Let α' be the number such that $\overline{u}(\alpha') = s + s - m(\alpha')$. That is

$$\alpha' = \frac{s-l}{2h-s-l}.$$

Then, since $\overline{u}$ is strictly increasing, and

$$s + s - m(\alpha) < s + (\sigma_1/\sigma_2)^2(s - m(\alpha)),$$

it is clear that $\alpha' < \alpha$. Hence,

$$\begin{aligned}
\eta(\xi) &= \overline{c} - \underline{c} \\
&\geq \alpha' - c_2^* \\
&= \left[\frac{(s-l)(h-s)}{(h-s)+(h-l)}\right]\left[\frac{3-\zeta}{(\zeta-1)(s-l)+(\zeta+1)(h-s)}\right].
\end{aligned}$$

Since $\zeta > 1$, the last expression will be positive if $\zeta < 3$ which is equivalent to

$\sigma_2 < (h-l)/r^{1/2}$. Therefore, if $\sigma_1 < (h-l)/r^{1/2}$, then

$$\zeta < \left(1 + \frac{8r\sigma_1^2}{(h-l)^2}\right)^{1/2} < 3.$$

Let $\zeta^* = (1 + 8r\sigma_1^2/(h-l)^2)^{1/2}$. Then, since

$$\frac{3-\zeta}{(\zeta-1)(s-l)+(\zeta+1)(h-s)}$$

is strictly decreasing for $\zeta \in [1,3]$,

$$\begin{aligned} \eta(\xi) &\geq \left[\frac{(s-l)(h-s)}{(h-s)+(h-l)}\right]\left[\frac{(3-\zeta)}{(\zeta-1)(s-l)+(\zeta+1)(h-s)}\right] \\ &> \left[\frac{(s-l)(h-s)}{(h-s)+(h-l)}\right]\left[\frac{(3-\zeta^*)}{(\zeta^*-1)(s-l)+(\zeta^*+1)(h-s)}\right]. \end{aligned} \tag{2.17}$$

Now suppose that $\xi = (\xi_1, \xi_2)$ is an SPS equilibrium with player 1 as a leader and player 2 as a follower. Again $u_1^* \leq u_1$ implies that $\underline{c} \leq c_1^*$. Note that c_1^* will be determined in the same way as c_2^*, and the result is

$$c_1^* = \frac{(\zeta^*-1)(s-l)}{(\zeta^*-1)(s-l)+(\zeta^*+1)(h-s)}.$$

Now by Theorem 2.21, at $p = \overline{c}$

$$u_2 = s + (\frac{\sigma_2}{\sigma_1})^2(s-m).$$

The basic idea of the proof is almost the same as in the previous case. We will show that there is $\sigma_2^* < \sigma_1$ such that for all $\sigma_2 > \sigma_2^*$, the distance between c_1^* and the intersection of $s + (\sigma_2/\sigma_1)^2(s-m)$ and $\overline{u}$ is bounded away from zero.

Since u_2 is an increasing convex function and $u_2 \leq \overline{u} = ph + (1-p)s$, $\overline{c}$ should be greater than or equal to β such that

$$\overline{u}(\beta) = s + (\sigma_2/\sigma_1)^2(s - m(\beta)).$$

It is easy to see that

$$\beta = \frac{s-l}{(\sigma_1/\sigma_2)^2(h-s)+(h-l)}.$$

Hence,

$$\begin{aligned}
\eta(\xi) &\geq \beta - c_1^* \\
&= \left[\frac{(s-l)(h-s)}{(\zeta^*+1)(h-s)+(\zeta^*-1)(s-l)}\right]\left[\frac{2-(\sigma_1/\sigma_2)^2(\zeta^*-1)}{(\sigma_1/\sigma_2)^2(h-s)+(h-l)}\right].
\end{aligned}$$

If $\sigma_1 < (h-l)/r^{1/2}$, then $\zeta^* < 3$, and hence, there exists $\sigma_2^* < \sigma_1$ such that

$$2 - (\frac{\sigma_1}{\sigma_2})^2(\zeta^*-1) > 2 - (\frac{\sigma_1}{\sigma_2^*})^2(\zeta^*-1) > 0$$

for all $\sigma_2^* < \sigma_2 < \sigma_1$.

Consequently,

$$\begin{aligned}
&\eta(\xi) \qquad (2.18)\\
\geq\ &\left[\frac{(s-l)(h-s)}{(\zeta^*+1)(h-s)+(\zeta^*-1)(s-l)}\right]\left[\frac{2-(\sigma_1/\sigma_2)^2(\zeta^*-1)}{(\sigma_1/\sigma_2)^2(h-s)+(h-l)}\right] \\
>\ &\left[\frac{(s-l)(h-s)}{(\zeta^*+1)(h-s)+(\zeta^*-1)(s-l)}\right]\left[\frac{2-(\sigma_1/\sigma_2^*)^2(\zeta^*-1)}{(\sigma_1/\sigma_2^*)^2(h-s)+(h-l)}\right]
\end{aligned}$$

for all σ_2 such that $\sigma_2^* < \sigma_2 < \sigma_1$.

From (2.17) and (2.18), it is clearly true that

$$\lim \inf_{\sigma_2 \to \sigma_1} l(\sigma_1, \sigma_2) > 0.$$

■

2.7 The Value Functions at SPS Equilibria

Suppose that a strategy profile $\xi = (\xi_1, \xi_2)$ is an SPS equilibrium with player 1 as a leader and player 2 as a follower. Then, since player 2 is experimenting less and free riding on player 1's experimentation for some range of beliefs, we could conjecture that the payoff of player 2 is greater than that of player 1. We will prove in this section that this conjecture is indeed true. With this result, LFMSE can be understood as a simple coordination game. There are only two kinds of equilibria, and who will have a better payoff is determined according to which type of equilibria they are playing. That is, the player who is allowed to free ride will have a higher payoff.

We need the following lemma, which is interesting in itself, to prove the main result in this section.

Lemma 2.24 *Suppose that a strategy profile $\xi = (\xi_1, \xi_2)$ is an SPS equilibrium, and that $\xi_1 \neq \xi_2$ if and only if $p \in (c_1, c_2]$. Let u_i be the value function of player i when she plays a best response to ξ_{-i}, $i = 1, 2$. Then,*

$$\begin{aligned} & u_2''(p)[u_1(p) - \{(1-\xi_1(p))s + \xi_1(p)m(p)\}] \\ = \ & u_1''(p)[u_2(p) - \{(1-\xi_2(p))s + \xi_2(p)m(p)\}] \end{aligned}$$

for all $p \in [0, 1] \backslash \{c_1, c_2\}$.

Proof . We will prove for the case when player 1 is a leader. The proof for the case when player 2 is a leader is similar. Suppose that $\xi = (\xi_1, \xi_2)$ is an SPS equilibrium with player 1 as a leader and player 2 as a follower, and that $\xi_1(p) \neq \xi_2(p)$ if and only if $p \in (c_1, c_2]$. Then,

$$\xi_1 = \begin{cases} 0 & \text{if } \ 0 \leq p \leq c_1 \\ 1 & \text{if } \ c_1 < p \leq 1, \end{cases}$$

and

$$\xi_2 = \begin{cases} 0 & \text{if } \ 0 \leq p \leq c_2 \\ 1 & \text{if } \ c_2 < p \leq 1. \end{cases}$$

Since u_i is the value function for player i,

$$u_1 = \begin{cases} s & \text{if } \ 0 \leq p \leq c_1 \\ m + (1/(r\sigma_1^2))\Phi u_1''/2 & \text{if } \ c_1 < p \leq c_2 \\ m + (1/r)(1/\sigma_1^2 + 1/\sigma_2^2)\Phi u_1''/2 & \text{if } \ c_2 < p \leq 1, \end{cases}$$

and

$$u_2 = \begin{cases} s & \text{if } \ 0 \leq p \leq c_1 \\ s + (1/(r\sigma_1^2))\Phi u_2''/2 & \text{if } \ c_1 < p \leq c_2 \\ m + (1/r)(1/\sigma_1^2 + 1/\sigma_2^2)\Phi u_2''/2 & \text{if } \ c_2 < p \leq 1. \end{cases}$$

From this, the statement is clearly true. ■

Since $u_i'' > 0$ for $p \in (c_1, 1]\backslash\{c_2\}$, we obtain

$$\begin{aligned} & \frac{u_1(p) - \{(1-\xi_1(p))s + \xi_1(p)m(p)\}}{u_1''(p)} \\ = \ & \frac{u_2(p) - \{(1-\xi_2(p))s + \xi_2(p)m(p)\}}{u_2''(p)} \end{aligned}$$

for $p \in (c_1, 1]\backslash\{c_2\}$. The numerators are the differences between overall optimal payoff and the instant expected payoff. That is, it measures the gain from additional information generated by experimenting at equilibrium. The denominators are private shadow prices of the information for each player. Hence, at equilibrium, when information has a positive shadow price for each player, the ratio of the gain from information to the private shadow price of the information is equalized across the players.

The following is the main result of this section.

Theorem 2.25 *Suppose that a strategy profile $\xi = (\xi_1, \xi_2)$ is an SPS equilibrium and that $\xi_1 \neq \xi_2$ if and only if $p \in (c_1, c_2]$. Let u_i be the value function of player i when she plays a best response against ξ_{-i}, $i = 1, 2$. Then, $u_i \leq u_{-i}$ if and only if ξ is an SPS equilibrium with player i as a leader and player $3 - i$ as a follower, $i = 1, 2$. The inequality is strict for $p \in (c_1, 1)$.*

Proof . We will prove the theorem for the case when player 2 is a leader. The proof for the case when player 1 is a leader is similar. Suppose that $\xi = (\xi_1, \xi_2)$ is an SPS equilibrium with player 2 as a leader and player 1 as a follower. Note that $c_2 \leq b$ by Proposition 2.6.

We will first show that $u_1 > u_2$ locally to the right of c_1. And then, by showing that it is impossible for $u_1(p) = u_2(p)$ for some p such that $c_1 < p < c_2$, we will derive the conclusion that $u_1 > u_2$ for every $p \in (c_1, 1)$.

From

$$u_1 = s + \frac{1}{r\sigma_2^2}\Phi\frac{u_1''}{2},$$

and

$$u_2 = m + \frac{1}{r\sigma_2^2}\Phi\frac{u_2''}{2}$$

for $p \in (c_1, c_2)$, we obtain

$$u_1 = s + \beta_1 f(p) + \beta_2 g(p),$$

and

$$u_2 = m + \hat{\beta}_1 f(p) + \hat{\beta}_2 g(p),$$

where

$$\begin{aligned} f(p) &= p^{-(\zeta-1)/2}(1-p)^{(\zeta+1)/2}, \\ g(p) &= p^{(\zeta+1)/2}(1-p)^{-(\zeta-1)/2}, \end{aligned}$$

and

$$\zeta = \left(1 + \frac{8r\sigma_2^2}{(h-l)^2}\right)^{1/2}.$$

In the above, $\beta_1, \beta_2, \hat{\beta}_1$, and $\hat{\beta}_2$ are parameters that will be determined by the boundary conditions.

Note that $f, g > 0$ for $p \in (c_1, c_2]$. Therefore, $u_1(c_1) = s$ implies that β_1 and β_2 have opposite signs. Since f is increasing, and g is decreasing, for u_1 to be increasing, it should be that

$$\beta_1 < 0, \quad \text{and} \quad \beta_2 > 0.$$

From this, it follows that the right derivative of u_1 for $p = c_1$ is strictly positive. The smooth pasting condition for player 2, however, implies that $u_2' = 0$ for $p = c_1$.

