Essays on Information Economics

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree
Doctor of Philosophy in the Graduate School of The Ohio State
University

By

Yeochang Yoon, B.A., M.A.

Graduate Program in Economics

The Ohio State University

2016

Dissertation Committee:

James D. Peck, Advisor
Paul J. Healy
Huanxing Yang
© Copyright by
Yeochang Yoon
2016
Abstract

In the first chapter, I develop a theoretical model to investigate why and how information senders are biased. In this paper, a rational Bayesian consumer decides whether or not to purchase a new product. His utility from purchasing depends on the quality of the product and his idiosyncratic preference for the product. Before making his decision, the consumer can receive a signal of the product’s quality by actively choosing an information sender. An information sender would like to attract more consumers by providing more accurate signals, but it is costly. In the paper, an information structure consists of a probability of recommending the product when the quality is high and that of not recommending it when the quality is low. An information sender’s bias is defined as the difference between the accuracy of the signal in the high quality state and that in the low quality state. On the other hand, the sender’s overall accuracy depends on the sum of the accuracy of the signals. A consumer does not have direct utility from biased information but it is shown that his expected utility from an information sender is increasing in the sender’s bias when the consumer subscribe to a like-minded information sender holding the sender’s overall accuracy constant. The indirect demand for information bias gives an incentive to the sender to be biased. As a result, no matter how many information senders are in the market, they have an incentive to be biased. Moreover, as more information
senders are potentially able to enter the market, overall accuracy is weakly increasing and bias is weakly decreasing due to competition effects.

In the second chapter, I explore a rational social learning model in which a consumer can observe other consumers’ ratings for a product and past purchase decisions. In this paper, I demonstrate how ratings work as an additional information source in a social learning model, and investigate whether or not additional ratings information improves learning. It is common in the social learning literature to model the history of the purchase as the observable, but in this paper, product ratings are also incorporated as an additional source of information; consumers can observe the history of both purchase decisions and ratings. Ratings provide additional information about how previous buyers felt about the product they purchased. I find that with ratings, low quality products are not chosen, but high quality products may also be ignored which causes incorrect herds in the long-run. Interestingly, the probability of high quality products being ignored is not monotonic in the accuracy of ratings. The more accurate information might lower the probability of making correct decisions in the long-run. When consumers believe that ratings are more accurate than private signals, they rely more on ratings and ignore other information sources. In that case, a small number of negative ratings can prevent later buyers from purchasing high quality products. Moreover, the probability of incorrect herds with high quality products might be higher in an environment with ratings, even when the ratings are more accurate than private signals. These results imply that the additional information source does not always guarantee better learning.

Finally, the third chapter studies the case in which observing others actions is noisy in diffusion of innovations. I consider a three stage game where only finite
number of early adopters receive private signals about the state of nature and make their decision about whether to adopt the new technology in the first stage. In the second stage, all remaining decision makers observe the early adopters’ decisions, but the observation is noisy. I show how the noisy observation affects the dynamics in diffusion of innovations. The dynamics are related to inefficiency of the model because it is desirable that the majority adopts the new technology earlier when its quality is good. The inefficiency tends to decrease as agents can observe others’ actions with higher probability, however, it might locally increase in some areas.
For Danbee
Acknowledgments

I would like to thank James D. Peck for invaluable guidance and support. I am also thankful to Paul J. Healy and Huanxing Yang for helpful advice and feedback. This work benefited substantially from the comments and assistance of Yaron Azrieli, Lixin Ye, Dan Levin, Lucas Coffman, Semin Kim, Daeyoung Jeong, John Conlon, and the participants of the microeconomics brownbag seminar, theory seminar, and theory/experimental colloquium at Ohio State, as well as the participants of Midwest Economic Theory and Trade Meetings 2015, International Conference on Game Theory 2014, and the Midwest Economics Association Annual Meeting 2014. I gratefully acknowledge research funding from the JMCB Grants for Graduate Student Research Program.

Finally, I am thankful to my family, especially my wife, Danbee, and my parents. My brother, Yeohwan, deserves my wholehearted thanks as well. Without their love and support this dissertation my have never been completed.
Vita

2008 ................................. B.A. Economics, Yonsei University

2011 ................................. M.A. Economics, The Ohio State University

2011 to present ..................... Graduate Teaching Associate, Department of Economics, The Ohio State University

Fields of Study

Major Field: Economics
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>v</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>vi</td>
</tr>
<tr>
<td>Vita</td>
<td>vii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>x</td>
</tr>
</tbody>
</table>

**Chapters**

1. **Why Are Information Senders Biased? Indirect Demand for Information Bias** ........................................ 1
   1.1 Introduction ........................................................................ 1
   1.2 Related Literature ................................................................ 4
   1.3 Model .................................................................................. 7
      1.3.1 Decision Maker’s payoff-relevant action ............................ 7
      1.3.2 Information Structure .................................................. 7
      1.3.3 Information Choice ........................................................ 8
      1.3.4 Timing .......................................................................... 9
      1.3.5 Bias and Overall Accuracy ............................................. 10
   1.4 Analysis .............................................................................. 11
      1.4.1 Posterior Beliefs .......................................................... 11
      1.4.2 Threshold Preference .................................................... 11
      1.4.3 Indirect Demand for Informational Bias ............................. 12
   1.5 Equilibrium .......................................................................... 13
      1.5.1 Monopoly ...................................................................... 13

viii
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Equilibrium Reporting Strategy in Monopoly</td>
<td>16</td>
</tr>
<tr>
<td>1.2 Overall Accuracy and Bias in Monopoly</td>
<td>17</td>
</tr>
<tr>
<td>1.3 Relationship between Overall Accuracy and Bias</td>
<td>26</td>
</tr>
<tr>
<td>1.4 Overall Accuracy and Bias over Market Structures</td>
<td>26</td>
</tr>
<tr>
<td>3.1 Example of Optimal Choices for High Signal Agents</td>
<td>61</td>
</tr>
<tr>
<td>3.2 Example of the Cutoff-Observed Adoption Changed By $\pi$</td>
<td>62</td>
</tr>
<tr>
<td>3.3 Example of the Expected Fraction who Adopt the New Technology at Stage 2 Changed By $\pi$</td>
<td>63</td>
</tr>
<tr>
<td>3.4 Binomial Distributions when $\pi = 0.4$ and $\pi = 0.5$ and the Corresponding Cutoffs</td>
<td>64</td>
</tr>
<tr>
<td>3.5 Binomial Distributions when $\pi = 0.5$ and $\pi = 0.6$ and the Corresponding Cutoffs</td>
<td>64</td>
</tr>
</tbody>
</table>
Chapter 1: Why Are Information Senders Biased? Indirect Demand for Information Bias

1.1 Introduction

In many cases, consumers are active information receivers. Information sources are not simply given to them, but they choose information senders such as consumer reports, magazines, and consulting firms when they need to collect information about the product they would like to purchase. However, it is not guaranteed that the information sources are unbiased. Dewenter and Heimeshoff (2014) analyze automobile reviews in German car magazines and show the evidence of biased ratings. Reuter (2009) also finds the consistent evidence of bias in wine ratings by comparing major wine publications. Still, consumers seem to demand and rely on the information senders in spite of their information bias. This paper builds a theoretical model to investigate why information bias is desired by rational Bayesian consumers and how the bias is endogenously determined by information senders.

In the formal setting, a brand new car is introduced and its quality depends on states of nature that are binary, high or low, but unknown. The game has three stages. First, an information sender determines her information structure that consists of accuracy of signals conditional on each state. The true state is not observable to the
sender, thus she invests her resources such as journalists or experts to investigate the quality of the product. The sender can attract more consumers by providing more accurate signals, but this is more costly. In the second stage, a consumer observes all information structures in the market and chooses which information sender he receives a signal from. Last, after receiving a signal from the sender, the consumer decides whether or not to purchase the new car. This is his payoff-relevant action. His payoff depends on the quality of the product and on his idiosyncratic term regarding the product.

We define an information sender’s bias as the difference between the accuracy of the signal when the quality is high and the accuracy of the signal when the quality is low.\(^1\) As the bias increases, the information sender treats the high quality state and the low quality state differently. When the information sender has a positive bias towards a new product, she is more likely to recommend the product, and when the sender has a negative bias, she is less likely to recommend it. On the other hand, the sender’s overall accuracy depends on the sum of the accuracy of the signals. Higher overall accuracy implies that it is more likely for a decision maker to receive a correct signal from the sender.

In the model, for a consumer, making a correct decision is more important than his own preference for the product. Hence, the consumer prefers an information sender with higher overall accuracy. Holding the sender’s overall accuracy, however, it is shown that the expected utility of subscribing to an information sender increases

\(^1\)In this paper, the fact that a car magazine tends to favor one car manufacturer does not mean that the information sender withholds evidence about the quality of the car. If the sender withholds the evidence, the sender’s information structure is distorted from the original and it is dominated by the original for rational Bayesian consumers. We do not consider this type of distortion bias but the case of filtering bias in this paper. In the case of filtering bias, the sender provides the different accuracy of signals for the different states. We will discuss the detail in Chapter 2.
with the sender’s bias if the sender’s bias has the same sign with the consumer’s idiosyncratic term. When the sender is biased towards a direction that matches a consumer’s preference for the product, the consumer is more likely to receive a signal that increases their expected utility. This implies that the consumer has indirect demand for informational bias. The intuition is that, for a consumer with a positive idiosyncratic preference for the product and who adopts the sender’s recommendation, a mistake to recommend the product is less costly than a mistake to recommend against purchasing the product. Hence, the consumer tends to choose a like-minded information sender and this gives an incentive to information senders to be biased.

In this paper, we consider monopoly, duopoly and multiple senders’ case. The equilibrium reporting strategy for an information sender depends on the market structure. However, regardless of the different market structure, one common result of all market structures is that information senders have an incentive to be biased. This result holds true even in a monopoly; the monopolistic sender is biased towards one of two states. Since consumers tend to choose a like-minded information sender, the information sender can easily attract more consumers with low accuracy cost by taking high bias. In duopoly, two senders are biased in exactly opposite directions. In the multiple senders’ case, the market is simply segmented into two parts. All information senders provide the same overall accuracy, and a half of them have the same positive bias and another half of them have the same negative bias.

We compare how overall accuracy and bias change as more information senders potentially enter the market. Overall accuracy is weakly increased with the number of senders in the market through the competition effect. Hence, the probability of
making a correct decision for a consumer also increases. In addition, bias is weakly reduced with the number of potential competitors.