Consequently, $u_1 > u_2$ locally to the right of c_1.

Now we will show that $u_1 > u_2$ for all $p \in (c_1, c_2)$.

Suppose there exists a $p \in (c_1, c_2)$ such that $u_1(p) = u_2(p)$. Let $\tilde{p}$ be the infimum of those p's. That is, $u_1 > u_2$ for $p \in (c_1, \tilde{p})$, and $u_1 = u_2$ for $p = \tilde{p}$. From Lemma 2.24, for $p = \tilde{p}$,

$$u_1''(\tilde{p})[u_2(\tilde{p}) - m(\tilde{p})] = u_2''(\tilde{p})[u_1(\tilde{p}) - s].$$

Since $s > m(\tilde{p})$, we have

$$u_1''(\tilde{p}) < u_2''(\tilde{p}). \tag{2.19}$$

Since $u_1 > u_2$ for $p \in (c_1, \tilde{p})$, and $u_1 = u_2$ for $p = \tilde{p}$, and since the value functions are convex, (2.19) implies that

$$u_1'(\tilde{p}) < u_2'(\tilde{p}).$$

Therefore, there exists $c \leq c_2$ such that for $p \in (\tilde{p}, c)$,

$$\begin{aligned} u_2 &= m + \frac{1}{r\sigma_2^2}\Phi\frac{u_2''}{2} \\ &> u_1 = s + \frac{1}{r\sigma_2^2}\Phi\frac{u_1''}{2}, \end{aligned}$$

which in turn implies from Lemma 2.24 that $u_2'' > u_1''$ for $p \in (\tilde{p}, c)$. That is, $u_2 - u_1$ is strictly convex on $(\tilde{p}, c)$, $u_2 - u_1 > 0$ for $p \in (\tilde{p}, c)$, and $u_2 - u_1 = 0$ for $p = \tilde{p}$. Hence, it is easy to see that $u_2 > u_1$ for all $p \in (c_1, c_2]$.

Now, for $p \in (c_2, 1]$,

$$u_1 = m + \frac{1}{r}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)\Phi\frac{u_1''}{2},$$

and

$$u_2 = m + \frac{1}{r}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)\Phi\frac{u_2''}{2}.$$

Thus, for $p \in (c_2, 1]$,

$$u_1 = m + \gamma_1 \tilde{f}(p),$$

and

$$u_2 = m + \gamma_2 \tilde{f}(p),$$

where

$$\tilde{f}(p) = p^{-(\zeta_0 - 1)/2}(1-p)^{(\zeta_0+1)/2},$$

and

$$\zeta_0 = \left(1 + \frac{8r}{(h-l)^2(1/\sigma_1^2 + 1/\sigma_2^2)}\right)^{1/2}.$$

Here γ_i's are positive parameters. Then, that $u_2(c_2) > u_1(c_2)$ implies that $\gamma_2 > \gamma_1$, which again implies that

$$u_2'' - u_1'' = (\gamma_2 - \gamma_1)\tilde{f}''(p) > 0$$

for all $p \in (c_2, 1]$. Recall that by Theorem 2.18 and the smooth pasting condition, the right and left derivative of u_1 and u_2 at c_2 should be equal to each other. Hence, overall, $u_2 - u_1$ is increasing, strictly convex, and differentiable for $p \in (\tilde{p}, 1)$, and $u_2 - u_1 = 0$ at $p = \tilde{p}$, which contradicts the fact that $u_2 - u_1 = 0$ at $p = 1$.

We've shown that $u_1 > u_2$ for $p \in (c_1, c_2)$. Now we will show that it is impossible for $u_1(c_1) = u_2(c_2)$, which in turn implies that $u_1 > u_2$ for all $p \in (c_1, 1)$, since if $u_1(c_1) > u_2(c_2)$, then $\gamma_1 > \gamma_2$ so that $u_1 = m + \gamma_1 \tilde{f} > u_2 = m + \gamma_2 \tilde{f}$.

Suppose $u_1(c_1) = u_2(c_2)$. Then, $\gamma_1 = \gamma_2$, and hence $u_1 = u_2$ for all $p \geq c_2$. At $p = c_2$, since $u_1 > u_2$ for $p \in (c_1, c_2)$, $u_1 = u_2$ for $p \in (c_2, 1]$, and since p follows a Brownian motion so that there is a positive probability that p will fall below c_2, it is impossible that $u_1(c_1) = u_2(c_2)$. Contradiction. ■

In the proof of Theorem 2.25, we showed that if $\xi = (\xi_1, \xi_2)$ is an SPS equilibrium with player 2 as a leader and player 1 as a follower, and if $\xi_1 \neq \xi_2$ if and only if $p \in (c_1, c_2]$, then the right derivative of the value function of player 1, u_1, at $p = c_1$ is strictly positive. Therefore, $u_1'(c_1) + u_2'(c_1) > 0$. In the analysis of the Team Problem, however, we showed that $u_*'(\hat{c}) = 0$ when $\hat{c}$ is the cutoff point for player 2's experimentation. Since an SPS equilibrium with player 1 as a leader and player 2 as a follower is obviously inefficient, overall, every equilibrium in LFMSE is not efficient. As $u_1 + u_2 \leq 2u_*$ at both equilibria, we will have less experimentation than optimal in the neighborhood of c_1 at which the leader begins to experiment.

2.8 Existence of an SPS Equilibrium

The existence of both kinds of SPS equilibria can be easily shown by a simple application of Knaster-Tarski's fixed point theorem. Knaster-Tarski's fixed point theorem is for an increasing function defined on a partially ordered set. Suppose that $f : \mathcal{W} \to \mathcal{W}$ is a non decreasing function, where $\mathcal{W}$ is a partially ordered set. That is, if $w_1 \leq w_2$, $w_1, w_2 \in \mathcal{W}$, then $f(w_1) \leq f(w_2)$. Suppose also that there exists $\tilde{w} \in \mathcal{W}$ such that $\tilde{w} \leq f(\tilde{w})$, and that every linearly ordered chain in $\mathcal{W}$ has a supremum in $\mathcal{W}$. Then, f has a fixed point in $\mathcal{W}$.[4]

Theorem 2.26 *There exist an SPS equilibrium with player 1 as a leader and player 2 as a follower and an SPS equilibrium with player 2 as a leader and player 1 as a follower in LFMSE.*

Proof . Let $\mathcal{U}$ be the set of Lipschitz continuous functions $u : [0,1] \to [l,h]$ such that $0 \leq u' \leq h - l$ almost everywhere on $[0,1]$, and let $\mathcal{S}$ be the set of simple strategies $\xi : [0,1] \to [0,1]$ with $\xi^0, \xi^1 : [0,1] \to [0,1]$, where $\xi^0 = 0$ and $\xi^1 = 1$ for all $p \in [0,1]$. Here $\mathcal{U}$ and $\mathcal{S}$ may be interpreted, respectively, as the space of the value functions and the space of the SPS. Let's define three functions,

[4] For proof, see Dugundji and Granas (1982).

$\nu_L, \nu_{1F}, \nu_{2F} : \mathcal{U} \to \mathcal{S}$ as follows:

$$\nu_L(u)(p) = \begin{cases} 0 & \text{if } \ u(p) \leq s \ \text{ and } p < b \\ 1 & \text{otherwise,} \end{cases}$$

$$\nu_{1F}(u)(p) = \begin{cases} 0 & \text{if } \ (u(p) - s)/(s - m(p)) \leq (\sigma_1/\sigma_2)^2 \ \text{ and } p < b \\ 1 & \text{otherwise,} \end{cases}$$

and

$$\nu_{2F}(u)(p) = \begin{cases} 0 & \text{if } \ (u(p) - s)/(s - m(p)) \leq (\sigma_2/\sigma_1)^2 \ \text{ and } p < b \\ 1 & \text{otherwise.} \end{cases}$$

Let $\psi_1 : \mathcal{U}^2 \to \mathcal{S}^2$ be the function defined as

$$\psi_1(u_1, u_2) = \begin{cases} (\nu_L(u_1), \nu_{2F}(u_2)) & \text{if } \ \nu_L(u_1) \geq \nu_{2F}(u_2) \\ (\eta_L, \eta_F) & \text{otherwise,} \end{cases}$$

where $\eta_L, \eta_F \in \mathcal{S}$ are defined as

$$\eta_L(p) = \begin{cases} 0 & \text{if } p \in [0, (1 + b)/2] \\ 1 & \text{if } \ p \in ((1 + b)/2, 1], \end{cases}$$

and

$$\eta_F(p) = \begin{cases} 0 & \text{if } p \in [0, (2 + b)/3] \\ 1 & \text{if } \ p \in ((2 + b)/3, 1]. \end{cases}$$

Also, let $\psi_2 : \mathcal{U}^2 \to \mathcal{S}^2$ be the function defined as

$$\psi_2(u_1, u_2) = \begin{cases} (\nu_{1F}(u_1), \nu_L(u_2)) & \text{if } \ \nu_L(u_2) \geq \nu_{1F}(u_1) \\ (\eta_F, \eta_L) & \text{otherwise.} \end{cases}$$

Then, the mapping $\psi : \mathcal{U}^2 \to \mathcal{S}^4$ defined as $\psi(u_1, u_2) = (\psi_1(u_1, u_2), \psi_2(u_1, u_2))$ will map each pair of value functions into a pair of strategy profiles, the first of which is an SPS with player 1 as a leader and player 2 as a follower, and the second of which is an SPS with player 2 as a leader and player 1 as a follower. Note that, by construction, ψ is increasing.

Let $\lambda : \mathcal{S}^2 \to \mathcal{U}^2$ be defined as $\lambda(\xi_1, \xi_2) = (u_1, u_2)$, where u_i is the value function of player i when she plays a best response to ξ_{-i}. And define $\Lambda : \mathcal{S}^4 \to \mathcal{U}^4$ as $\Lambda(\xi_1, \xi_2, \xi_3, \xi_4) = (\lambda(\xi_1, \xi_2), \lambda(\xi_3, \xi_4))$. Note again that Λ is increasing by Lemma 2.7.

Let $\rho_1, \rho_2 : \mathcal{U}^4 \to \mathcal{U}^2$ be the projection mappings defined as follows.

$$\rho_1(u_1, u_2, u_3, u_4) = (u_1, u_2),$$

and

$$\rho_2(u_1, u_2, u_3, u_4) = (u_3, u_4).$$

Now we are ready to define our main mapping $h_1, h_2 : \mathcal{U}^2 \to \mathcal{U}^2$ which will be shown to have a fixed point. Let's define h_1 and h_2 as

$$h_1 = \rho_1 \circ \Lambda \circ \psi,$$

and

$$h_2 = \rho_2 \circ \Lambda \circ \psi.$$

From Theorem 2.21, it is clear that if h_1 has a fixed point (u_1, u_2) such that $\psi_1(u_1, u_2) \neq (\eta_L, \eta_F)$, then there is an SPS equilibrium with player 1 as a leader and player 2 as a follower. Also it is true that if h_2 has a fixed point (u_1, u_2) such that $\psi_2(u_1, u_2) \neq (\eta_F, \eta_L)$, then there is an SPS equilibrium with player 2 as a leader and player 1 as a follower.

The conclusion of the theorem follows immediately from Knaster-Tarski's fixed point theorem and Proposition 2.6.