The remainder of the paper is organized as follows. In Section 1.2 we review the related literature. Section 1.3 describes the basic model setting and defines bias and overall accuracy. Section 1.4 explains how decision makers update their posterior beliefs given information structures and how they make subscription and action choices. In Section 1.5, the optimal reporting strategies for information senders are presented in various market structures. Section 1.6 examines how overall accuracy and bias change over different market structures. In Section 1.7, the possibility of cross-checking is introduced, and Section 1.8 concludes the paper.

1.2 Related Literature

This paper is topically related to studies on media bias because the role of media firms is mainly considered as an information sender in the literature. The media market is widely regarded as biased by the public. According to the Pew Research Center (2012), 75% of people said that media tend to favor one side, which is compared with 53% who said so in 1985. This number has gradually increased. Empirical evidence also supports rising bias in the news media market (Groseclose and Milyo 2005, Lott and Hassett 2004). Furthermore, media bias exists not only in political situations but also in consumer’s problems. (See Dewenter and Heimeshoff 2014 and Reuter 2009)

Most of theoretical studies on media bias, however, assume that agents only receive utility from their own choice and not from the aggregate choice, which fits better the example of a commodity rather than a political contest or a public policy.
Media bias can be defined in various ways. Gentzkow, Shapiro and Stone (2014) categorize bias into two different types: distortion bias and filtering bias. In the case of distortion bias, information senders intentionally omit certain information or put a slant on a story. When an information sender distorts the original reporting strategy, the distorted one is Blackwell less informative than the original. In the case of filtering bias, information senders provide different accuracy of signals for different quality states, and this does not generate the order of two different reporting strategies by Blackwell informativeness. In our paper, we consider rational Bayesian decision makers for whom making correct decisions is important. For them, the distorted information structure is dominated by the original information structure and thus, they never choose the distorted one. Hence, we do not use the concept of distortion but that of filtering to define bias.

The reason for media bias has been theoretically studied in various aspects. First, informational bias could be directly driven by information senders. According to Baron (2006), journalists can easily publish biased reports on the front page, which can advance their careers. Chan and Suen (2009) study the political voting situation when media outlets are biased towards the incumbent or the challenger, and how the biased media affects the party’s policy. Anderson and McLaren (2012) show that media owners who have political and profit motives can manipulate political outcomes by distorting information. The second possible reason of media bias is the government. Gehlbach and Sonin (2014) show that mobilizing governments increases media bias but that it can be reduced by competition in the media market. Last, information receivers could be the reason of media bias. In Mullainathan and Shleifer (2005), biased consumers get direct utility from receiving biased information, which
presents an incentive to media firms to be biased. Bernhardt et al. (2008) develop a model in which partisan consumers’ entertainment utility from watching biased media increases with their watching time. These papers consider the case where decision makers receive direct utility from watching biased information. On the other hand, Gentzkow and Shapiro (2006) consider consumers who are interested in receiving accurate information without direct utility from biased information, and they show that media firms distort information to increase their reputation among heterogeneous consumers. In their paper, a media firm’s information structure is unknown and consumers intertemporally learn it, which, however, is incomplete. In our paper, on the other hand, the information structure is known in advance among decision makers. Even when they know how information senders are biased, decision makers still have an incentive to choose a biased sender.

Burke (2008) is the study most related to this paper; he also investigates how media firms are biased with rational Bayesian decision makers. However, our work mainly differs from his paper in two important aspects. First, in his work, the overall accuracy of information structure is exogenously given and information senders only determine the difference in the accuracy of the signals. In this paper, however, information senders endogenously determine overall accuracy and bias together, and more accurate signals are costly. With this feature of our paper, we can analyze how overall accuracy and bias are related to each other and how they change with market structures and an increase in the accuracy cost. In his paper, the competition may not decrease bias while in our paper, bias is weakly decreased as more information senders appear in the market. Second, his model considers decision makers who have homogeneous preferences but heterogeneous prior beliefs while we study
decision makers who have heterogeneous preferences but common prior beliefs. This difference allows us to have a simple expected utility form of a decision maker from subscribing to an information sender.

1.3 Model

1.3.1 Decision Maker’s payoff-relevant action

A new product is introduced and its quality depends on states of nature. Suppose that there are binary states, \( \omega \in \{-1, 1\} \), and they are equally likely.\(^3\) Consider a continuum of decision makers who make a payoff-relevant choice, \( x \in \{-1, 1\} \).\(^4\) If they make a correct decision, i.e., \( x = \omega \), then they receive \( u > 0 \) and otherwise \(-u\). They also have idiosyncratic term of the product, \( v \), for the product which is uniformly distributed, \( v \sim U[-u, u] \). Note that since \( v \) is between \(-u\) and \( u \), making a correct decision is more important than their idiosyncratic term. In sum, a decision maker’s payoff is denoted by

\[
u(x|\omega, v) = (u \omega + v)x, \tag{1.1}\]

which combines a payoff from making a correct decision and that from matching a choice with their preference.

1.3.2 Information Structure

Consider a finite number of information senders, \( k \in \{1, \cdots, K\} \). They send a binary signal, \( s \in \{-1, 1\} \) to decision makers.\(^5\) Before decision makers choose

\(^3\)We can consider that in \( \omega = 1 \) the new product is better while in \( \omega = -1 \) the old one is better.

\(^4\)If a decision maker chooses \( x = 1 \), it implies adoption of the new product and \( x = -1 \) is to keep the old one.

\(^5\)Again, \( s = 1 \) implies that the sender recommends consumers to buy the new product while \( s = -1 \) implies that the sender recommends consumers to keep the old product.
them, information senders determine how accurately they provide information for
each state respectively, $Pr_k(s = 1|\omega = 1) = \mu_k$ and $Pr_k(s = -1|\omega = -1) = \lambda_k$, where $\mu_k, \lambda_k \in [\frac{1}{2}, 1]$. The fact that a sender provides a more accurate signal for a state can be interpreted as that the sender collects more clues to send a signal for the state.\(^6\)

Given information structures, decision makers choose which information sender they receive a signal from, $y \in \{0, 1, \cdots, K\}$. A decision maker is allowed to choose $y = 0$ that means not subscribing to any information providers.

For an information sender, her revenue is her fraction of subscribers, i.e., how many decision makers choose her, $\int I\{y = k\}$. Moreover, providing more accurate signals is costly, $c(\mu_k, \lambda_k)$. In this paper, we assume a specific cost function, $c(\mu_k, \lambda_k) = \gamma[\frac{1}{2} - (\mu_k + \lambda_k)] - 1$, where $\gamma$ is a cost parameter. With this cost function, providing perfectly inaccurate signals for both states, $\mu_k = \lambda_k = \frac{1}{2}$, costs zero while providing perfectly accurate signals, $\mu_k = \lambda_k = 1$, is infinitely costly. This cost function is increasing and convex in sum of accuracy of signals for each state, $\mu_k + \lambda_k$. For instance, information senders need to hire more journalists in order to investigate which state is correct and it must be costly. Higher number of journalists means that it is more likely that the sender provides a correct signal, but in order to provide perfectly accurate signals, they need infinite number of journalists.

1.3.3 Information Choice

Each decision maker chooses which information sender he receives a signal from, $y \in \{0, 1, \cdots, K\}$. A decision maker is limited to choose only one information sender, $y = 0$ that means not subscribing to any information providers.

\(^6\)If $\mu_k$ or $\lambda_k$ is smaller than $\frac{1}{2}$, it means that information providers send an opposite signal, i.e., $Pr_k(s = -1|\omega = 1) = 1 - \mu_k > 1/2$ or $Pr_k(s = 1|\omega = -1) = 1 - \lambda_k > 1/2$. This eventually has the same interpretation with the above, so we restrict $\mu_k$ and $\lambda_k$ between $[\frac{1}{2}, 1]$. 

8
but this assumption will be relaxed in Chapter 7. In this paper, we assume that receiving a signal is slightly costly for a decision maker.

**Assumption 1.** A decision maker’s subscription choice, $y$, incurs a cost $\epsilon$ to receive a signal from an information sender $y \in \{1, \cdots, K\}$, where $\epsilon > 0$ is sufficiently close to zero.\textsuperscript{7}

Hence, when his expected utility from subscribing to any information senders is the same with that from not subscribing, he does not have an incentive to subscribe to any information senders. If a decision maker does not strictly benefit from any information senders, he will not subscribe to any information senders. In reality, people with extreme preference still consume media even for the entertainment reasons. With this assumption, however, we can avoid the situation in which the extreme decision makers consume the media which is biased towards the opposite of their favorites.

### 1.3.4 Timing

We are now ready to summarize the timing of the model.

1. First, nature selects the state, $\omega \in \{-1, 1\}$, and decision makers’ idiosyncratic term of the product, $v \sim U[-u, u]$.

2. Without knowing the true state, an information sender determines her information structure to maximize her expected profits,

$$\{\mu_k^*, \lambda_k^*\} \in \arg\max_{\{\mu_k, \lambda_k\}} \left[ \int I\{y(v) = k\} dv - c(\mu_k, \lambda_k) \right]. \quad (1.2)$$

\textsuperscript{7}Even if we consider the fixed price of the signals, results would be qualitatively unchanged. The only difference is that the larger set of consumers would not subscribe to any information senders. But for simplicity, we assume the price is small enough to be ignored but still positive.
Next, observing the information structures, \(\{\mu_k, \lambda_k\}\) and his own idiosyncratic term, a decision maker chooses which information provider he subscribes to,

\[
g^*(v) \in \arg\max_{y \in \{0, 1, \ldots, K\}} Eu(y|v). \tag{1.3}
\]

After then, a decision maker receives his signal, \(s \in \{-1, 1\}\), from the sender and makes a payoff-relevant action,

\[
x^*(v, k, s) \in \arg\max_{x \in \{-1, 1\}} Eu(x|v, k). \tag{1.4}
\]

### 1.3.5 Bias and Overall Accuracy

We define an information sender’s bias as the difference between the accuracy of signals for each state, \(\beta_k\), where \(\beta_k = \mu_k - \lambda_k\) and \(\beta_k \in [-\frac{1}{2}, \frac{1}{2}]\). We say that the information sender \(k\) is biased towards the state 1 (-1) if \(\beta_k\) is positive (negative). It implies that if an information sender is biased towards the state 1, \(\beta_k > 0\), (ex-ante) probability of sending the signal 1 is greater than that of sending the signal -1; \(Pr_k(s = 1) > Pr_k(s = -1)\).\(^8\) Thus, agents are more likely to receive the signal 1 from the sender. When \(\beta_k\) is zero, we say that the information sender is unbiased.

On the other hand, we also characterize the sender’s overall accuracy depending on the sum of the accuracy of signals for both states, \(\alpha_k = \mu_k + \lambda_k - 1\), where \(\alpha_k \in [0, 1]\). Note that the bias and the overall accuracy are now distinguished with each other, i.e., the lower overall accuracy does not directly imply the lower bias. More specifically, when \(\alpha_k < 1/2\), the overall accuracy and the bias have the positive relationship; the higher overall accuracy implies the higher bias. When \(\alpha_k > 1/2\), however, it is opposite; the higher overall accuracy implies the lower bias.

\(^8\)The probability that the sender \(k\) sends the signal 1 is, \(Pr_k(s = 1) = \sum_{\omega} Pr(\omega)Pr_k(s = 1|\omega) = \frac{1}{2} + \frac{1}{2}\beta_k\).
Definition 1.1. [Overall Accuracy and Bias] For an information sender $k$,

1. overall accuracy depends on the sum of the accuracy for both states, $\alpha_k = \mu_k + \lambda_k - 1$.
2. bias is the difference between the accuracy for each state, $\beta_k = \mu_k - \lambda_k$.

1.4 Analysis

1.4.1 Posterior Beliefs

As mentioned, two states are equally likely and the common prior belief is $1/2$. Suppose that a decision maker receives a signal from an information sender $k$ whose information structure is $\{\mu_k, \lambda_k\}$. Given the information structure, Bayes’ rule provides the updating rule for the posterior belief that state is $1$ after receiving signal $1$ and $-1$, respectively, which is the following,

$$Pr_k(\omega = 1|s = 1) = \frac{\mu_k}{\mu_k + (1 - \lambda_k)} \quad \text{and} \quad Pr_k(\omega = 1|s = -1) = \frac{1 - \mu_k}{(1 - \mu_k) + \lambda_k}. \quad (1.5)$$

Since the information structure does not have to be symmetric, we have the above posterior beliefs. If it is symmetric, $\mu_k = \lambda_k$, then, we simply have $Pr_k(\omega = 1|s = 1) = \mu_k = \lambda_k$ and $Pr_k(\omega = 1|s = -1) = 1 - \mu_k = 1 - \lambda_k$.

1.4.2 Threshold Preference

Based on the posterior beliefs, we obtain a decision maker’s expected payoff from making a payoff-relevant action. Hence, given the information structure, we can find the region between thresholds, $[\bar{v}_k, \bar{v}_k]$: a decision maker within the thresholds follows his signal while out of the range, he does not follow it.
Lemma 1.1. Given the information structure $k$, the optimal action for a decision maker within the thresholds, $[v_k, \bar{v}_k]$, is following his signal. When his preference $v$ is higher than $\bar{v}_k$, he chooses 1 for any signal he receives. When his preference $v$ is lower than $v_k$, he chooses $-1$ for any signal he receives.

$$x^*(v, k, s) = \begin{cases} 1 & \text{if } v > \bar{v}_k \\ -1 & \text{if } v < v_k \\ s & \text{if } v \in [v_k, \bar{v}_k] \end{cases} \quad (1.6)$$

where $\bar{v}_k = \frac{\mu_k - (1 - \lambda_k)}{1 - \mu_k + \lambda_k} u$ and $v_k = \frac{1 - \mu_k - \lambda_k}{\mu_k + (1 - \lambda_k)} u$.

1.4.3 Indirect Demand for Informational Bias

From the assumption 1, only decision makers within $[v_k, \bar{v}_k]$ have an incentive to subscribe to information $k$ and those out of the range do not. Suppose that a decision maker has an incentive to subscribe to the information sender $k$, $v \in [v_k, \bar{v}_k]$. Then, the following lemma shows his expected utility from subscribing to $k$.

Lemma 1.2. Given the information structure $k$ and a decision maker’s preference $v \in [v_k, \bar{v}_k]$, his expected utility from subscribing to $k$ is

$$Eu(y = k|v) = (\mu_k + \lambda_k - 1)u + (\mu_k - \lambda_k)v = \alpha_k u + \beta_k v \quad (1.7)$$

Proof. Attached in Appendix.

This lemma implies that a decision maker’s expected utility from his subscription choice is increasing in the overall accuracy of the information structure $k$ as well as increasing in the bias if the sender’s bias has the same sign with the decision

$^9$If the information sender $k$ provides perfectly accurate signals for both states, $\mu_k = \lambda_k = 1$, then $\bar{v}_k = u$ and $v_k = -u$, which implies that all decision makers will follow a signal from $k$, $x^*(v, k, s) = s$ for $\forall v$. Otherwise, $-u < v_k < \bar{v}_k < u$. 

12
maker’s idiosyncratic preference. Hence, the bias and the overall accuracy could be substitutes with each other for a decision maker. When an information structure $k$ is biased towards the state 1, it is more likely for the decision maker to receive a signal 1 from $k$. The decision maker knows this fact in advance but he still updates his posterior belief for the state 1 after receiving $s = 1$ from $k$. For $v > 0$, the higher $\beta_k$ increases his expected utility from the sender $k$. This is because for $v > 0$, his utility from $x = 1$ is higher in $\omega = 1$ than that from $x = -1$ in $\omega = -1$, he wants to receive more accurate signal for the state 1. Hence, a decision maker has indirect demand for information bias which was not directly in his utility function.

When a decision maker’s preference $v$ is out of the range $[\bar{v}_k, \underline{v}_k]$ for information structure $k$, his expected utility from subscribing to $k$ is simply the same with not subscribing, i.e., $Eu(y = k|v) = u$, and thus he has no incentive to subscribe to $k$. As a result, given $\{\mu_k, \lambda_k\}$ for $k = 1, \cdots, K$, decision makers with $v \notin [\min v_k, \max \bar{v}_k]$ do not subscribe to any information senders.

**Lemma 1.3.** For $v \notin [\min_k v_k, \max_k \bar{v}_k]$, $y^*(v) = 0$.

In addition, the thresholds $\bar{v}_k$ and $\underline{v}_k$ are increasing with the bias $\beta_k$, which means that more decision makers with positive preference have incentive to subscribe to $k$ while less decision makers with negative preference do not.

### 1.5 Equilibrium

#### 1.5.1 Monopoly

Now, consider the equilibrium reporting strategy for an information sender. Suppose that there is only one potential information sender in the market who determines how accurately she provides a signal to a decision maker in order to maximize her
profits. The sender is active in the market as long as her profit is positive. Since decision makers only with \( v \in [\underline{v}, \bar{v}] \) subscribe to the sender, her expected profit will be,

\[
\max_{\mu, \lambda} [U(\bar{v}) - U(v)] - c(\mu, \lambda) = \frac{1}{2} \left[ \frac{\alpha}{1 - \beta} + \frac{\alpha}{1 + \beta} \right] - \gamma \left[ \frac{1}{1 - \alpha} - 1 \right] \quad (1.8)
\]

**Lemma 1.4.** In monopoly, the unbiased information structure, \( \beta = 0 \), cannot be equilibrium if \( \gamma > 0 \).

**Proof.** Consider the overall accuracy and the bias as control variables instead of \( \mu \) and \( \lambda \). Then the first order condition for \( \beta \) indicates \( \frac{\alpha}{2} \frac{1}{(1-\beta)^2} - \frac{1}{(1+\beta)^2} \) which is zero when \( \beta = 0 \). The second order condition, however, is positive, which implies that unbiasedness minimizes the expected profit for the monopolistic sender and it is increasing as the sender becomes more biased.

The above lemma shows that even the monopolistic sender always has an incentive to be randomly biased towards either direction. This is because bias and overall accuracy are substitutes for decision makers with each other. Hence, if the sender takes the higher bias, she can easily attract more decision makers with the same overall accuracy and the same cost. Therefore, even though the sender cares only about her profit, she will not be unbiased. Since all settings are symmetric, the sender is randomly biased towards one of the states. The state which the information sender is biased towards is called her *primary*. For example, if \( \beta_k > 0 \), then the sender’s primary is the state 1 and if \( \beta_k < 0 \), then the sender’s primary is the state \(-1\).

**Proposition 1.1.** In monopoly, the equilibrium reporting strategy is the following.

\[^{10}\text{Since } \bar{v} = \frac{\mu - (1 - \lambda)}{(1 - \mu) + \lambda} u \text{ and } \underline{v} = \frac{1 - \mu - \lambda}{\mu + (1 - \lambda)} u, \text{ we can rewrite them by } \bar{v} = \frac{\alpha}{1 - \beta} u \text{ and } \underline{v} = \frac{-\alpha}{1 + \beta} u.]^{10}
1. For \( \gamma \in (0, \frac{1}{9}) \), the sender provides a perfect signal for her primary while accuracy for the other is decreasing in information cost, \( \gamma \).

2. For \( \gamma \in (\frac{1}{9}, \frac{5}{9}) \), the sender provides a perfect signal for her primary and perfectly inaccurate signal for the other even when \( \gamma \) increases.

3. For \( \gamma \in (\frac{5}{9}, 1) \), the sender reduces accuracy for her primary and provides perfectly inaccurate signal for the other.

4. For \( \gamma > 1 \), the sender is not active and provides perfectly inaccurate signals for both.

Figure 1.1 represents the case where the sender is biased towards the state 1; for \( \gamma \in (0, \frac{1}{9}) \), \( \mu^* = 1 \) and \( \lambda^* = \frac{1-\sqrt{\gamma}}{1-\sqrt{\frac{1}{9}}} \), and for \( \gamma \in (\frac{5}{9}, 1) \), \( \mu^* = \frac{2}{1+\sqrt{2}\gamma-1} - \frac{1}{2} \) and \( \lambda^* = \frac{1}{2} \).

In equilibrium, when the accuracy cost is sufficiently low, she provides a perfectly accurate signal for her primary while reduces the accuracy of the signal conditional on the other state as the cost increases. However, when the cost parameter \( \gamma \) reaches at \( \frac{1}{9} \), the sender cannot reduce the accuracy of signal which she has decreased to \( 1/2 \). In this case, if the sender decreases the accuracy of signal for her primary, then the bias and the overall accuracy are reduced together, which makes the information structure less attractive to all decision makers. In the region, \( \gamma \in (\frac{1}{9}, \frac{5}{9}) \), the sender has no sufficient incentive to lower both the overall accuracy and the bias. A decrease in her market share is greater than a decrease in the accuracy cost by reducing the overall accuracy and the bias. Thus, she provides a perfect signal for her primary and perfectly inaccurate signal for the other even when the accuracy cost increases. The region is defined as the extreme bias region. When \( \gamma \) is sufficiently high, however, it is better for the sender to reduce both of them. Eventually, if \( \gamma \) is greater than 1, the
sender provides uninformative signals, \( \{\mu^*, \lambda^*\} = \{ \frac{1}{2}, \frac{1}{2} \} \), and her share and her cost are zero, and we can say that the sender is inactive.