Since ψ, Λ, and ρ_i's are increasing, the h_i are increasing. Moreover, $(\underline{u}, \underline{u}) \leq h_i(\underline{u}, \underline{u})$ by Proposition 2.4. Hence, by Knaster-Tarski's fixed point theorem, there exist fixed points for the h_i. Now from Proposition 2.6, it is clear that if (u_1, u_2) is a fixed point for h_1, then $\psi_1(u_1, u_2) \neq (\eta_L, \eta_F)$. Similarly, if (u_1, u_2) is a fixed point for h_2, then $\psi_2(u_1, u_2) \neq (\eta_F, \eta_L)$. ■

The above proof is also valid for the homogeneity case, i.e. when $\sigma_1 = \sigma_2$. Therefore, if the number of players in Bolton and Harris (1999) is two, then there will be asymmetric equilibria in addition to the symmetric equilibrium.

At the symmetric mixed strategy equilibrium in Bolton and Harris (1999), once the two players begin to choose the risky option, they keep choosing it indefinitely.[5] At an SPS equilibrium, however, the probability of event $\{p(t) \leq \underline{c}$, for some $t \leq T\}$ is always positive for all $T > 0$ and for all $p(0) > \underline{c}$. Thus, at asymmetric equilibria, even if the prior belief is high enough that one or both players do experimentation, there is a positive probability that they will end up with the safe option, and that they keep choosing it, which is impossible at the symmetric equilibrium.

We conjecture that in an N-player game with $0 < \sigma_N < ... < \sigma_1$, there are $N!$ types of asymmetric pure strategy equilibria, each of which has the same

---

[5] See Bolton and Harris (1999), p 363.

structure. That is, at each type of equilibria, there exist $(i_1, i_2, ..., i_N)$ which is a permutation of $(1, 2, ..., N)$ and N cutoff points $0 < c_N < ... < c_1 < 1$ such that player i_k is choosing the risky option if and only if $p \in (c_k, 1]$.

Lastly, it should not be puzzling much that LFMSE has two asymmetric equilibria. Since we have focused only on Markov strategies, many other equilibria may have been ignored. Fully characterizing all the equilibria is, however, certainly beyond the scope of this chapter.

Chapter 3

Strategic Experimentation in Markets

3.1 Introduction

It has long been observed that the existence of informational externality is often the source of inefficient resource allocation. For example, from Bolton and Harris (1999) and Chapter 2 of this dissertation, we see that the existence of informational externality will result in less than optimal level of social learning when people behave strategically. As it is impossible for people to fully appropriate the informational benefits of their experimentations in these models, the production of information at equilibrium will be suboptimal.

In this chapter, we will show that price competition *a la* Bertrand between firms can lead to the efficient level of social learning in the setting of Bolton and Harris (1999). Thus, the informational externality Bolton and Harris (1999) study is of such a kind that could be overcome by introducing a market competition. By contrast, it will be shown that the informational externality such as we study in Chapter 2 of this dissertation can't be remedied by introducing price competition. In Bolton and Harris (1999), one buyer's experimentation could be perfectly substituted by another buyer's experimentation. In the model of Chapter 2 of this dissertation, however, due to the differences in the qualities of the information from each buyer's experimentation, buyer 2's experimentation can't be perfectly substituted by buyer 1's experimentation. Hence, unless buyer 1 and 2 have opportunities to transfer between them in order to internalize this externality, mere price competition between firms will not guarantee efficiency.

This chapter is organized as follows. Section 2 describes the game. The efficient allocation will be reported in Section 3. In Section 4, we will show that if there is a symmetric equilibrium at which the incumbent firm's smooth pasting condition is satisfied, it is efficient. We will then comment on Bergemann and Välimäki (2000). In Bergemann and Välimäki (2000), they study the same model independently and conclude that equilibrium allocations are not efficient. We will argue that their reasoning is incorrect. In Section 5, it will be shown that linear prices equilibrium exists. In so doing, we resolve the existence problem, and at the same time, show how to calculate an equilibrium. In Section 6, we show that it is impossible to achieve efficient allocation with heterogeneous buyers.

3.2 A Continuous-Time Market Game

3.2.1 The Model

There are N infinitely-lived risk-neutral buyers who will be indexed by $i = 1, 2, ..., N$. There are two firms selling differentiated products, the incumbent (firm I) and the entrant (firm E). At each time period $[t, t + dt)$, each buyer has a unit demand at maximum for products of firm I and E. The flow payoff from the product of firm I is known to be s. The flow payoff from the product of firm E, μ, is, however, unknown to all the buyers and sellers at time 0, although μ is fixed at h or l. We will assume that $0 < l < s < h$. Buyers and firms share the common prior probability of μ being h at time 0.

At the beginning of each time period $[t, t + dt)$, both firms will announce simultaneously their prices. Given those prices, each buyer will choose the product of which firm to buy. Let $p^I(t)$ and $p^E(t)$ be the prices firm I and E charge buyers at time period $[t, t + dt)$. Then, the flow payoff of buyer i at time period $[t, t + dt)$ will be

$$dv_i = \begin{cases} (s - p^I(t))dt & \text{if she buys at firm } I \\ (\mu - p^E(t))dt + \sigma dZ_i(t) & \text{if she buys at firm } E \\ 0 & \text{otherwise,} \end{cases}$$

where the $dZ_i(t)$ are the independent standard Brownian motions for $i = 1, 2, ..., N$. Standard Brownian motions could be understood as continuous-time version of random walks. To be precise, $dZ_i(t)$ will be distributed following normal distribution whose mean and variance are zero and dt, respectively. Thus, the unknown quality of the product of firm E can be learned, but not perfectly due to the noises added to the payoffs.

Marginal cost of production of each firm is normalized to be zero. Hence, the flow payoff of firm $J \in \{I, E\}$ at time period $[t, t + dt)$, when $0 \leq k \leq N$ buyers buy its product, will be

$$dw^J = kp^J(t)dt.$$

Buyers and sellers are assumed to maximize the present discounted value of their payoff streams, namely $E[\int_0^\infty re^{-rt}dv_i(t)]$ for buyer $1 \leq i \leq N$, and $E[\int_0^\infty re^{-rt}dw^J(t)]$ for firm $J \in \{I, E\}$. It will be assumed that at the beginning of the time period $[t, t + dt)$, all the prices announced by the sellers, decisions of the buyers', and payoffs to all the players in the past are common knowledge. Therefore, there is no hidden information. The assumption of complete observability enables the players to learn not only from his own experimentation, but also from other player's experimentation. Therefore, information about the unknown quality is a kind of public good, which will be provided when buyers experiment.

3.2.2 Belief, Strategies and Equilibrium

Since there is no hidden information, the buyers and the firms will have common posterior belief $\beta(t)$ at each time t, which will be our natural choice for the state variable. We will focus on symmetric Markov perfect equilibria: Strategies of buyers and sellers will not depend on payoff irrelevant variables. *A pricing strategy* of firm $J \in \{I, E\}$, p^J, is a measurable function from $[0, 1]$ to R. Thus, each firm's pricing strategy will depend only on the state variable $\beta(t)$. Buyers' decisions, however, will depend on the announced prices, too. Hence, *an acceptance policy* of buyer i, $d_i = (d_i^I, d_i^E)$, is defined to be a measurable function from $[0, 1] \times R^2$ to $\{(0, 0), (0, 1), (1, 0)\}$. Note that an acceptance policy of buyer i does not allow him to mix between the two products. As Bertrand competition will not let buyers mix at equilibrium, there will be no loss of generality, however. Given a strategy profile $((d_i), p^I, p^E)$, we will use d_{-i} or p^{-J} to denote the strategy profiles of all the buyers and sellers except buyer i or seller J, respectively. That is, d_{-i} and p^{-I} will stand for $((d_j)_{j\neq i}, p^I, p^E)$ and $((d_i), p^E)$, respectively. A strategy profile $((d_i), p^I, p^E)$ is a *Markov perfect equilibrium* if d_i is an acceptance policy of buyer $i \in \{1, 2, ..., N\}$, p^J is a pricing strategy of firm $J \in \{I, E\}$, d_i is a best response to d_{-i} for all $i \in \{1, 2, ..., N\}$, and if p^J is a best response to p^{-J} for all $j \in \{I, E\}$.

In a discrete-time model, it is difficult to describe the posterior beliefs in a tractable way. By contrast, in a continuous-time model, the law of motion of $\beta(t)$ can be nicely described, which explains the main reason why we are using a continuous-time model. The following proposition can be proved as Proposition 2.1.

Proposition 3.1 *Suppose that d_i is the acceptance policy of buyer $i = 1, 2, ..., N$. Then, conditional on the information available at time t, the change in belief $d\beta(t)$ is distributed normally with mean 0 and variance*

$$\frac{\left(\sum_{i=1}^{N} d_i^E\right)}{\sigma^2}\Phi(\beta(t))dt,$$

where $\Phi(\beta) = [\beta(1-\beta)(h-l)]^2$.

Note that $\Phi(0) = \Phi(1) = 0$. Therefore, once they become sure about μ, from that point on, there will be no further change in β, which is a common feature of Bayesian learning. The more accurate the information is, the more radically the posterior belief will change. As an extreme case, if players could observe μ accurately in the period $[t, t+dt)$, then the posterior at the period $[t+dt, t+2dt)$ will be either l or h so that the ratio of the change in the belief to the length of time would be infinity. Hence, the variance would also be infinite. Overall, the magnitude of the variance measures the amount of information. It is, then, clear why we have $\sum_{i=1}^{N} d_i^E$ in the variance of $\beta(t)$.

3.2.3 Hamilton-Jacobi-Bellman Equations

Due to Proposition 3.1, we can use Itô's lemma to derive the Hamilton-Jacobi-Bellman (HJB) equation of buyer i. HJB equation is the dynamic programming equation in our continuous-time setting.

Let $m(\beta) = \beta h + (1-\beta)l$ be the myopic expected payoff from the product of firm E when μ is believed to be h with probability β. Then, the value function of buyer i will be determined as

$$\begin{aligned} u_i(\beta) \quad = \quad & \max\{s - p^I + \frac{\left(\sum_{j\neq i} d_j^E\right)}{r\sigma^2}\Phi(\beta)\frac{u_i''(\beta)}{2}, \\ & m(\beta) - p^E + \frac{\left(1 + \sum_{j\neq i} d_j^E\right)}{r\sigma^2}\Phi(\beta)\frac{u_i''(\beta)}{2}\}. \end{aligned} \tag{3.1}$$

The HJB equation consists of two parts. The first portions, $s - p^I$ and $m(\beta) - p^E$, represent the instant expected flow payoff. The expected future benefits from the extra information about the unknown quality are

$$\frac{\left(\sum_{j\neq i} d_j^E\right)}{r\sigma^2}\Phi(\beta)\frac{u_i''(\beta)}{2}$$

and

$$\frac{\left(1 + \sum_{j\neq i} d_j^E\right)}{r\sigma^2}\Phi(\beta)\frac{u_i''(\beta)}{2}.$$

As the portion of the variance part measures the amount of information that will be generated from buyers' experimentations, we could interpret $u_i''(p)$ as shadow price of information.

Similarly, the HJB equation for firm J will be

$$\pi^J(\beta) = \max_{p^J}\{p^J \sum_{i=1}^{N} d_i^J + \frac{\left(\sum_{i=1}^{N} d_i^E\right)}{r\sigma^2}\Phi(\beta)\frac{(\pi^J)''(\beta)}{2}\}.$$

As seen in Chapter 2, equilibria of this game will be determined by a system of second order differential equations of $(u_i)_i, \pi^I$, and π^E with appropriate boundary conditions. In the remainder of this paper, for notational convenience, we will suppress the dependence of $u_1, u_2, \pi^I, \pi^I, m$, and Φ on β as long as there is no risk of confusion.