**Definition 1.2. [Extreme Bias Region]** When an information sender provides a perfect signal for her primary and a perfectly uninformative signal for the other, it is called that the sender’s information structure is in the *extreme bias region*.

From the equilibrium reporting strategy for the monopolistic sender, we can derive how the bias and the overall accuracy change over the accuracy cost, \( \gamma \). In monopoly, the equilibrium the overall accuracy is weakly decreasing as the accuracy cost increases. The equilibrium bias, however, is increasing before the extreme bias region and decreasing after the region as the cost increases. In the extreme bias region, the bias and the overall accuracy stay at the same even with changes in the cost.

**Corollary 1.1.** In monopoly, the equilibrium overall accuracy and bias are the following.
1. For $\gamma \in (0, \frac{1}{9})$, the overall accuracy is increasing while the bias is decreasing as the accuracy cost, $\gamma$, increases; $\alpha^* = \frac{1-2\sqrt{\gamma}}{1-\sqrt{\gamma}}$ and $|\beta^*| = 1 - \frac{1-2\sqrt{\gamma}}{1-\sqrt{\gamma}}$.

2. For $\gamma \in (\frac{1}{9}, \frac{5}{9})$, the overall accuracy and the bias are constant; $\alpha^* = |\beta^*| = \frac{1}{2}$.

3. For $\gamma \in (\frac{5}{9}, 1)$, the overall accuracy and the bias are decreasing in $\gamma$; $\alpha^* = |\beta^*| = \frac{2}{1+\sqrt{2\gamma}-1} - 1$.

![Graph showing Overall Accuracy and Bias in Monopoly](image)

Figure 1.2: Overall Accuracy and Bias in Monopoly

### 1.5.2 Duopoly

Suppose that there are two potential information senders in the market who provide $\{\mu_k, \lambda_k\}$ for $k = 1, 2$. Since we consider rational Bayesian decision makers, if an information structure is Blackwell more informative than the other, then no one chooses the less informative one. Without loss of generality, we can consider $\beta_1 \geq \beta_2$.

Then, we find a cutoff agent with $\bar{v}_{12}$ who is indifferent between subscribing to the sender 1 and the sender 2, where

$$\bar{v}_{12} = \frac{(\mu_2 + \lambda_2) - (\mu_1 + \lambda_1)}{(\mu_1 - \lambda_1) - (\mu_2 - \lambda_2)} u = \frac{\alpha_2 - \alpha_1}{\beta_1 - \beta_2} u.$$ 

(1.9)
When a decision maker’s preference is higher than \( \bar{v}_{12} \), he prefers the sender 1 and otherwise, prefers the sender 2. Hence, the market is now segmented into two parts; the sender 1’s market share \( S_1 \) is \( U(\bar{v}_1) - U(\bar{v}_{12}) \) and the sender 2’s share \( S_2 \) is \( U(\bar{v}_{12}) - U(\bar{v}_2) \). Their maximization problems are the following.

For sender 1, \( \max_{\mu_1, \lambda_1} [U(\bar{v}_1) - U(\bar{v}_{12})] - c(\mu_1, \lambda_1) = \frac{1}{2} \left[ \frac{\alpha_1}{1 - \beta_1} - \frac{\alpha_2 - \alpha_1}{\beta_1 - \beta_2} \right] - \gamma \left[ \frac{1}{1 - \alpha_1} - 1 \right] \).

For sender 2, \( \max_{\mu_2, \lambda_2} [U(\bar{v}_{12}) - U(\bar{v}_2)] - c(\mu_2, \lambda_2) = \frac{1}{2} \left[ \frac{\alpha_2 - \alpha_1}{\beta_1 - \beta_2} - \frac{-\alpha_2}{1 + \beta_2} \right] - \gamma \left[ \frac{1}{1 - \alpha_2} - 1 \right] \).

In duopoly, the senders are in the strategic situation and they compete with each other to take more market shares from the competitor.

**Lemma 1.5.** In duopoly,

1. the same bias, \( \beta_1 = \beta_2 \), cannot be equilibrium for the senders.

2. it cannot be equilibrium that one sender is unbiased while the other one is biased.

3. both senders will provide the same overall accuracy, \( \alpha^*_1 = \alpha^*_2 \).

**Proof.** Attached in Appendix.

The above lemma implies that duopolistic senders have an incentive to be biased towards the different directions. If they are biased towards the same direction, one of them can increase her market share by moving to the opposite direction. In addition, if one of them has the higher overall accuracy, she can increase her market share by taking the same bias direction with the competitor and dominating the competitor. Hence, both senders will provide the same overall accuracy not to be dominated by the competitor.
Proposition 1.2. In duopoly, the equilibrium reporting strategy for each sender is the following.

1. For $\gamma \in (0, \frac{1}{8})$, each sender oppositely provides a perfect signal for her own primary and reduces the accuracy for the other state; $\{\mu^*_1, \lambda^*_1\} = \{1, 1 - 4\gamma\}$ and $\{\mu^*_2, \lambda^*_2\} = \{1 - 4\gamma, 1\}$.

2. For $\gamma \in (\frac{1}{8}, \frac{1}{2})$, each sender oppositely focuses on her own primary and stays at the extreme bias region; $\{\mu^*_1, \lambda^*_1\} = \{1, \frac{1}{2}\}$ and $\{\mu^*_2, \lambda^*_2\} = \{\frac{1}{2}, 1\}$.

3. For $\gamma \in (\frac{1}{2}, \frac{5}{9})$, one of senders stays out and the other sender stays at the extreme bias region; $\{\mu^*, \lambda^*\} = \{1, \frac{1}{2}\}$ or $\{\frac{1}{2}, 1\}$.

4. For $\gamma \in (\frac{5}{9}, 1)$, one of senders stays out, and the other sender reduces the accuracy for her primary and provides perfectly uninformative signal for the other state,

5. for $\gamma > 1$, both senders stay out; $\{\mu^*_1, \lambda^*_1\} = \{\mu^*_2, \lambda^*_2\} = \{\frac{1}{2}, \frac{1}{2}\}$.

1.5.3 Multiple Senders’ Case

Now, consider that $K$ senders potentially enter the market. Given their information structures, we can make an order of information senders by their bias, $\beta_K \leq \cdots \leq \beta_1$, and find a cutoff agents, $\bar{v}_{k+1,k}$ who is indifferent between the sender $k$ and the sender $k+1$, where $\bar{v}_{k+1,k} = \frac{a_k-a_{k+1}}{\beta_{k+1}-\beta_k}$. Each sender’s maximization problem is the following.

1. For the sender 1, $\max_{\mu_1, \lambda_1} [U(\bar{v}_1) - U(\bar{v}_{12})] - c(\mu_1, \lambda_1) = \frac{1}{2} \left[ \frac{\alpha_1}{\beta_1 - \beta_1} - \frac{\alpha_1 - \alpha_2}{\beta_1 - \beta_2} \right] - \gamma \left[ \frac{1}{1 - \alpha_1} - 1 \right]$.

2. For the senders $k = 2, \cdots, K - 1$, $\max_{\mu_k, \lambda_k} [U(\bar{v}_{k-1,k}) - U(\bar{v}_{k,k+1})] - c(\mu_k, \lambda_k) = \frac{1}{2} \left[ \frac{\alpha_k - a_{k+1}}{\beta_{k+1} - \beta_k} - \frac{\alpha_k - a_{k+1}}{\beta_{k+1} - \beta_k} \right] - \gamma \left[ \frac{1}{1 - \alpha_k} - 1 \right]$.
3. For the sender $K$, $\max_{\mu_K,\lambda_K} [U(\bar{v}_{K-1,K}) - U(\bar{v}_K)] - c(\mu_K, \lambda_K) = \frac{1}{2}[\frac{\alpha_{K-1} - \alpha_K}{\beta_{K-1} - \beta_K} - \frac{-\alpha_K}{1+\beta_K}] - \gamma[\frac{1}{1-\alpha_K} - 1]$.

In this case, $\beta_k$ is between $\beta_{k+1}$ and $\beta_{k-1}$, which implies that the information sender $k$ is less biased than one of others or both. The following lemma shows that if one sender’s bias is between others, her overall accuracy should not be lowest. When it is lower than all others, she will be dominated by the more biased and more accurate information sender and thus, no one chooses the sender $k$; her market share is zero, $S_k = 0$. Hence, given the bias order, she will provide the more accurate information structure than the more biased senders to be chosen by some decision makers. In addition, if she is biased towards the same direction with other information senders but her bias is between them, her accuracy should be highest.

**Lemma 1.6.** Suppose that the information sender $k$’s bias is fixed between $\beta_{k-1}$ and $\beta_{k+1}$, $\beta_k \in (\beta_{k+1}, \beta_{k-1})$. The sender $k$’s best response is the following.

1. The information sender $k$’s overall accuracy will be higher than one of others to be chosen, $\alpha_k > \min\{\alpha_{k-1}, \alpha_{k+1}\}$.

2. If they are biased towards the same direction, $\beta_{k-1} < 0$ or $\beta_{k+1} > 0$, the information sender $k$’s overall accuracy will be highest, $\alpha_k > \max\{\alpha_{k-1}, \alpha_{k+1}\}$.

Now, consider an information sender $k$’s bias in the case of $\beta_k \in [\beta_{k-1}, \beta_{k+1}]$. From lemma 1.7, since her bias is between others, she will provide the more accurate information than at least one of them. If $\alpha_k > \alpha_{k+1}$ and $|\beta_k| < |\beta_{k+1}|$, the sender $k$ can dominate the sender $k + 1$ by increasing her bias to $|\beta_{k+1}|$ while holding her overall accuracy constant at $\alpha_k$. It does not change the sender $k$’s accuracy cost but
increase her market share and profit. Hence, the sender $k$ has no incentive to keep her bias at $\beta_k$ and thus, $\beta_k \in (\beta_{k-1}, \beta_{k+1})$ cannot be equilibrium for the sender $k$.

**Lemma 1.7.** Suppose that an information sender $k$’s bias is between sender $k − 1$’s and sender $k + 1$’s, $\beta_k \in [\beta_{k+1}, \beta_{k-1}]$. Then, $\beta_k \in (\beta_{k-1}, \beta_{k+1})$ is not optimal for the sender $k$.

As a result, there is no information sender who chooses the unbiased information structure and the market is segmented into two opposite directions. Suppose that $N$ is the equilibrium number of information senders in the market, where $N \geq 3$.\(^{11}\)

When $N$ is even, each segment has the same number of biased information senders. A half of senders have the same positive bias but the other half have the same negative bias; $\beta_i^* = \max \{x_N, \frac{1}{2}\}$ and $\beta_j^* = -\beta_i^*$, where $x_N = \sqrt{\frac{2N}{N-2}} \gamma + \left[\frac{N}{4(N-2)}\right]^2 - \frac{N}{4(N-2)}$ for $i = 1, \cdots, \frac{N}{2}$ and $j = \frac{N}{2} + 1, \cdots, N$. When $N$ is odd, however, one of segments must have one more sender than the other and thus, $\frac{N+1}{2}$ of information senders have the same $\beta$ while another $\frac{N-1}{2}$ of them have the opposite $\beta$; $\beta_i^* = \max \{x_{N+1}, \frac{1}{2}\}$ and $\beta_j^* = -\beta_i^*$ for $i = 1, \cdots, \frac{N-1}{2}$ and $j = \frac{N+1}{2}, \cdots, N$ or for $i = 1, \cdots, \frac{N+1}{2}$ and $j = \frac{N+3}{2}, \cdots, N$. Furthermore, the equilibrium overall accuracy depends on whether $N$ is even or odd, but in both cases, all active information senders are motivated to provide the same overall accuracy; $\alpha_i^* = \cdots = \alpha_N^* = \begin{cases} 1 - x_N & \text{if } N \text{ is even} \\ 1 - x_{N+1} & \text{if } N \text{ is odd} \end{cases}$. Otherwise, the less accurate information sender is dominated by others since all they have the same bias and a decision maker’s expected utility increases with the overall accuracy and with the bias as well. In the paper, a decision maker has continuous preference for the choices, however, all information senders are located at the two opposites.

\(^{11}\)Remark 1.1 describes how the equilibrium number of information senders is determined.
Proposition 1.3. Suppose that $N$ is the equilibrium number of active information senders in the market, where $N \geq 3$. Then, in equilibrium,

1. (a) when $N$ is even, a half of information senders have the same positive $\beta$ while another half of them have the same negative $\beta$;

$$\beta_i^* = \max\{x_N, \frac{1}{2}\} \text{ and } \beta_j^* = -\beta_i^*, \text{ where } x_N = \sqrt{\frac{2N}{N-2} \gamma + \left[\frac{N}{4(N-2)}\right]^2 - \frac{N}{4(N-2)}}$$

for $i = 1, \cdots, \frac{N}{2}$ and $j = \frac{N}{2} + 1, \cdots, N$.

(a) when $N$ is odd, $\frac{N+1}{2}$ of information senders have the same $\beta$ while another $\frac{N-1}{2}$ of them have the opposite $\beta$;

$$\beta_i^* = \max\{x_{N+1}, \frac{1}{2}\} \text{ and } \beta_j^* = -\beta_i^*$$

for $i = 1, \cdots, \frac{N-1}{2}$ and $j = \frac{N+1}{2}, \cdots, N$ or for $i = 1, \cdots, \frac{N+1}{2}$ and $j = \frac{N+3}{2}, \cdots, N$.

2. All active information senders will provide the same overall accuracy;

$$\alpha_1^* = \cdots = \alpha_N^* = \begin{cases} 1 - x_N & \text{if } N \text{ is even} \\ 1 - x_{N+1} & \text{if } N \text{ is odd} \end{cases}$$

Proof. Attached in Appendix

The equilibrium number of active information senders is determined by whether their profits are positive or not. The following remark shows how the equilibrium number of senders in the market is determined.

Remark 1.1. This remark represents how the equilibrium number of active information senders in the market, $N$, is determined. Start with $N = K$, where $K$ is the potential number of information senders and $K$ is even. Since all they have the same overall
accuracy, they have the same accuracy cost as well. Moreover, the market is exactly segmented into two opposite directions with the same absolute value of bias, so their fraction of subscribers is also the same and thus, all they have the same profit. As long as their profits are positive, no information senders leave the market. Once their profit is negative as the accuracy cost increases, however, one of senders from each segment leaves, and there will be \( K - 2 \) senders in the market. If their profit is positive with \( K - 2 \) senders, then the equilibrium number of senders is now \( K - 2 \), and this process is repeated with \( N = K - 2 \). On the other hand, when \( K \) is odd, one of segments has one more information sender, so their fraction of subscribers and profit in this segment are lower than the other’s. Hence, an information sender’s profit in this segment will turn out to be negative first and thus, one of senders only in this segment leaves the market. Then, \( K - 1 \) senders remain in the market, where \( K - 1 \) is even, and the case is exactly the same with above; repeat this process with \( N = K - 1 \) which is even.

From proposition 1.3, we can characterize the equilibrium reporting strategy for the special case of \( K \)-senders, triopoly; \( K = 3 \). This gives a good intuition for the general case. Since the market is segmented into two opposite directions and there are three senders in the market, two of them should be biased towards the same direction and conduct accuracy competition. This also affects the sender who is biased towards the opposite. After the market reaches at the extreme bias region, senders start to leave the market and eventually, only one of them is active and the market becomes monopoly.

**Corollary 1.2.** In triopoly, the equilibrium reporting strategy is the following.
1. For $\gamma \in (0, \frac{3}{16})$, two senders provide a perfect signal for the same state and reduce the accuracy for the other state as $\gamma$ increases while the other sender is the opposite.

2. For $\gamma \in (\frac{3}{16}, \frac{1}{4})$, two senders stay at the same extreme bias region and the sender stays at the extreme bias region with the opposite direction.

3. For $\gamma \in (\frac{1}{4}, \frac{1}{2})$, one of senders stays out, and two senders stay at the extreme bias region with opposite directions.

4. For $\gamma \in (\frac{1}{2}, \frac{5}{9})$, two of senders stay out, and the other sender stays at the extreme bias region.

5. For $\gamma \in (\frac{5}{9}, 1)$, two of senders stay out, and the other sender reduces the accuracy for her primary and provides perfectly uninformative signal for the other state.

6. For $\gamma > 1$, all information senders stay out of the market.

Proof. Attached in Appendix.

1.6 Overall Accuracy and Bias

When the accuracy cost is sufficiently high, information senders start to leave the market one by one until only one sender remains in the market. Once one of the information senders leaves the market, $K$ potential senders case becomes $K - 1$ senders case. Before the sender leaves the market, the market reaches at the extreme bias region. Thus, it is meaningful to find when the market reaches at the region, $\bar{\Gamma}_K$, which is different in different market structures. In monopoly and duopoly, the extreme bias region starts when $\gamma$ is equal to $1/9$ and $1/8$, respectively, hence $\bar{\Gamma}_1 = 1/9$.
and $\bar{\Gamma}_2 = 1/8$. For triopoly case, $\bar{\Gamma}_3 = 3/16$. From monopoly to triopoly, the cutoff is increasing in the number of potential information senders, which implies that as more information senders potentially enter the market, the higher cost is required for the market to reach at the extreme bias region. Overall accuracy before the market reaches at the extreme bias region, it is increasing in the number of potential senders $K$ as well. Therefore, overall accuracy is weakly increasing as more potential information senders enter the market. In addition, the probability of making a correct decision for a decision maker is also weakly increasing in $K$ because it is simply a function of the overall accuracy. Thus, for a consumer, it is more likely to make a correct decision as more information senders are in the market.

In this paper, information bias is driven by the demand-side, decision makers. However, bias is also weakly decreasing as there are more potential information senders. Information senders only can choose the combination in the triangle in Figure 1.3. In equilibrium, they will choose the combination on the line to maximize their profits. Hence, when the overall accuracy is higher than $1/2$, the higher accuracy means the smaller room for the bias. The overall accuracy has the negative relationship with the bias and thus, a decrease in the overall accuracy results in an increase in the bias. Figure 1.3 represents how the overall accuracy and the bias change with the various market structures, especially in the cases of monopoly, duopoly and triopoly.

1.7 Multiple Subscription Case

A decision maker is limited to subscribe to only one information sender so far. In this chapter, we relax the assumption and consider the possibility of cross-checking; a decision maker can simultaneously choose two different information senders. For
The higher overall accuracy, the smaller bias.

Figure 1.3: Relationship between Overall Accuracy and Bias

Overall Accuracy

Bias

Figure 1.4: Overall Accuracy and Bias over Market Structures
this case, we assume that there are at least four potential information senders in the market, \( K \geq 4 \).

Suppose that a decision maker chooses information senders, \( i \) and \( j \) who provide \( \{\mu_i, \lambda_i\} \) and \( \{\mu_j, \lambda_j\} \), respectively for \( i, j = 1, \ldots, K \).\(^{12}\) Without loss of generality, we assume that \( \mu_i \geq \mu_j \) and \( \lambda_i \leq \lambda_j \). In this case, four different scenarios from these two information senders can be considered; (i) \( s_i = s_j = 1 \), (ii) \( s_i = s_j = -1 \), (iii) \( s_i = 1 \) and \( s_j = -1 \), and (iv) \( s_i = -1 \) and \( s_j = 1 \). Hence, we obtain the posterior beliefs for these four scenarios as the following.

1. \( \Pr(\omega = 1 | s_i = s_j = 1) = \frac{\mu_i \mu_j}{\mu_i \mu_j + (1-\lambda_i)(1-\lambda_j)} \).

2. \( \Pr(\omega = 1 | s_i = 1, s_j = -1) = \frac{\lambda_j (1-\mu_j)}{\mu_i (1-\mu_j) + (1-\lambda_i)\lambda_j} \).

3. \( \Pr(\omega = 1 | s_i = -1, s_j = 1) = \frac{(1-\mu_i)\mu_j}{(1-\mu_i)\mu_j + \lambda_i (1-\lambda_j)} \).

4. \( \Pr(\omega = 1 | s_i = s_j = -1) = \frac{(1-\mu_i)(1-\mu_j)}{(1-\mu_i)(1-\mu_j) + \lambda_i \lambda_j} \).

Given the information structures, we can find the optimal action strategy for the decision maker which depends on his preference and signals he receives. If his preference is sufficiently high, then he always chooses \( x = 1 \) for any signals he receives, and if his preference is high but not sufficiently high, then he chooses \( x = -1 \) when he receives two negative signals, i.e., \( s_i = s_j = -1 \). Otherwise, he chooses \( x = 1 \). It is similar for the case in which his preference is low or sufficiently low. If his preference is moderate, on the other hand, he only follows the more accurate information sender.