3.3 Efficient Allocation

As a benchmark for our equilibrium analysis, we will investigate the team problem first. In the team problem, a social planner will maximize the average payoff of all the players. Since payments from the buyers to the sellers will be cancelled out, the social planner's problem is equivalent to maximizing the average payoff of the buyers when the two products are freely available. It is easy to see that this problem will be an optimal stopping problem. Therefore, the efficient allocation will be represented by a single cutoff $\hat{\beta}$; all the buyers should choose the product of firm E if and only if $\beta \in (\hat{\beta}, 1]$. Let u_* be the value function of the social planner's problem when the two products are for free. Thus, Nu_* will be the maximized sum of payoffs of all the players. For the proof of the following result, see Bolton and Harris (1999).

Theorem 3.2 *The efficient cutoff is*

$$\hat{\beta} = \frac{(s-l)(\lambda-1)}{(h-l)(\lambda-1)+2(h-s)},$$

where

$$\lambda = \sqrt{1+\frac{8r\sigma^2}{N(h-l)^2}}.$$

The value function of the social planner's problem is

$$u_* = \begin{cases} s & \text{for } \beta \in [0, \hat{\beta}] \\ m + a\beta^{\frac{1}{2}-\frac{1}{2}\lambda}(1-\beta)^{\frac{1}{2}\lambda+\frac{1}{2}} & \text{for } \beta \in (\hat{\beta}, 1], \end{cases}$$

where

$$a = \frac{2(h-s)}{\lambda-1}\left(\frac{\hat{\beta}}{1-\hat{\beta}}\right)^{\frac{1}{2}+\frac{1}{2}\lambda}.$$

Note that $\hat{\beta} < \beta_M$ where β_M is the myopically break-even point such that $m(\beta_M) = s$. As the information about the unknown quality μ is valuable, at optimal allocation, social learning should occur even if myopically it may not be worth choosing the product of firm E. As we noted before, u_*'' could be interpreted as shadow price of information. By direct calculation, it can be seen that $u''_*(\beta) > 0$ for $\beta \in (\hat{\beta}, 1)$.

3.4 Symmetric Equilibria

Due to price competition *a la* Bertrand, multiplicity of equilibria is inevitable. Therefore, to narrow down the set of equilibria, we need to put some restrictions on it. As we do not have a satisfactory refinement concept for continuous-time games yet, we will put only the weakest restrictions. We require that whenever players cease to learn, the equilibria should be Nash equilibria in undominated strategies in the corresponding static game. Therefore, for instance, if $\beta = 0$, i.e. the quality of the product of firm E is believed to be l for sure, then we will choose as the equilibrium $(p^I, p^E) = (s - l, 0)$ with all the buyers choosing firm E. We believe that undominatedness in static case is the minimum that should be satisfied by any attempt to refine equilibria of continuous-time games.

Under this minimum restriction, we will show in the following that if a symmetric equilibrium is characterized by a cutoff β^* and if the value function of the incumbent is smooth at the cutoff, then the equilibrium cutoff β^* is identical with $\hat{\beta}$. Therefore, all the symmetric equilibria of this type will be efficient. Bertrand competition adjust prices so that the resulting social learning is to the adequate amount.

3.4.1 Equilibrium Prices

Suppose that there is a symmetric equilibrium with a cutoff β^* such that all the buyers will choose firm E if and only if $\beta \in (\beta^*, 1]$. Let π^J and u be the value function of firm $J \in \{I, E\}$ and the common value function of the buyers at this symmetric equilibrium, respectively.

Then, we will have from (3.1)

$$\begin{aligned} u &= m - p^E + \frac{N}{r\sigma^2}\Phi\frac{u''}{2} \\ &\geq s - p^I + \frac{N-1}{r\sigma^2}\Phi\frac{u''}{2}. \end{aligned}$$

for $\beta \in (\beta^*, 1]$. At equilibrium, price competition between two firms will force the above inequality to hold with equality, which implies that the opportunity cost of choosing the product of firm E, $(s - p^I) - (m - p^E)$, will be equal to the informational benefit, $\frac{1}{r\sigma^2}\Phi\frac{u''}{2}$. By rearranging terms, we have

$$p^I = p^E + (s - m) - \frac{1}{r\sigma^2}\Phi\frac{u''}{2}. \tag{3.2}$$

In a similar way, for $\beta \in (\beta^*, 1]$, we have

$$\begin{aligned} \pi^I &= \frac{N}{r\sigma^2}\Phi\frac{(\pi^I)''}{2} \\ &\geq Np^I, \end{aligned}$$

and

$$\begin{aligned}\pi^E &= Np^E + \frac{N}{r\sigma^2}\Phi\frac{(\pi^E)''}{2} \\ &\geq 0.\end{aligned}$$

Thus,

$$\begin{aligned}\frac{1}{r\sigma^2}\Phi\frac{(\pi^I)''}{2} &\geq p^I \\ p^E + \frac{1}{r\sigma^2}\Phi\frac{(\pi^E)''}{2} &\geq 0.\end{aligned}$$

By combining these two inequalities with (3.2), we have the following necessary condition for an equilibrium price p^I:

$$(s-m) - \frac{1}{r\sigma^2}\Phi\frac{(\pi^E)''}{2} - \frac{1}{r\sigma^2}\Phi\frac{u''}{2} \leq p^I \leq \frac{1}{r\sigma^2}\Phi\frac{(\pi^I)''}{2}. \tag{3.3}$$

3.4.2 Efficiency

If $\beta \in [0, \beta^*]$, there will be no experimentation. Hence, from undominatedness, we obtain

$$\begin{aligned}p^I &= s - m(\beta) \\ p^E &= 0\end{aligned}$$

as the equilibrium prices. It is immediate that $u(\beta) = m(\beta)$ for $\beta \in [0, \beta^*]$. The payoffs of the buyers will be exactly equal to the value of their outside options. It is also clear that

$$\begin{aligned}\pi^I &= N(s-m)\\ \pi^E &= 0\end{aligned}$$

for $\beta \in [0, \beta^*]$.

To find an equilibrium, we have to solve a system of second order differential equations for $\beta \in (\beta^*, 1]$. Boundary conditions for this system are from the value matching conditions and the smooth pasting conditions. The value matching conditions are simply conditions about continuity of the value functions at the cutoff point. The smooth pasting conditions are a kind of first order conditions for our stochastic dynamic programming problem, which require that the left derivative and the right derivative of each value function at the cutoff point should be the same.

By direct substitution, it can be seen that the general solution of the following differential equation

$$\pi^I = \frac{N}{r\sigma^2}\Phi\frac{(\pi^I)''}{2}$$

will be

$$\pi^I = c_1\beta^{\frac{1}{2}-\frac{1}{2}\lambda}(1-\beta)^{\frac{1}{2}\lambda+\frac{1}{2}} + c_2\beta^{\frac{1}{2}\lambda+\frac{1}{2}}(1-\beta)^{\frac{1}{2}-\frac{1}{2}\lambda},$$

where

$$\lambda = \sqrt{1+\frac{8r\sigma^2}{N(h-l)^2}}.$$

Since π^I is bounded for all $\beta \in (\beta^*, 1]$, we can see that $c_2 = 0$. Thus,

$$\pi^I = c_1 \beta^{\frac{1}{2} - \frac{1}{2}\lambda} (1-\beta)^{\frac{1}{2}\lambda + \frac{1}{2}}.$$

The value matching condition and the smooth pasting condition for firm I are

$$\begin{aligned} \pi^I(\beta^*) &= N(s - m(\beta^*)) \\ (\pi^I)'(\beta^*) &= [N(s - m(\beta^*))]' = N(l - h). \end{aligned} \tag{3.4}$$

Therefore, two unknowns, c_1 and β^*, will be determined from the above boundary conditions of firm I. It can be shown that the equilibrium cutoff β^* coincides with the socially efficient cutoff level.

Theorem 3.3 *If there is a symmetric equilibrium at which the smooth pasting condition of the incumbent is satisfied, it is efficient.*

Proof . Note that

$$(\pi^I)' = \frac{\pi^I}{2}\left(\frac{1-\lambda}{\beta} - \frac{1+\lambda}{\beta}\right).$$

Therefore,

$$\begin{aligned} (\pi^I)'(\beta^*) &= \frac{\pi^I(\beta^*)}{2}\left(\frac{1-\lambda}{\beta^*} - \frac{1+\lambda}{\beta^*}\right) \\ &= \frac{N(s - m(\beta^*))}{2}\left(\frac{1-\lambda}{\beta^*} - \frac{1+\lambda}{\beta^*}\right). \end{aligned}$$

From the smooth pasting condition (3.4), we have

$$\frac{(s - m(\beta^*))}{2}\left(\frac{1-\lambda}{\beta^*} - \frac{1+\lambda}{\beta^*}\right) = (l - h).$$

It is immediate that

$$\beta^* = \frac{(s-l)(\lambda-1)}{(h-l)(\lambda-1)+2(h-s)} = \hat{\beta}.$$

■

The intuition behind this efficiency result is rather simple. In the social planner's problem, the opportunity cost of buyers' choosing the risky option is $N(s - m(\beta))$. As $\pi^I(\beta) = N(s - m(\beta))$, the private cost of the incumbent firm for letting buyers choose firm E will be exactly equal to the social opportunity cost. Therefore, in the profit maximization problem of firm I, the social opportunity cost is fully reflected. As the results of buyers' experimentation shift β up or down, the resulting social pessimism or optimism about the product of firm E will be fully reflected in p^E. Thus, when buyers are buying at firm E, the sum of the informational benefits for buyers and that for firm E will be zero. Hence, when firm E is selling, the total social informational benefit will be equal to the informational benefit for firm I who can fully observe buyers' experimentation. Indeed, from $\beta^* = \hat{\beta}$ and (3.4), we can see that

$$c_1 = \frac{2N(h-s)}{\lambda-1}\left(\frac{\hat{\beta}}{1-\hat{\beta}}\right)^{\frac{1}{2}+\frac{1}{2}\lambda}.$$

Hence,

$$\frac{N}{r\sigma^2}\Phi\frac{(\pi^I)''}{2} = \frac{N}{r\sigma^2}\Phi\frac{u_*''}{2}.$$

In summary, since firm I's private cost and benefit of letting buyers experiment are equal to the social opportunity cost of and the social informational benefit from buyers' experimentation, firm I's optimal choice will be identical with socially optimal timing.

It is also notable that for the efficiency result, we do not need firm E. What is crucial is that the private cost of firm I for letting buyers to choose the risky option is equal to the social cost. Therefore, even in the case where buyers have free access to the risky product whereas the safe option is provided by firm I, we could still obtain the efficiency result. Similarly, it can be shown that we could achieve the efficient allocation with only firm E selling the risky product, while the buyers have free access to the safe one.

3.4.3 Comments on Bergemann and Välimäki (2000)

In an earlier but independent work, Bergemann and Välimäki (2000) analyze the same game with ours. They conclude that there will be excessive experimentation at the symmetric equilibrium. Their paper is, however, flawed by a serious mistake.