**Lemma 1.8.** Given information structures, \( \{\mu_i, \lambda_i\} \) and \( \{\mu_j, \lambda_j\} \), the optimal action choice for a decision maker is the following,

\(^{12}\)The decision maker is restricted to subscribe to two information senders. This assumption will be relaxed to choose one or two senders including not subscription choice.
1. for \( v > \bar{v}_{ij} \), \( x^*(v, i, j, s_i, s_j) = 1 \) for all \( s_i, s_j \), where \( \bar{v}_{ij} = \frac{\lambda_i \lambda_j - (1 - \mu_i) (1 - \mu_j)}{\lambda_i \lambda_j + (1 - \mu_i) (1 - \mu_j)} u \) 

2. for \( v \in (\bar{v}_{ij}, \bar{v}_{ij}) \), \( x^*(v, i, j, s_i, s_j) = \begin{cases} -1 & \text{if } s_i = s_j = -1 \\ 1 & \text{otherwise} \end{cases} \), where \( \bar{v}_{ij} = \max\{\hat{v}_{1, -1}, \hat{v}_{-1, 1}\}^{13} \)

3. for \( v \in (v_{ij}, \bar{v}_{ij}) \), \( x^*(v, i, j, s_i, s_j) = s_k \), where \( s_k = \begin{cases} s_i & \text{if } \hat{v}_{1, -1} > \hat{v}_{-1, 1} \\ s_j & \text{if } \hat{v}_{-1, 1} > \hat{v}_{1, -1} \end{cases} \)

4. for \( v \in (v_{ij}, v_{ij}) \), \( x^*(v, i, j, s_i, s_j) = \begin{cases} 1 & \text{if } s_i = s_j = 1 \\ -1 & \text{otherwise} \end{cases} \), where \( v_{ij} = \min\{\hat{v}_{1, -1}, \hat{v}_{-1, 1}\} \)

5. for \( v < v_{ij} \), \( x^*(v, i, j, s_i, s_j) = -1 \) for all \( s_i, s_j \), where \( v_{ij} = \frac{(1 - \lambda_i)(1 - \lambda_j) - \mu_i \mu_j}{(1 - \lambda_i)(1 - \lambda_j) + \mu_i \mu_j} u \).

**Proof.** Attached in Appendix

Since the moderate decision maker just follows one of the signals depending on \( \hat{v}_{1, -1} \) and \( \hat{v}_{-1, 1} \), his expected utility from choosing \( i \) and \( j \) is exactly the same with that from subscribing to one of them; hence, it is still increasing in both the overall accuracy and the bias. For the decision maker in \( v \in (\bar{v}_{ij}, \bar{v}_{ij}) \), his expected utility from choosing \( i \) and \( j \) is

\[
Eu(y = \{i, j\} | v \in (\bar{v}_{ij}, \bar{v}_{ij})) = A_{ij} u + B_{ij} v
\]

where \( A_{ij} = \lambda_i \lambda_j - (1 - \mu_i)(1 - \mu_j) \) and \( B_{ij} = 1 - \lambda_i \lambda_j - (1 - \mu_i)(1 - \mu_j) \).

Since this decision maker chooses \( x = -1 \) only when both information senders provide negative signals, it is important for him to receive correct signals for the state \(-1\). Hence, we have the above result. Similarly, for the decision maker in \( (v_{ij}, v_{ij}) \),

\[
Eu(y = \{i, j\} | v \in (v_{ij}, v_{ij})) = A'_{ij} u + B'_{ij} v
\]

where \( A'_{ij} = \mu_i \mu_j - (1 - \lambda_i)(1 - \lambda_j) \) and \( B'_{ij} = \mu_i \mu_j + (1 - \lambda_i)(1 - \lambda_j) - 1 \).

\(^{13}\hat{v}_{1, -1} = \frac{(1 - \lambda_i) \lambda_j - \mu_i (1 - \mu_j)}{(1 - \lambda_i) \lambda_j + \mu_i (1 - \mu_j)} u \) and \( \hat{v}_{-1, 1} = \frac{\lambda_i (1 - \lambda_j) - (1 - \mu_i) \mu_j}{\lambda_i (1 - \lambda_j) + (1 - \mu_i) \mu_j} u \)
Now, compare a decision maker’s expected utility from a single sender and that from multiple senders. The decision maker has an incentive to subscribe to multiple senders only when the expected utility from multiple senders is higher than that from the single sender who provides the highest expected utility in a single sender case. The following lemma shows the condition for that the decision maker has an incentive to choose an additional sender, given the optimal sender in a single sender case.

**Lemma 1.9.** Suppose that \( \{\mu^*_i, \lambda^*_i\} \) is a decision maker’s equilibrium information structure in a single sender case.

1. If \( v \in (\tilde{v}_{ij}, \bar{v}_{ij}) \) or \( v \in (v_{ij}, \tilde{v}_{ij}) \), the decision maker has an incentive to subscribe to an additional information sender \( j \).

2. If \( v \in (v_{ij}, \bar{v}_{ij},) \), the decision maker has no incentive to subscribe to an additional information sender \( j \).

The following proposition shows that the decision maker as no incentive to choose an oppositely biased sender (say, sender \( j \)) with the optimal sender (say, sender \( i^* \)) in a single sender case. This is because a signal from the additional sender never convinces the decision maker. Whenever a signal from the sender conflicts with a signal from the sender \( i^* \), the decision maker will follow the signal from the sender \( i^* \). Hence, the signal from the sender \( j \) is useless for the decision maker and thus, the decision maker has no incentive to subscribe to the additional sender who is oppositely biased with sender \( i^* \). The proposition also implies that if there is the same information structure with the sender \( i^* \), the decision maker is always better off with the additional sender \( j \), which gives an incentive to choose the additional information sender to the decision maker.
Proposition 1.4. Suppose that \( \{\mu^*_i, \lambda^*_i\} \) is a decision maker’s equilibrium information structure in single subscription case.

1. The decision maker always has an incentive to subscribe to an additional information sender \( j \) whose information structure is the same with the sender \( i \), \( \mu_j = \mu^*_i \) and \( \lambda_j = \lambda^*_i \).

2. The decision maker has no incentive to subscribe to an additional information sender \( j \) whose information structure is opposite with the sender \( i \), \( \mu_j = \lambda^*_i \) and \( \lambda_j = \mu^*_i \).

1.8 Conclusion

This paper investigates why and how information senders are biased. In the paper, we do not allow an information sender’s own preference for the product and a decision maker’s psychological utility from biased information. Nevertheless, a rational Bayesian decision maker has indirect demand for information bias. Though making a correct decision is more important than his idiosyncratic term, a decision maker chooses a like-minded information sender.

By making bias, information senders can easily attract decision makers with the low accuracy cost. Hence, no matter how many senders potentially are available, all of them have an incentive to be biased and thus, the market is segmented into two opposite directions. In the multiple senders case, all active information senders will provide the same overall accuracy, otherwise she will be strictly dominated by others.

We also see the competition effect; as more potential information senders enter the market, the overall accuracy is weakly increasing and a probability of making a correct decision by decision makers is weakly increasing as well. In addition, even
though the bias is driven by the demand-side, it is weakly decreasing in the number of potential information senders.
Chapter 2: Social Learning With a Ratings System

2.1 Introduction

Suppose that people consider purchasing a product, like a book, with limited information about quality. They can buy a best-seller, because a high number of sales implies that the book has already been purchased by many people who believed that it has high quality. In this process, they can partially extract others’ information by observing the number of sales. For this information, they can visit on-line markets such as Amazon.com or BarnesAndNoble.com, and sort products by number of sales, referred to as popularity. Another piece of information which is provided by the internet retailers is ratings. When they visit these websites, they can also see the average ratings of the products. This is another way of collecting information and is a type of social learning that is commonly used in purchase decisions nowadays. The ratings can provide information about how the previous buyers feel about the products. However, average ratings and popularity provide sometimes different results, since some popular products have lower ratings while less popular ones have higher ratings. In this case, the readers might be confused about which source provides the correct information. The main difference between popularity and ratings is that the

\[14\] For example, tripadvisor.com, expedia.com, yelp.com, and rottentomatoes.com provide various types of ratings information.
purchase decision is made based on ex-ante utilities of buyers, while ratings reflect
their ex-post utilities. The study of observing actions as aggregators of information
has been developed extensively, but the study of information conveyed by ratings
has been less extensively examined, although ratings are commonly used in the real
world.

The social learning literature has emphasized the possibility of observing other
people's actions to collect information. In the basic setting, infinite individuals are
exogenously ordered and they take binary choices such as purchase or investment de-
cisions. The decisions are made based on their own private signals and the history of
the past actions observed. Banerjee (1992) and Bikhchandani, Hirshleifer and Welch
(1992) investigated herding, where individuals take the same actions by observing
others' decisions and ignoring their own private information. Smith and Sorensen
(2000) explored how individuals learn sequentially from the discrete actions of others
and how the public belief converges over time. When public beliefs converge to a
limit point and no one's action delivers any further information, the belief conver-
gence forces action convergence and herds arise. They also investigated the existence
of confounded learning when there are different types of agents in the economy. Ali
(2014), Hendrick, Sorensen and Wiseman (2012), and Mueller and Pai (2014) intro-
duced costly information acquisition or search in the social learning model. In all
these studies, the information source for the observational learning is others' actions
only.

In the past, it was difficult to observe user-generated ratings in large samples.
Only professional reviews and small sample size of word-of-mouth were available.
However, as internet service has now grown, we can easily give and observe ratings for
products. Many internet retailers provide user-generated ratings for their customers and ratings are now one of the most important information sources that many buyers use for purchase decisions. There are empirical studies that emphasize the importance of ratings in purchase decisions. Chevalier and Mayzlin (2006) showed that average ratings have a significant impact on book sales. Li and Hitt (2008) examined that average ratings and product sales have a positive relationship. Dellarocas, Zhang and Awad (2007) found that user-generated reviews are good predictors for box office sales. These studies on ratings focused on empirical approaches, but theoretical studies on ratings have been less extensively examined.

In this paper we develop a social learning model with a ratings system where buyers can observe not only past history of purchase decisions but also ratings of previous buyers. We investigate how ratings work in a social learning model and whether they improve learning or not. Throughout this paper, two different information environments are considered:

No Ratings Information Environment: Individuals observe the past history of others’ actions without ratings.

Ratings Information Environment: Individuals give ratings for the product based on their ex-post payoffs, and others can observe the history of ratings and past actions.\textsuperscript{15}

We set up a social learning model with a ratings system in section 2.2. Individuals not only observe predecessors’ actions, but also see their ratings. For simplicity, we restrict attention to the case in which only those who purchased the product can give ratings. Moreover, the frequency of giving ratings is assumed to be one. Hence,\textsuperscript{15}

\textsuperscript{15}This model looks at a stylized version where everyone leaves a review and all of the decisions to purchase and not to purchase are observed.
ratings are automatically provided to successors whenever predecessors purchase the product. It implies that a purchase by a predecessor is a chance for successors to observe his decision and feedback at the same time. If a buyer does not purchase the item, successors only observe his decision without ratings.

In section 2.3, we consider bounded private beliefs which implies there is a limited amount of information for individuals. There is no one whose private signal has perfect accuracy and thus, agents want to collect more information by observing others’ actions. In this case, ratings prevent consumers from purchasing low quality products, and sometimes, even high quality ones. That is because there might be negative feedbacks even for high quality products which lower the public beliefs. Moreover, the probability of high quality products being ignored is not monotonic in the accuracy of ratings. When ratings information is accurate, people rely on ratings much more than the history of past actions and their own private signals. In that case, a small number of negative ratings prevent people from purchasing high quality products and thus, the probability of incorrect herds might be higher than the case where ratings are less accurate than private signals. More interestingly, ratings do not guarantee better learning and the probability of incorrect herds might be higher with ratings than that without ratings. Incorrect herds are more likely to occur with ratings even when the ratings are more accurate than private signals. The additional information is accurate, but it might hurt the buyers’ payoffs.

Extensions are briefly discussed in section 2.4. We relax assumptions on ratings system that is considered in sections 2.2 and 2.3. The frequency of giving ratings

\[16\] \text{With unbounded private beliefs, individuals have no maximum amount of information about the state of the world.}
might be less than 1 and not every buyer leaves feedback for the product. In addition, the possibilities of manipulation and endogenous timing are also considered.

2.2 Model

A new product is introduced and its quality depends on an underlying state of the world. There are two states of nature which are payoff relevant, indexed by \( \omega \in \{H, L\} \), and they are equally likely, \( pr(\omega = H) = pr(\omega = L) = \frac{1}{2} \). In state \( H \), the product provides higher quality than in state \( L \), \( u_H > u_L \). To simplify the notation of the exposition, we assume that \( u_H = u > 0 \) and \( u_L = -u < 0 \). An infinite sequence of exogenously ordered individuals, indexed by \( n \in \mathbb{N} \) sequentially takes a one-shot binary action which is not reversible, \( x_n \in \{0, 1\} \). We can consider \( x_n = 1 \) as purchase of the new product and \( x_n = 0 \) as no purchase.

There is an idiosyncratic shock from the purchase decision, \( v_n \) which is realized conditional on \( x_n = 1 \). We suppose that the idiosyncratic shocks are i.i.d. across individuals, following the cumulative distribution, \( G \), where \( [-u, u] \subset supp(G) \), and \( E(v_n) = 0 \). The idiosyncratic shocks are unknown before purchase and thus, they do not affect individuals’ purchase decisions.

The payoff of individual \( n \) depends on the quality of the product he purchased, and his idiosyncratic shock which is revealed after purchase, \( (u_\omega + v_n)x_n \), where \( x_n \in \{0, 1\} \). Hence, if the agent \( n \) chooses \( x_n = 1 \), his ex-post payoff will be \( u_\omega + v_n \).

---

17After purchase, individuals have enough time to experience and scrutinize the product, so that he could realize that there are some aspects which were not considered before purchase. This is captured by his private idiosyncratic shocks.

18Since the support for idiosyncratic shocks is greater than that for common qualities, consumers might have negative payoffs from their purchase decisions even in state \( H \). Similarly, it is also possible that they get positive payoffs from their purchase decisions when the quality is low.

19Agents can observe the entire ex-post payoffs after purchase, not the idiosyncratic shocks. Without the idiosyncratic shocks, ratings simply reveal the true state.
where $u_\omega$ is a common value which is identical across all agents conditional on the state, and $v_n$ is his private value. On the other hand, if he chooses $x_n = 0$, his payoff is simply zero in both states, so it is always a safe choice.

In contrast to the previous literature on social learning, we consider the information environment where individuals can observe ratings from the past buyers as well as the history of past purchases. Buyers give ratings for the product, $y_n \in \{-1, 1\}$, only when they purchase the product, $x_n = 1$. We investigate the case of sincere feedbacks. For the agent who did not purchase the item, no ratings information is provided to the later agents and $y_n$ is given as zero. In addition, all buyers must leave their ratings.\textsuperscript{20} Hence when public beliefs about the quality is intermediate, a purchase decision may deliver information about one’s private signals to later agents and change the public beliefs. At the same time, the purchase decision provides an opportunity for later agents to receive additional information, ratings. If the agent does not purchase, this information is not provided.

$$y_n = \begin{cases} 
1 & \text{if } x_n = 1 \text{ and } u_\omega + v_n \geq 0 \\
-1 & \text{if } x_n = 1 \text{ and } u_\omega + v_n < 0 \\
0 & \text{if } x_n = 0
\end{cases}$$

Throughout the paper two different information environments are considered:

**No Ratings Information Environment (NR):** For individual $n$, the past action history, $\mathbb{H}_n = \{x_1, \ldots, x_{n-1}\}$ is the only information aggregator.

**Ratings Information Environment (R):** Individuals leave ratings for the product based on their ex-post payoffs, and others can observe ratings and past actions, $\mathbb{H}_n = \{(x_1, y_1), \ldots, (x_{n-1}, y_{n-1})\}$.

\textsuperscript{20}We will discuss the relaxation of these assumptions in Chapter 2.4.
The state, \( \omega \), is unknown to agents, but they receive their own private signals, \( s_n \), and so they form their private beliefs based on the signals. The individuals' private beliefs, \( p_n \), are identically and independently generated according to a distribution \( F_\omega \) depending on the state, \( \omega \in \{ H, L \} \). We assume that \( F_H \) and \( F_L \) are absolutely continuous with respect to each other. Since they are not identical distributions, we can say that the signals are informative. We also define bounded private beliefs if the convex hull of common support of \( F_H \) and \( F_L \) is an interval \([p, \bar{p}]\) where \( 0 < p < \frac{1}{2} < \bar{p} < 1 \). On the other hand, the private beliefs are unbounded if the convex hull of common support is \([0, 1]\).

At the period \( n \), individuals observe the history of information and update the public beliefs to \( q_n = \text{prob}(\omega = H \mid \mathbb{H}_n) \). The agent \( n \) combines the public belief and his own private belief, \( p_n \), to construct the posterior belief \( r_n \). Given the public belief \( q_n \),

\[
r_n(p_n, q_n) = \frac{p_n q_n}{p_n q_n + (1 - p_n)(1 - q_n)}
\]

Each individual makes his purchase decision, \( x_n \in \{0, 1\} \), based on his expected payoff and on the posterior belief, \( r_n \). When the expected payoff is zero, the individual chooses \( x_n = 1 \) with probability \( \frac{1}{2} \) and \( x_n = 0 \) with probability \( \frac{1}{2} \). 21

2.3 Equilibrium Dynamics

In this section we characterize the equilibrium dynamics and the outcomes of the model. We are interested in exploring long-run behavior of buyers and, in particular, the likelihood of “incorrect herds.” Incorrect herds occur when high quality products are ignored or when low quality products are chosen. A Bayesian equilibrium is

21This decision is referred to as ”flip a coin”.

38
considered in which agents update their beliefs using Bayes’ rule, based on the history of information. We look for the dynamics of beliefs and actions which follows Smith and Sorensen (2000). Define the public likelihood ratio as

$$\ell_n^H = \frac{1 - q_n}{q_n} = \frac{\text{prob}(\omega = L | H_n)}{\text{prob}(\omega = H | H_n)}$$

and let $\ell_0^H$ denote the initial likelihood.\textsuperscript{22} If this likelihood ratio converges to zero, the true state is perfectly revealed to be state $H$ and we can say that learning is complete. On the other hand, if the public belief converges to 1, the likelihood ratio goes to infinity, so that we should consider the inverse of likelihood ratio in state $L$, $\ell_n^L = \frac{q_n}{1 - q_n}$. As the inverse converges to zero, the true state is revealed to be state $L$. The likelihood ratios, $\ell_n^H$ and $\ell_n^L$ are, respectively, martingale stochastic processes conditional on, respectively, state $H$ and state $L$.\textsuperscript{23}

Next, we define the learning region where private information can be partially or completely captured by purchase and no-purchase decisions. In the learning region, whenever the predecessors take their actions, the actions reveal something about their own private signals and the successors update their beliefs. Out of the learning region, in contrast, all types choose the same action so public beliefs do not change. Hence, the public beliefs could stop changing when they are out of the learning region. The range of the learning region depends on whether the private beliefs are bounded or not. If the private beliefs are unbounded, the learning region is simply $(0, 1)$ and learning stops only when the likelihood ratios are 0 or 1. With bounded private beliefs, in contrast,

\textsuperscript{22}$\ell_0^H = 1$ if the prior belief for state $H$ is $\frac{1}{2}$.

\textsuperscript{23}The general proof for this discussion is provided in Smith and Sorensen(2000) and the specific case will be discussed in the Appendix.
Lemma 2.1. [Learning Region with bounded private beliefs]

1. Suppose that the state is $H$. When $\ell^H_n \leq \ell^H \leq \bar{\ell}^H$, agents update the public beliefs after observing agent $n$’s action, where $\ell^H = \frac{1 - \bar{p}}{\bar{p}}$ and $\bar{\ell}^H = \frac{1 - p}{p}$.

2. Suppose that the state is $L$. When $\ell^L_n \leq \ell^L \leq \bar{\ell}^L$, agents update the public beliefs after observing agent $n$’s action, where $\ell^L = \frac{\bar{p}}{1 - \bar{p}}$ and $\bar{\ell}^L = \frac{p}{1 - p}$.

Beliefs outside the learning region in the no ratings information environment stop the observational learning and no further information about their private signals is delivered by any decisions of buyers. However, in the ratings information environment, buyers give their ratings for the product when they purchase it, so that learning still occurs about the quality of the product even when all agents purchase regardless of their private signals. In the learning region, the agent $n$’s purchase decision conveys information about both his own private signal and ex-post payoff. With significantly higher public beliefs, in addition, all agents simply follow the majority and purchase the item. Hence, his rating for the product delivers his ex-post payoff information which includes the information about the state.

For the ratings information environment, we need an additional assumption on the distribution of idiosyncratic shocks to simplify our discussion.

Assumption 2. The distribution for idiosyncratic shock, $G$, is symmetric;

\[ G(v) = 1 - G(-v) \text{ for all } v \in \text{supp}(g). \]

Both normal and uniform distributions satisfy this assumption. It also implies that $0 < G(-u) < \frac{1}{2} < G(u) < 1$. 

40
In the ratings information environment, three different information sources are available for individuals: the previous agents’ purchase decisions and ratings, and their own private signals. Since the previous agents’ signals and their own signals have the same accuracy, the relative accuracy between a private signal and a rating determines the posterior public belief.