To overcome the multiplicity of equilibria, they introduce the concept of cautious strategies in their paper. To be brief, cautious strategies are those satisfying the following relation:

$$p^I = \frac{1}{r\sigma^2}\Phi\frac{(\pi^I)''}{2}.$$

By focusing on cautious equilibrium, however, they kill one degree of freedom that is indispensable to satisfy the two boundary conditions for π^E. Since (3.3) is the main source for multiplicity of equilibria, they seem to attempt to overcome this problem by choosing a specific p^I. In so doing, however, they throw away too much so that they can't have enough degree of freedom to satisfy all the boundary conditions. For instance, at the state they claim to be an equilibrium, we can show that the smooth pasting condition for firm I is violated. In private correspondence, they argue that smooth pasting condition is not a necessary condition for optimality for the incumbent since the payoff from the stopped process of the incumbent is kinked at the cutoff of the cautious equilibrium. This observation is correct. We do not know yet the general necessary conditions for optimality in case the payoff from the stopped process is not differentiable, which is the reason why we have an additional assumption about the smoothness of the value function of the incumbent in Theorem 3.3. Without a doubt, however, in order to claim that a state is an equilibrium, we ought to show explicitly that all the people in the model are indeed optimizing. Bergemann and Välimäki mistakenly identify equation (18) with (19) in their paper to conclude that they can omit to check the optimality conditions for the incumbent. This is a pity, since still we do not know if the cautious equilibrium in Bergemann and Välimäki (2000) is really an equilibrium.

3.5 Linear Prices

Now we have two more differential equations:

$$\begin{aligned} u &= m - p^E + \frac{N}{r\sigma^2}\Phi\frac{u''}{2} \\ \pi^E &= Np^E + \frac{N}{r\sigma^2}\Phi\frac{(\pi^E)''}{2}. \end{aligned}$$

We know from Theorem 3.3, all the symmetric equilibria are efficient if the value function of the incumbent is smooth. Therefore, we do not have to solve both of the above two equations in this case. When we have π^E, from the accounting identity

$$\pi^I + \pi^E + Nu = Nu_*,$$

we can get u as a residual. Since the value matching condition and the smooth pasting condition for firm E are

$$\begin{aligned} \pi^E(\beta^*) &= 0 \\ (\pi^E)'(\beta^*) &= 0, \end{aligned}$$

and since u_*, the solution to the Team problem, will satisfy its boundary conditions

$$\begin{aligned} u_*(\beta^*) &= s \\ u_*'(\beta^*) &= 0, \end{aligned}$$

the value function u obtained as a residual will automatically satisfy its own value matching condition and smooth pasting condition, which are

$$\begin{aligned} u(\beta^*) &= m(\beta^*) \\ u'(\beta^*) &= h - l. \end{aligned}$$

With fixed p^E, p^I will be determined by (3.2). Therefore, to find an equilibrium is to find a pricing policy p^E such that the resulting π^E and π^I will satisfy all the boundary conditions while p^I determined by (3.2) will satisfy (3.3). Since the binding restrictions for p^E are local (it is required only at β^*), there will be plethora of equilibria.

In this section, we will show that there is a very simple pricing policies to support the equilibrium cutoff: A selling firm's price will be linear function of the posterior belief β. We have already chosen

$$p^I = s - m(\beta)$$

as the equilibrium price of firm I for $\beta \in [0, \beta^*]$. Thus, it suffices to show the existence of equilibrium p^E which is a linear function of β for $\beta \in (\beta^*, 1]$.

Theorem 3.4 *There is a symmetric equilibrium where*

$$p^E = m(\beta) - s$$

for $\beta \in (\beta^*, 1]$. *There is no other equilibrium policies where* p^E *is a linear function of* β *for* $\beta \in (\beta^*, 1]$.

Proof . Suppose that $p^E = a + b\beta$ is an equilibrium price. Due to the undominatedness in static Bertrand competition, we have

$$p^E(1) = a + b = h - s.$$

Since

$$\pi^E = Np^E + \frac{N}{r\sigma^2}\Phi\frac{(\pi^E)''}{2}$$

for $\beta \in (\beta^*, 1]$, and since π^E is bounded, the general solution for π^E will be

$$\pi^E = N(h-s) + Nb(\beta - 1) + c\beta^{\frac{1}{2}-\frac{1}{2}\lambda}(1-\beta)^{\frac{1}{2}\lambda+\frac{1}{2}}.$$

Thus, π^E will have two parameters to determine, and these will be fixed by the following conditions.

$$\begin{aligned} \pi^E(\beta^*) &= 0 \\ (\pi^E)'(\beta^*) &= 0. \end{aligned}$$

The result is

$$\begin{aligned} b &= h - l \\ c &= \frac{2N(h-s)}{\lambda - 1}\left(\frac{\beta^*}{1-\beta^*}\right)^{\frac{1}{2}+\frac{1}{2}\lambda} = Na. \end{aligned}$$

Recall that β^* is chosen already so that the optimality condition for firm I is satisfied. Also, by defining u as

$$u = u_* - \frac{1}{N}(\pi^I + \pi^E),$$

it can be seen that the optimality conditions for the buyers are automatically satisfied.

To verify that this is indeed an equilibrium, we have to show that (3.3) is not violated. Since

$$\begin{aligned} p^I &= p^E + (s - m) - \frac{1}{r\sigma^2}\Phi\frac{u''}{2} \\ &= -\frac{1}{r\sigma^2}\Phi\frac{u''}{2}, \end{aligned}$$

it suffices to show that

$$-\frac{1}{r\sigma^2}\Phi\frac{u''}{2} \leq \frac{1}{r\sigma^2}\Phi\frac{(\pi^I)''}{2}.$$

From Theorem 3.3, it can be shown that

$$\pi^I = Na\beta^{\frac{1}{2}-\frac{1}{2}\lambda}(1-\beta)^{\frac{1}{2}\lambda+\frac{1}{2}}.$$

From

$$u = u_* - \frac{1}{N}(\pi^I + \pi^E),$$

we have

$$u = s - a\beta^{\frac{1}{2}-\frac{1}{2}\lambda}(1-\beta)^{\frac{1}{2}\lambda+\frac{1}{2}}.$$

Therefore,

$$(\pi^I)'' = N(-u'') > -u''.$$

■

In the above proof, we have shown that if the price of the selling firm is linear in β, then for $\beta \geq \beta^*$,

$$\pi^I = Naf(\beta)$$

$$\begin{aligned}\pi^E &= N[m(\beta) - s] + Naf(\beta) \\ u &= s - af(\beta),\end{aligned}$$

where

$$f(\beta) = \beta^{\frac{1}{2}-\frac{1}{2}\lambda}(1-\beta)^{\frac{1}{2}\lambda+\frac{1}{2}}.$$

Therefore,

$$\frac{1}{r\sigma^2}\Phi\frac{(\pi^E)''}{2} + \frac{N}{r\sigma^2}\Phi\frac{u''}{2} = 0.$$

That is, the sum of the informational benefit for firm E and the informational benefit of all buyers will always be zero when firm E is selling. This again verifies our intuition in Section 3.4.2.

3.6 Heterogeneous Buyers

Instead of N homogeneous buyers, suppose that we have 2 buyers who have different abilities to evaluate the uncertain quality of the product of firm E. More precisely, buyer i's instant flow payoff dv_i is assumed to be

$$dv_i = \begin{cases} (s - p^I(t))dt & \text{if she buys at firm } I \\ (\mu - p^E(t))dt + \sigma_i dZ_i(t) & \text{if she buys at firm } E \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < \sigma_2 < \sigma_1$. Hence, buyer 2's experimentation will provide more precise information about the true quality of the product of firm E. In fact, this is the same model we study in Chapter 2 except that two options are now being sold by two competing firms. From Proposition 2.1, it is immediate that $d\beta(t)$ will be distributed normally with mean 0 and variance

$$(\frac{d_1^E}{\sigma_1^2}+\frac{d_2^E}{\sigma_2^2})\Phi(\beta(t))dt.$$

In this section, we will show that unlike the case of homogeneous buyers it is impossible to achieve the efficient allocation.

We will allow the possibility of price discrimination. Thus, a pricing strategy of firm $J \in \{I, E\}$ is now defined as $p^J = (p_1^J, p_2^J)$, where p_i^J is the price that firm J charges buyer i. Then, buyer 1's HJB equation will be

$$u_1 = \max\{s - p_1^I + \frac{d_2^E}{r\sigma_2^2}\Phi\frac{u_1''}{2}, m - p_1^E + \frac{1}{r}(\frac{1}{\sigma_1^2}+\frac{d_2^E}{\sigma_2^2})\Phi\frac{u_1''}{2}\}.$$

Price competition between two firms will make the two terms in max operator equal to each other at equilibrium. Hence,

$$p_1^E = m - s + p_1^I + \frac{1}{r\sigma_1^2}\Phi\frac{u_1''}{2}. \tag{3.5}$$

Similarly, from buyer 2's HJB equation, we have

$$p_2^E = m - s + p_2^I + \frac{1}{r\sigma_2^2}\Phi\frac{u_2''}{2}. \tag{3.6}$$

What (3.5) and (3.6) imply is simply that private opportunity cost of choosing the risky option is equal to the private informational benefit from experimentation, which, obviously, ought to be true at equilibrium.

HJB equations for firm I and E will be

$$\begin{aligned}
\pi^I &= \max_{p_1^I, p_2^I}\{p_1^I + p_2^I, p_1^I + \frac{1}{r\sigma_2^2}\Phi\frac{(\pi^I)''}{2}, \\
&\quad p_2^I + \frac{1}{r\sigma_1^2}\Phi\frac{(\pi^I)''}{2}, \frac{1}{r}(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2})\Phi\frac{(\pi^I)''}{2}\},
\end{aligned}$$

and

$$\begin{aligned}
\pi^E &= \max_{p_1^E, p_2^E}\{0, p_2^E + \frac{1}{r\sigma_2^2}\Phi\frac{(\pi^E)''}{2}, \\
&\quad p_1^E + \frac{1}{r\sigma_1^2}\Phi\frac{(\pi^E)''}{2}, p_1^E + p_2^E\frac{1}{r}(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2})\Phi\frac{(\pi^E)''}{2}\}.
\end{aligned}$$

Now we will prove that it is impossible to obtain efficient allocation at equilibrium.

Theorem 3.5 When buyers are heterogeneous, there is no equilibrium which is efficient.

Proof . With (3.5) and (3.6), the HJB equation of firm E will be

$$\begin{aligned}
\pi^E &= \max_{p_1^E, p_2^E}\{0, m - s + p_2^I + \frac{1}{r\sigma_2^2}\Phi\frac{u_2''}{2} + \frac{1}{r\sigma_2^2}\Phi\frac{(\pi^E)''}{2}, \\
&\quad m - s + p_1^I + \frac{1}{r\sigma_1^2}\Phi\frac{u_1''}{2} + \frac{1}{r\sigma_1^2}\Phi\frac{(\pi^E)''}{2}, \\
&\quad 2(m-s) + p_1^I + p_2^I + \frac{1}{r\sigma_1^2}\Phi\frac{u_1''}{2} + \frac{1}{r\sigma_2^2}\Phi\frac{u_2''}{2} + \frac{1}{r}(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2})\Phi\frac{(\pi^E)''}{2}\}.
\end{aligned}$$

We know from the analysis of Team problem in Chapter 2 that the efficient allocation will be represented by two cutoffs $0 < \beta_2 < \beta_1 < 1$ such that buyer i selects the product of firm E if and only if $\beta > \beta_i$. Hence, suppose that we have an equilibrium with two cutoffs $0 < \beta'_2 < \beta'_1$, and that buyer i selects the product of firm E if and only if $\beta > \beta'_i$. This implies that

$$\begin{aligned} & \text{the second term in firm } I\text{'s HJB equation} \\ \gtreqless \quad & \text{the last term in firm } I\text{'s HJB equation} \end{aligned}$$

if and only if

$$\begin{aligned} & \text{the second term in firm } E\text{'s HJB equation} \\ \gtreqless \quad & \text{the last term in firm } E\text{'s HJB equation.} \end{aligned}$$

Therefore, the cutoff β'_1 will be determined by comparing

$$\frac{1}{r\sigma_1^2}\Phi\left(\frac{u_1'' + (\pi^I)'' + (\pi^E)''}{2}\right)$$

and

$$s - m.$$

Similarly, the cutoff β'_2 will be determined by comparing

$$\frac{1}{r\sigma_2^2}\Phi\left(\frac{u_2'' + (\pi^I)'' + (\pi^E)''}{2}\right)$$

and

$$s - m.$$

In the social planner's problem, however, we compare the social informational benefit, which is

$$\frac{1}{r\sigma_i^2}\Phi\left(\frac{u_1'' + u_2'' + (\pi^I)'' + (\pi^E)''}{2}\right),$$

with $(s-m)$ to determine β_i. The value functions of both players can be shown to be concave as in Chapter 2. Thus, $u_i'' \neq 0$ so that the equilibrium allocation will not be efficient. ■

Unlike the homogeneous case, price competition will not guarantee efficiency here. In the homogeneous case, one buyer's experimentation could be perfectly substituted with another buyer's. In the heterogeneous case, however, due to the differences in the qualities of the information their experimentations will generate, buyer 2's experimentation can only partially be substituted with buyer 1's. Hence, without having opportunities for buyer 1 and buyer 2 to sign a contract in order to internalize this externality, it is impossible for market equilibria to be efficient.

Chapter 4

Existence of Pure Markov Strategy Equilibrium in Discrete-Time Multi-Player Multi-Armed Bandit Problems

4.1 Introduction

In Chapter 2 and 3, we analyze strategic experimentation with continuous-time models, even though it is still an open question how to generally set up continuous-time multi-player multi-armed bandit problems. Most difficulties are around the notion of strategies. Nevertheless, in the models we use in Chapter 2 and 3, we could represent the law of motion of the posterior in a closed form, which explains why we adopt continuous-time models in spite of all the technical difficulties.

In this chapter, we turn to discrete-time models. Unlike continuous-time models, we can provide a general setting for multi-player multi-armed bandit problems in discrete-time set up. We will assume perfect observability. That is, at any period m, all the previous selections of all the players and the results of their choices are commonly known to all the players. Therefore, the players will share the same information about the k alternatives, and thus, hidden information will not be an issue. Under this assumption, we will generalize Section 2.2 in Berry and Fristedt (1985) to n player case. Then, we will show that there exists pure Markov strategy equilibria.

This chapter is organized as follows. General setting is described in Section 2. Best response is defined and shown to be non-empty in Section 3. In Section 4, we will show that value functions are continuous in other players' strategies. Existence of pure Markov strategy equilibrium is proved in Section 5.

4.2 General Setting

There are n players. Each of them has k alternatives to select at each time $t = 1, 2, \ldots$. If player i selects j-th option at $t = m$, his payoff will be $X^i_{j,m}$, which is a random variable. The distribution of $\{X^i_{j,m}\}$ is not known to any player at $t = 0$. We assume that the players have common prior belief about $\{X^i_{j,m}\}$ at $t = 0$. We also assume that $\{X^i_{j,m}\}$ are independent, and that the distribution of $X^i_{j,m}$ is equal to that of $X^{i'}_{j',m'}$ if and only if $j = j'$. Hence, each player will be in an identical situation, and his payoff relevant variable will not be directly influenced by other players' actions. The actions of other players will have effects on player i's decision not because his payoff will vary according to their choices, but because he might get some information about the uncertain alternatives by observing their selections and the results of their decisions.

Let $\mathcal{D}$ represent the space of probability distributions on R, the set of real numbers. We will use the topology of convergence in distribution on $\mathcal{D}$. That is, with $Q_n, Q \in \mathcal{D}$, $Q_n \to Q$ as $n \to \infty$ if and only if $\int_R h dQ_n \to \int_R h dQ$ for all bounded continuous function h on R. The space of ordered k-tuples of $\mathcal{D}$, $\mathcal{D}^k$, will be considered to have the product topology arising from the topology on $\mathcal{D}$. The coordinate Q_j of $Q = (Q_1, Q_2, \ldots, Q_k) \in \mathcal{D}^k$ is interpreted as the true but unknown distribution governing the payoff of j-th alternative. The space of probability distributions on $\mathcal{D}^k$ is $\mathcal{D}(\mathcal{D}^k)$. The common prior belief about Q will be described by an element of $\mathcal{D}(\mathcal{D}^k)$. We will use the topology of convergence in distribution on $\mathcal{D}(\mathcal{D}^k)$.

Now we will define the probability space Ω. Let Ω be defined as

$$\Omega = \mathcal{D}^k \times \Pi_{i=1}^n \Pi_{j=1}^k \Pi_{m=1}^\infty (0,1).$$

The probability measure P on Ω is the product of a member of $\mathcal{D}(\mathcal{D}^k)$ and Lebesgue measure on each unit intervals. We can denote an element of Ω as

$$\omega = (Q_i, 1 \le i \le k; \omega_{j,m}^i, 1 \le i \le n, 1 \le j \le k, m = 1, 2, ...),$$

where each $Q_i \in \mathcal{D}$ and each $\omega_{j,m}^i \in (0,1)$. Let the random variable $X_{j,m}^i$, the payoff that player i will get if he selects j-th option at $t = m$, be defined as

$$X_{j,m}^i(\omega) = Q_j^{-1}(\omega_{j,m}^i),$$

where Q_j^{-1} is the usual right continuous inverse function of Q_j. The structure of Ω reflects the idea that the payoff of each option is determined when $\omega \in \Omega$ is fixed whether or not they are observed. For player i to select j-th alternative at $t = m$ is to get to observe the value $X_{j,m}^i(\omega)$. Note that $\{X_{j,m}^i\}$ are independent conditional on $(Q_1, Q_2, ..., Q_k)$.

When a player selects some option, all the information that will be relevant later on is which alternative he selects and what his payoff from that option is. Thus, his experimentation can be summarized by $e \in E$, where

$$E = \{(e_1, e_2) : e_1 \in \{1, 2, ..., k\}, e_2 \in R\}.$$

We will equip E with the measure μ defined as follows: For all $i \in \{1, 2, ..., k\}$, and for all Borel subset $A \subset R$,

$$\mu(i, A) = \mu_i(A),$$

where μ_i is the measure induced by Q_i on R. One shot experimentation of n players will be represented by a member of E^n. We will equip E^n with the product measure arising from μ. Let $\mathcal{H}_m$ be the set of histories at the beginning of period m. Hence, $\mathcal{H}_m = E^{n(m-1)}$. Let $\mathcal{H}_1$ be defined as $\varnothing$. A *player i's strategy at period m*, τ_m^i, is a measurable mapping from $\mathcal{H}_m$ into Δ, where Δ is the $(k-1)$-dimensional simplex with a measure induced from Lebesgue measure on R^{k-1}:

$$\Delta = \{(p_1, p_2, ..., p_k) \in R^k : \sum_{i=1}^{k} p_i = 1, p_j \geq 0 \text{ for all } j = 1, 2, ..., k\}.$$

A *pure strategy of player i at period m* is a player i's strategy which maps $\mathcal{H}_m$ into vertices of Δ. A *strategy of player i* $\tau^i = (\tau_1^i, \tau_2^i, ...)$ is defined as a collection of player i's strategies at each period. τ^i is a *pure strategy* if all the τ_m^i are pure. A *strategy profile* $\tau = (\tau^1, ..., \tau^n)$ is a collection of strategies of all the players, and it is pure if all the τ^i are pure. Following convention, we will use the notation $\tau = (\tau^i, \tau^{-i})$ to decompose τ into player i's strategy and all the other players' strategies. Let $\mathcal{T}$ be the set of strategies of each player. Hence, $\mathcal{T}^n$ will be the set of all the strategy profiles.

Let Z_m^i be the realized payoff of player i at period m, which will depend on the strategy profile the players are adopting. Suppose for the time being that τ is a pure strategy profile. Then, Z_m^i are defined recursively as

$$\begin{aligned} Z_1^i &= X^i_{\tau_1^i(\varnothing),1} \\ Z_1 &= (Z_1^1, ..., Z_1^n) \\ Z_m^i &= X^i_{\tau_m^i(Z_1,...,Z_{m-1}),m} \text{ for } m > 1, \end{aligned}$$

where as an abuse of notation, we use $\tau_m^i(Z_1, ..., Z_{m-1})$ to denote player i's selection at period m, which could be justified by the observation that, given τ, knowing all the values of $(Z_1, ..., Z_{m-1})$ is equivalent to knowing the history at period m. From the definition, it is clear that Z_m^i is dependent on τ. For notational convenience, however, we will suppress the dependence of Z_m^i on τ as long as there is no risk of confusion. Player i's payoff when a strategy profile τ is played is

$$E_\tau^i(\sum_{m=1}^{\infty} \alpha_m Z_m^i),$$

where $A = (a_1, \alpha_2, ...)$ is the sequence of discount factors and the subscript τ represents the dependence of the expectation on τ. The above definition will generalize to the case when τ is not pure. If τ is not a pure strategy profile, then it will be a mixture of pure strategy profiles. The payoff will be then the average of the each payoff from the constituting pure strategy profiles.

4.3 Best Responses

Let G be the common prior that the players have at the beginning of the game. Let the value function of player i given τ^{-i}, $V^i(G;\tau^{-i})$, be defined as

$$V^i(G;\tau^{-i}) = \sup_{\tau^i} E^i_\tau(\sum_{m=1}^{\infty} \alpha_m Z^i_m). \tag{4.1}$$

Assumption 4.1 Each element α_m of the discount sequence A is nonnegative, and $\sum_{m=1}^{\infty} \alpha_m < \infty$.

Assumption 4.2 Each component Q_i of $(Q_1, Q_2, ..., Q_k) \in \mathcal{D}^k$ has finite first absolute moment with G-probability one, and that this moment has finite G-expectation.

The previous assumptions guarantee that $V^i(G;\tau^{-i})$ is bounded for all G and τ^{-i}.[6] To further our analysis, we need one more technical result. Since the players keep updating their beliefs as they get to observe the experimentations at each period, we want this updating process measurable. A modification of Theorem V.8.1 in Parthasarathy (1967) and Lemma 2.2.1 in Berry and Fristedt (1985) will give us what we need. Since we are using $\mathcal{D}(\mathcal{D}^k)$ as the set of state variables, all of these technicalities are unavoidable.

[6] This could be proved in a similar way as Theorem 2.5.1 is proved in Berry and Fristedt (1985).

Proposition 4.3 *For all $c = (c_1, ..., c_n) \in \{1, 2, ..., k\}^n$, there exists a measurable function $f_c : R^n \times \mathcal{D}(\mathcal{D}^k) \to \mathcal{D}(\mathcal{D}^k)$ such that for every Borel set $D \in \mathcal{D}^k$,*

$$\begin{aligned} & f_c(X^1_{c_1,1}(\omega), ..., X^n_{c_n,1}(\omega); G)(D) \\ = \quad & P(D \times \Pi^n_{i=1}\Pi^k_{j=1}\Pi^\infty_{m=1}(0,1)|X^1_{c_1,1}, ..., X^n_{c_n,1})(\omega) \ \text{ a.e. } \omega. \end{aligned}$$

If player i selects j-th option at $t = 1$, $(Z^1_{\tau^1_1(\varnothing),1}, ..., Z^i_{j,1}, ..., Z^1_{\tau^n_1(\varnothing),1})$ will be the outcome. We will denote the common random posterior at the beginning of period 2 as $G^{(1),j}$, where superscript j indicates that player i selected j-th option previously. Recall that the updating process is measurable. Then, it is not difficult to see that

$$\begin{aligned} V^i(G; \tau^{-i}) &= \sup_{\tau^i} E^i_\tau(\sum_{m=1}^\infty \alpha_m Z^i_m) \\ &= \vee^k_{j=1}[\alpha_1 E(X^i_{j,1}) + \sup_{\tau^i_1(\varnothing)=j} E_\tau(\sum_{m=2}^\infty \alpha_m Z^i_m)] \\ &= \vee^k_{j=1}[\alpha_1 E(X^i_{j,1}) + E\left(V^i(G^{(1),j}; \tau^{-i})\right)]. \end{aligned}$$

Now we will define the best response correspondence of player i.

Definition 4.4 Player i's **best response correspondence** $B^i : \mathcal{T}^{n-1} \to 2^{\mathcal{T}}$ is defined as

$$B^i(\tau^{-i}) = \{\tau^i \in \mathcal{T} : V^i(G; \tau^{-i}) = E^i_{(\tau^i,\tau^{-i})}(\sum_{m=1}^\infty \alpha_m Z^i_m)\}.$$

Lemma 4.5 *Player i's best response correspondence B^i is non-empty for all $\tau^{-i} \in T^{n-1}$ and for all i.*

Proof . We will construct $\hat{\tau}^i \in B^i(\tau^{-i})$ by induction as follows. Let $\hat{\tau}_1^i(\varnothing)$ be the smallest j such that

$$V^i(G;\tau^{-i}) = \alpha_1 E(X_{j,1}^i) + E\left(V^i(G^{(1),j};\tau^{-i})\right).$$

Let $G^{(1)}$ be the random posterior at period 2. Then, we have

$$V^i(G;\tau^{-i}) = E_{\hat{\tau}}\left(\alpha_1 Z_1^i + V^i(G^{(1)};\tau^{-i})\right),$$

where the subscript $\hat{\tau}$ indicates that the expectation is dependent on $\hat{\tau}_1^i$ and τ^{-i}. After $G^{(n)}$ and $\hat{\tau}_n^i$ are defined, $\hat{\tau}_{n+1}^i$ will be defined similarly. By induction, we will have

$$V^i(G;\tau^{-i}) = E_{\hat{\tau}}\left(\sum_{m=1}^{n} \alpha_m Z_m^i + V^i(G^{(n)};\tau^{-i})\right).$$

The proof will be completed by showing that

$$\lim_{n\to\infty} E_{\hat{\tau}}\Big(\sum_{m=n+1}^{\infty} \alpha_m Z_m^i - V^i(G^{(n)};\tau^{-i})\Big) = 0.$$

The argument is standard and we will omit the detail.[7] ∎

We used the finiteness of the set of the alternatives in the proof of Lemma 4.5 in an essential way. As a matter of fact, if the set of the alternatives is not finite, optimal policies may not exist. Indeed, Easley and Kiefer (1989) shows that if the set of alternatives are uncountable, optimal policies may not exist even in

[7] For example, see p.44, Berry and Fristedt (1985).

one-player case.

In the proof of Proposition 4.5, we constructed a best response which is a pure strategy. Hence,

Proposition 4.6 *Against every $\tau^{-i} \in T^{n-1}$, there exists at least one pure strategy best response for player i.*

4.4 Value Functions

In general, player i's value function $V^i(G; \tau^{-i})$ is not a continuous function in G.[8] In the following special case, however, $V^i(G; \tau^{-i})$ will be continuous in G.

Lemma 4.7 *Suppose that for some $G \in \mathcal{D}(\mathcal{D}^k)$,*

$$G^{(1),n} = \sum_c \alpha_c^n f_c(X_{c_1,1}^1, ..., X_{c_n,1}^n; G),$$

and

$$G^{(1)} = \sum_c \alpha_c f_c(X_{c_1,1}^1, ..., X_{c_n,1}^n; G),$$

where f_c is the function defined in Proposition 4.3, the summations are taken over all possible $c \in \{1, 2, ..., k\}^n$, and $\sum_c \alpha_c^n = \sum_c \alpha_c = 1$. If $\alpha_c^n \to \alpha_c$ as $n \to \infty$ for all c, then for all τ^{-i},

$$V^i(G^{(1),n}; \tau^{-i}) \to V^i(G^{(1)}; \tau^{-i})$$

[8] See Example 2.5.1 in Berry and Fristedt (1985).

as $n \to \infty$.

Proof . Note that

$$\begin{aligned} V^i(G^{(1),n};\tau^{-i}) &= \sup_{\tau^i} E_{\tau,G^{(1),n}}\left(\sum_{m=1}^{\infty}\alpha_m Z_m\right) \\ &= \sup_{\tau^i}\sum_c \alpha_c^n E_\tau\left(\sum_{m=1}^{\infty}\alpha_m Z_m|X_{c_1,1}^1,...,X_{c_n,1}^n\right), \end{aligned}$$

where the summations are taken over all possible $c \in \{1,2,...,k\}^n$. Likewise,

$$V^i(G^{(1)};\tau^{-i}) = \sup_{\tau^i}\sum_c \alpha_c E_\tau\left(\sum_{m=1}^{\infty}\alpha_m Z_m|X_{c_1,1}^1,...,X_{c_n,1}^n\right).$$

Hence, it is clear that $\alpha_c^n \to \alpha_c$ for all c implies that $V^i(G^{(1),n};\tau^{-i}) \to V^i(G^{(1)};\tau^{-i})$.

■

Now we turn to the set of strategies. Given $\tau = (\tau_1,\tau_2,...)$ and $\tau' = (\tau'_1,\tau'_2,...)$, recall that τ_m and τ'_m are R^k valued measurable functions. The distance between τ_m and τ'_m, $d_m(\tau_m,\tau'_m)$, is defined as

$$d_m(\tau_m,\tau'_m) = \max_{1\le j\le k}\int |\tau_{m,j}-\tau'_{m,j}|,$$

where $\tau_{m,j}$ and $\tau'_{m,j}$ are j-th coordinate of τ_m and τ'_m, respectively. Next, we will define a metric on $\mathcal{T}$. The distance between strategy profiles τ and τ', $d(\tau,\tau')$, is define as

$$d(\tau,\tau') = \sum_{m=1}^{\infty} 2^{-m} d_m(\tau_m,\tau'_m).$$

Therefore, $d(\tau^n,\tau) \to 0$ as $n \to 0$ is equivalent to L^1 convergence of each component of τ^n to the corresponding component of τ. If $\tau_m^n \to \tau_m$ in L^1, there exists a subsequence of τ_m^n which will converge to τ_m almost surely. Therefore, τ_m will also be a strategy at period m. From this observation, it is clear that $\mathcal{T}$ is a complete metric space. For strategy profiles $\tau, \hat{\tau} \in \mathcal{T}^{n-1}($ or $\mathcal{T}^n$), we will use the product metric arising from d on $\mathcal{T}$. For example, for $\tau = (\tau^1, ..., \tau^{n-1}), \hat{\tau} = (\hat{\tau}^1, ..., \hat{\tau}^{n-1}) \in \mathcal{T}^{n-1}$

$$d(\tau, \hat{\tau}) = \max_i d(\tau^i, \hat{\tau}^i).$$

Changes in τ^{-i} do not have direct effects on the instant payoff of player i at each stage, since τ^{-i} only determines the law by which player i will obtain extra information. Thus, we would conjecture that the value function of player i is a continuous function of τ^{-i}. The following lemma shows that this conjecture is true.

Lemma 4.8 *For fixed G, $V^i(G; \cdot)$ is a continuous function.*

Proof . Given $A = (\alpha_1, \alpha_2, ...)$, let $V^i(G; \tau^{-i}; A^{(n)})$ be defined as

$$V^i(G; \tau^{-i}; A^{(n)}) = \sup_{\tau^i} E_\tau^i(\sum_{m=1}^n \alpha_m Z_m^i).$$

We will first show that for all A, and for all $n < \infty$, $V^i(G; \tau^{-i}; A^{(n)})$ is continuous in τ^{-i} uniformly in G. We prove this by induction. For $n = 1$, the optimal strategy for player i is to select the myopic best choice. Thus,

$$\left|V^i(G; \tau^{-i}; A^{(1)}) - V^i(G; \hat{\tau}^{-i}; A^{(1)})\right| = 0 \text{ for all } G, \tau^{-i}, \hat{\tau}^{-i}, A.$$

Suppose that it is true for all $1 \leq k \leq n$, and for all A that $V^i(G; \tau^{-i}; A^{(k)})$ is continuous in τ^{-i} uniformly in G. Note that

$$\begin{aligned}
& \left|V^i(G;\tau^{-i};A^{(n+1)}) - V^i(G;\hat{\tau}^{-i};A^{(n+1)})\right| \\
= & \left| \begin{array}{c} \vee_{j=1}^k[\alpha_1 E(X_{j,1}^i) + E\left(V^i(G^{(1),j};\tau^{-i};A^{(n)})\right)] \\ - \vee_{j=1}^k\left[\alpha_1 E(X_{j,1}^i) + E\left(V^i(\hat{G}^{(1),j};\hat{\tau}^{-i};A^{(n)})\right)\right] \end{array} \right| \\
\leq & \left| \begin{array}{c} \vee_{j=1}^k[\alpha_1 E(X_{j,1}^i) + E\left(V^i(G^{(1),j};\tau^{-i};A^{(n)})\right)] \\ - \vee_{j=1}^k\left[\alpha_1 E(X_{j,1}^i) + E\left(V^i(G^{(1),j};\hat{\tau}^{-i};A^{(n)})\right)\right] \end{array} \right| \\
& + \left| \begin{array}{c} \vee_{j=1}^k[\alpha_1 E(X_{j,1}^i) + E\left(V^i(G^{(1),j};\hat{\tau}^{-i};A^{(n)})\right)] \\ - \vee_{j=1}^k\left[\alpha_1 E(X_{j,1}^i) + E\left(V^i(\hat{G}^{(1),j};\hat{\tau}^{-i};A^{(n)})\right)\right] \end{array} \right|.
\end{aligned}$$

The first term in the right hand side of the inequality will be arbitrarily small by induction hypothesis if $\hat{\tau}^{-i}$ is close enough to τ^{-i}. As $G^{(1),j}$ and $\hat{G}^{(1),j}$ are posteriors from the same prior G, we can apply Lemma 4.7, and hence, the last term will also be arbitrarily small if $\hat{\tau}^{-i}$ is close enough to τ^{-i}.

Now,

$$\begin{aligned}
& \left|V^i(G;\tau^{-i}) - V^i(G;\hat{\tau}^{-i})\right| \\
\leq & \left|V^i(G;\tau^{-i}) - V^i(G;\tau^{-i};A^{(n)})\right| \\
& + \left|V^i(G;\tau^{-i};A^{(n)}) - V^i(G;\hat{\tau}^{-i};A^{(n)})\right| \\
& + \left|V^i(G;\hat{\tau}^{-i};A^{(n)}) - V^i(G;\hat{\tau}^{-i})\right|.
\end{aligned}$$

As $V^i(G;\tau^{-i};A^{(n)}) \to V^i(G;\tau^{-i})$ uniformly in τ^{-i}, [9] by choosing n appropriately, the first and the last term can be made less than $\varepsilon/3$ for all $\varepsilon > 0$. Given that n and ε, by continuity of $V^i(G;\cdot;A^{(n)})$, the middle term will be less than $\varepsilon/3$ if $\hat{\tau}^{-i}$

9 See Theorem 2.5.1 in Berry and Fristedt (1985).

is close enough to τ^{-i}. ■

4.5 Existence of Pure Markov Strategy Equilibria

For $h_m \in \mathcal{H}_m$, let G^{h_m} be the posterior belief at the beginning of stage m. A strategy $\tau = (\tau_1, \tau_2, ...) \in \mathcal{T}$ is a *Markov strategy* if there exists $\tilde{\tau} : \mathcal{D}(\mathcal{D}^k) \to \Delta$ such that $\tau_m(h_m) = \tilde{\tau}(G^{h_m})$ for all $h_m \in \mathcal{H}_m$, and m. Thus, if player i is adopting a Markov strategy, then his selection will depend only on the posterior at that stage. As long as the posteriors are the same, regardless of the time he is going to make a choice and the history before the stage, his selection will be the same. A Markov strategy $\tilde{\tau}$ is a *pure Markov strategy* if $\tilde{\tau}$ maps $\mathcal{D}(\mathcal{D}^k)$ into vertices of Δ.

In this section, we will provide the main result of this chapter: There exist pure Markov strategy Equilibria in multi-player multi-armed bandit problems. In order to invoke the usual fixed point theorem, we will show that the set of pure Markov strategies is compact subset of the set of strategies. Then, it will be shown that the best response correspondence restricted on the set of pure Markov strategies is non-empty, convex, and upper semi-continuous. Recall that, given a sequence of sets $\{A_k\}$, $\limsup A_k$ and $\liminf A_k$ are defined as

$$\limsup A_k = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k$$

and

$$\liminf A_k = \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k.$$

We say that $A_k \to A$ if $\limsup A_k = \liminf A_k$

Lemma 4.9 $\mathcal{M}$ is a compact subset of $\mathcal{T}$.

Proof . Since $\mathcal{T}$ is a complete metric space, it suffices to show that $\mathcal{M}$ is sequentially compact. That is, it suffices to show that for every sequence $\{\tilde{\tau}^n\} \subset \mathcal{M}$, there exists a subsequence of $\{\tilde{\tau}^n\}$ that converges in $\mathcal{M}$. Note that a pure Markov strategy $\tilde{\tau}^n \in \mathcal{M}$ can be represented as

$$\tilde{\tau}^n = \sum_{j=1}^{k} e_j 1_{G_j^n},$$

where the e_j are unit vectors in R^k, and the G_j^n are the inverse images of e_j by $\tilde{\tau}^n$. By definition, $\cup G_j^n = \mathcal{D}(\mathcal{D}^k)$ and $G_i^n \cap G_j^n = \varnothing$ for $i \neq j$.

Let

$$G_1 = \limsup G_1^n$$

and select a subsequence $\{\tilde{\tau}^{n_l}\}$of $\{\tilde{\tau}^n\}$ such that $G_1^{n_l} \to G_1$ as $l \to \infty$. Given $\{\tilde{\tau}^{n_l}\}$, let's define G_2 as

$$G_2 = \limsup G_2^{n_l},$$

and select a subsequence $\{\tilde{\tau}^{n_{l_k}}\}$of $\{\tilde{\tau}^{n_l}\}$ such that $G_2^{n_{l_k}} \to G_2$ as $k \to \infty$. Repeat this process until we have subsequence $\{\tilde{\tau}^{n_q}\}$of $\{\tilde{\tau}^n\}$ such that $G_j^{n_q} \to G_j \subset \mathcal{D}(\mathcal{D}^k)$

for every j. It is clear that $G_i \cap G_j = \varnothing$ for $i \neq j$. Now let

$$\tilde{\tau} = \sum_{j=1}^{k} e_j 1_{G_j}.$$

Then, by construction, $\tilde{\tau}$ is a random variable. Since $\{G_j\}$ are disjoint, $\tilde{\tau}$ is a Markov strategy. ■

In the remainder, we will consider a modified game of the original multi-player multi-armed bandit problems. In this modified game, the set of strategies of the players are restricted to $\mathcal{M}$. Payoffs will be determined as in the original game. Let $\hat{B}^i$ be player i's best response in the modified game. We could prove that $\hat{B}^i$ is non-empty as in Lemma 4.5. The proof of Lemma 4.5 will go through without much change. As the strategies are stationary, it is obvious that $\hat{B}^i(\tau^{-i})$ is convex for every $\tau^{-i} \in \mathcal{M}^{n-1}$. We already showed that $\mathcal{M}$ is compact. What remains to be shown is the following. Recall that a correspondence $\Gamma : X \to 2^Y$ is called *upper semicontinuous* if $\{x : \Gamma(x) \subset W\}$ is open in X for every open $W \subset Y$.

Lemma 4.10 *Player i's best response correspondence $\hat{B}^i$ is upper semicontinuous.*

Proof . By Theorem 7.1.14 in Klein and Thompson (1984), it suffices to show that $\hat{B}^i$ is upper hemicontinuous and $\hat{B}^i(\tau^{-i})$ is compact for every $\tau^{-i} \in \mathcal{M}^{n-1}$.

We prove upper hemicontinuity first. Suppose that

$$\begin{aligned} \tau^{-i,n} &\to \tau^{-i} \\ \tau^{i,n} &\in \hat{B}^i(\tau^{-i,n}), \end{aligned}$$

and

$$\tau^{i,n} \to \tau^i.$$

We have to show that

$$\tau^i \in \hat{B}^i(\tau^{-i}),$$

and hence, it suffices to show that

$$V^i(G;\tau^{-i}) = E^i_{(\tau^i,\tau^{-i})}(\sum_{m=1}^{\infty} \alpha_m Z^i_m).$$

Since $\tau^{i,n} \in \hat{B}^i(\tau^{-i,n})$, we have

$$V^i(G;\tau^{-i,n}) = E^i_{(\tau^{i,n},\tau^{-i,n})}(\sum_{m=1}^{\infty} \alpha_m Z^i_m).$$

It is clear that $E^i_{(\tau^{i,n},\tau^{-i,n})}(\sum_{m=1}^{\infty} \alpha_m Z^i_m) \to E^i_{(\tau^i,\tau^{-i})}(\sum_{m=1}^{\infty} \alpha_m Z^i_m)$. Also, by Lemma 4.8, $V^i(G;\tau^{-i,n}) \to V^i(G;\tau^{-i})$. Thus, $\hat{B}^i(\tau^{-i})$ is upper hemicontinuous.

For compactness of $\hat{B}^i(\tau^{-i})$, as $\mathcal{M}$ is compact, we have only to show that $\hat{B}^i(\tau^{-i})$ is closed. Suppose that $\tau^{i,n} \in \hat{B}^i(\tau^{-i})$ and $\tau^{i,n} \to \tau^i$. It is immediate that $\tau^i \in \hat{B}^i(\tau^{-i})$ since

$$\begin{aligned} V^i(G;\tau^{-i}) &= E^i_{(\tau^{i,n},\tau^{-i})}(\sum_{m=1}^{\infty} \alpha_m Z^i_m) \\ &\to E^i_{(\tau^i,\tau^{-i})}(\sum_{m=1}^{\infty} \alpha_m Z^i_m). \end{aligned}$$

■

Theorem 4.11 *There exist pure Markov strategy equilibria in multi-player multi-armed bandit problems.*

Proof . Let

$$\hat{B} = \Pi_{i=1}^{n} \hat{B}^i : \mathcal{M}^n \to (2^{\mathcal{M}})^n.$$

For proof, it suffices to show that $\hat{B}$ has a fixed point. Since the $\hat{B}^i$ are non-empty, and convex, $\hat{B}$ will also be non-empty, and convex. It is clear that $\hat{B}$ is also upper semicontinuous, as $\hat{B}^i$ is upper semicontinuous. It is obvious that $\hat{B}$ is closed. As $\mathcal{M}^n$ is compact and convex, by Theorem 11.4 in Dugundji and Granas (1982), $\hat{B}$ has a fixed point. ■

Chapter 5

Conclusion

In this dissertation, we investigated various situations of strategic experimentation. All these models, however, share one common feature: Players can observe perfectly others' actions and payoffs. We should point out that there is some literature on strategic experimentation with different assumptions. Most well known is the literature on herding. In the study of herd behavior, it is assumed that people can observe what others are doing, but *not* their payoffs. Under this assumption, players have private information, and they can only infer information about the respective payoffs of others from observed actions.

We believe that our world is somewhere in between these two extreme cases. If so, we will have another dimension of strategic behavior, and a lot of intriguing questions. How much information should I collect? How much information should

I release? If I have some control over the quality of the released information, when should I release truth and when should I release forged information? All of these questions are waiting to be answered.

REFERENCES

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

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

3. Bergemann, D. and Välimäki, J. (1997): "Market Diffusion with Two-Sided Learning," *Rand Journal of Economics* **28**, 773-795.

4. Bergemann, D. and Välimäki, J. (2000): "Experimentation in Markets," *Review of Economic Studies* **67**, 213-234.

5. Berry, D. A. and Fristedt, B. (1985): *Bandit Problems.* New York: Chapman and Hall.

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

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

8. Dixit, A. (1993): *The Art of Smooth Pasting*, Vol. 55 in *Fundamentals of Pure and Applied Economics*, eds. Lesourne, J. and Sonnenschein, H. Chur, Switzerland: Harwood Academic Publishers.

9. Dixit, A. and Pindyck, S. (1994): *Investment Under Uncertainty.* Princeton, New Jersey: Princeton University Press.

10. Dugundji, J. and Granas, A. (1982): *Fixed Point Theory.* PWN-Polish Scientific Publishers.

11. Easley, D. and Kiefer, N. M. (1989): "Optimal Learning with Endogenous Data," *International Economic Review* **30**, 963-978.

12. Folland, G.B. (1984): *Real Analysis: Modern Techniques and Their Applications.* New York: John Wiley & Sons.

13. Gale, D. (1996): "What Have We Learned from Social Learning?," *European Economic Review* **40**, 617-628

14. Gittins, J. C. and Jones, D. M. (1974): "A Dynamic Allocation Index for the Sequential Design of Experiments," in *Progress in Statistics* (eds J. Gani *et al.*), 241-266, Amsterdam: North-Holland.

15. Harrison, J.M. (1985): *Brownian Motion and Stochastic Flow Systems.* Malabar: Krieger.

16. Klein, E. and Thompson, A.C. (1984): *Theory of Correspondences.* New York: John Wiley & Sons.

17. Krylov, N.V. (1980): *Controlled Diffusion Process.* Berlin: Springer-Verlag.

18. McFadden, D. and Train, K. (1996): "Consumer's evaluation of New Products: Learning from Self and Others," *Journal of Political Economy* **104**, 683-703.

19. Mirman, L.J., Samuelson, L., and Urbano, A. (1993): "Monopoly Experimentation," *International Economic Review* **34**, 549-563.

20. Parathasarathy, K.R. (1967): *Probability Measures on Metric Spaces.* New York: Academic Press.

21. Rothschild, M. (1974): "A Two-Armed Bandit Theory of Market Pricing," *Journal of Economic Theory* **9**, 185-202.

22. Sobel, J. (2000): "Economists' Model of Learning," *Journal of Economic Theory* **94**, 241-261.