Given the observed information history, consumer \( n \) combines the public belief and his own private belief to build the posterior belief according to Bayes’ rule. Recall that if consumer \( n \) chooses \( x_n = 0 \), then his payoff is zero. Otherwise, the expected payoff is \( u \) in state \( H \) and \(-u\) in state \( L \). Hence, the criteria of the posterior beliefs for purchase decision is \( \frac{1}{2} \). When agents’ posterior beliefs are above \( \frac{1}{2} \), their optimal actions are purchasing the product. On the other hand, if their beliefs are lower than \( \frac{1}{2} \), they do not purchase. In addition, if their public beliefs are exactly \( \frac{1}{2} \), then they are indifferent between purchase and not, so that they simply flip a coin.

2.3.1 Binary Private Signals

In this subsection, we assume that agents receive binary private signals, \( s_n \in \{h, l\} \) and the accuracy of the signals is \( \sigma \).

\[
prob(s_n = h|\omega = H) = \text{prob}(s_n = l|\omega = L) = \sigma \in (\frac{1}{2}, 1)
\]

Without ratings information, the model was solved by Bikhchandani et al.(1992). In their paper, herds occur when the imbalance between two actions taken exceeds two. Since herds on two actions can occur in both states, incorrect herds might occur without ratings conditional on both states, \( H \) and \( L \). On the other hand, ratings are useful to prevent incorrect herds in state \( L \). Whenever buyers purchase low quality products, they are more likely to give negative ratings. Eventually, buyers
stop purchasing the low quality products with ratings. By contrast, incorrect herds might occur in state \( H \) even with ratings. Since we allow the possibility of negative payoffs for agents, it is also possible that they give negative ratings for high quality products and thus, they might ignore high quality products.

Next, we consider how the probability of incorrect herds changes with the accuracy of ratings. In order to get the probability of incorrect herds, we use the recursive method. Let \( \lambda_i(q_n, \omega) \) denote the probability of incorrect herds in state \( \omega \) when the current public belief is \( q_n \), given the information environment \( i \in \{NR, R\} \). The next lemma shows that the probability of incorrect herds in state \( H \) is weakly decreasing in the current public belief. When agents believe that the state is more likely to be \( H \), they would like to purchase the product and the incorrect herds are less likely to occur. On the other hand, the incorrect herding is more likely to occur in state \( L \) when agents believe that the state is more likely to be \( H \).

**Lemma 2.2.**

1. The probability of incorrect herds in state \( H \) is weakly decreasing in the public beliefs for state \( H \); \( \lambda_i(q_n, H) \) is weakly decreasing in \( q_n \), \( i \in \{NR, R\} \).

2. The probability of incorrect herds in state \( L \) is weakly increasing in the public beliefs for state \( H \); \( \lambda_i(q_n, L) \) is weakly increasing in \( q_n \), \( i \in \{NR, R\} \).

Bikhchandani et al. (1992) found that probability of incorrect herds decreases in accuracy of private signals without considering a ratings system. This implies that whenever consumers observe more accurate private signals, they are more likely to

---

24 Proposition 2.3 provides the proof for the results of the continuous private signals case.

25 We will discuss the general case in section 2.3.2, continuous private signals.
make correct choices. In ratings information environment, however, it is a different story. Proposition 2.1 shows that the probability of incorrect herds is not monotonic in accuracy of ratings, which means that the more accurate ratings information does not always improve the learning. More specifically, when ratings are slightly more accurate than private signals, the probability of incorrect herds increases. This is because consumers tend to excessively rely on ratings even when the information is just slightly more accurate than private signals. Hence, a small number of negative ratings lead later consumers to stop purchasing the high quality products and thus, incorrect herds occur.

**Proposition 2.1.** Suppose the state is $H$ and $\ell^H_0 \in (\underline{\ell}^H, \bar{\ell}^H)$. Then the probability of a herd on $x = 0$ occurring is not monotonic in $G(u)$. Specifically, when $G(u)$ is slightly greater than $\sigma$, the probability of a herd on $x = 0$ increases.

Now we investigate the usefulness of ratings, i.e., whether the additional information improves learning or not. In the ratings information environment, consumers receive the previous buyers’ ratings which provide additional information about their ex-post payoffs. For the low quality products, incorrect herds might occur without ratings, while ratings eventually prevent consumers from choosing the low quality products. However, there might be incorrect herds for high quality products with ratings, and the probability of incorrect herds might be higher in the ratings information environment. There are two reasons why ratings information might not improve

---

26Bikhchandani et al. (1992) have a similar result in their fashion leader example in which the increase in the accuracy of the first agent’s private signal can increase the probability of incorrect herds. However, they compare the case of different private signals while we consider consumers to have the same accuracy of private signals, but the user-generated feedbacks have different accuracy from private signals.
learning. The first reason is similar to the previous one; consumers are overly de-
dependent on ratings information when ratings are slightly more accurate than signals
and thus, a small number of negative ratings prevents buyers from purchasing the
product. The second reason is that it is difficult to achieve the correct learning for
high quality products. The complete learning stops only when the likelihood ratio
converges to certainty, $\ell_H^\infty = 0$. On the other hand, the convergent range for incorrect
herds is the interval, $(\ell_H^- \infty, \infty)$ where no buyer purchases the high quality products. For
these reasons, ratings might hurt learning even when they are more accurate than
private signals, which are the only clue on the state in the no ratings information
environment.

**Proposition 2.2.** 1. Suppose that the state is $L$. Then ratings prevent incorrect
herds and thus, $\lambda_{NR}(q_n = \frac{1}{2}, L) > \lambda_R(q_n = \frac{1}{2}, L) = 0$.

2. Suppose that the state is $H$. Then, ratings information does not guarantee
better learning, even when ratings are more accurate than private signals, i.e.,
the region exists where $\lambda_R(q_n = \frac{1}{2}, H) > \lambda_{NR}(q_n = \frac{1}{2}, H)$ for any accuracy of
private signals $\sigma$.

### 2.3.2 Continuous Private Signals

In this section, we generalize our discussion to the continuous private signals case.
Smith and Sorensen (2000) showed that learning is complete in unbounded belief
cases, even when consumers only observe the history of past actions, without ratings.
When we consider ratings as additional information, learning is still complete, so there
is no significant difference between no ratings and ratings information environments
in the unbounded beliefs case. The bounded beliefs case, however, is different. In the
no ratings information environment, there is a positive probability of incorrect herds on both high and low quality products. When public beliefs are sufficiently high or low, later buyers tend to follow the majority and their behavior is, sometimes, not appropriate for the true state. On the other hand, low quality products eventually will be ignored in the ratings information environment, i.e., incorrect herds almost surely do not occur in state $L$. This is because buyers are more likely to provide negative ratings whenever they purchase low quality products, and the negative ratings will lower the public belief. Eventually, learning stops, and consumers do not purchase the low quality products any more with ratings information. For high quality products, in constrast, it is still possible that incorrect herds occur even with ratings. After people purchase high quality products, they might have negative ex-post payoffs due to their idiosyncratic shocks and thus, they give negative ratings for high quality products. The negative ratings tend to prevent later consumers from purchasing the product and incorrect herds might occur even with ratings.

**Proposition 2.3.** Suppose that $\ell^H_0 \in (\ell^H, \bar{\ell}^H)$.

1. In state $H$, $\ell^H_n$ converges almost surely to a random variable $\ell^H_{\infty}$ with support in the set, $\{0\} \cup (\bar{\ell}^H, \infty)$.

2. In state $L$, $\ell^L_n$ converges almost surely to a random variable $\ell^L_{\infty}$ with support in the range, $[0, \ell^L]$, where $\ell^L = \frac{1}{\ell^H}$.

Moreover, when private beliefs are bounded only from below, i.e. $p_n \in (p, 1)$ and $\ell^H_n \in (0, \bar{\ell}^H)$, the learning could be complete in both states. In this case, some agents receive perfectly optimistic private signals and they always purchase the product no matter what the previous agents choose and which ratings they give. Even when the
public beliefs are significantly low, the agents with highly optimistic private beliefs purchase the product. They are more likely to give positive ratings for high quality products and thus, herds on $x = 0$ are stopped.

**Corollary 2.1.** Suppose that $\ell^H \equiv 0$ and $\ell^H_0 \in (0, \bar{\ell}^H)$. Then, in state $H$, $\ell^H_n$ converges almost surely to zero and in state $L$, $\ell^L_n$ converges almost surely to zero.

### 2.4 Extensions

A large social learning literature has focused on observing others’ actions as aggregation of information. Observational learning allows agents to collect more information on the state of the world. But the information in the literature has been limited only to history of past actions. As technology is now well-developed, however, buyers can share their ex-post payoff information in large numbers. In this paper, we studied a social learning model in which agents can observe not only others’ actions, but also their ratings for products.

The analysis can be extended in three directions; first to relax the sincere rating assumption, second to allow buyers not to leave their ratings, and third to endogenize agents’ decision timing. The first extension is allowing manipulation in rating. Throughout the paper, we assumed all buyers leave positive ratings whenever their ex-post payoffs are non-negative. However, there is a possibility of manipulation or opinion spams in the real world. Chevalier, Dover, and Mayzlin (2014) investigated the authenticity of online user reviews by examining hotel reviews. They show that there are significant signs of manipulations in hotel reviews. However, it is difficult for buyers to distinguish between manipulation and sincere rating. Hence, the manipulation can be regarded as a noise for agents and reduces the accuracy of ratings.
information. Second, the model assumed that all buyers leave their ratings; but this might be unrealistic. However, even if we relax the assumption, we will get similar results in the long run. The relaxation causes a delay in convergence, but not different convergent points. Finally, the analysis also can be extended to an endogenous timing model. In our paper, agents are exogenously ordered and delays are not allowed. This model is rarely able to explain the empirical results from Li and Hitt (2008) which showed that the average ratings tend to decrease over time. This trend can be explained in an endogenous timing model. If the endogenous timing is allowed and there is a discounting factor, then agents with optimistic signals or with favorable preferences tend to purchase the item earlier and those with pessimistic signals never purchase it. The existence of continuous private signals and endogenous timing causes higher ratings at the beginning. The ratings, however, will decrease over time as agents with intermediate beliefs make their decisions, and this can support the empirical data which is described in Li and Hitt (2008).
Chapter 3: Noisy Observation In Diffusion of Innovations

3.1 Introduction

New technology is introduced into our lives everyday. Some of them are adopted while others are not even noticed by people. Only few of people lead new trend and most of majority just follow their actions. When the new technology is unfamiliar to people, the majority cannot tell its true value, so they rely on those who have informational advantage. As people can observe the early adopter’ behavior, they can learn information about the new technology from their decisions. Hence, the observation can provide their private information.

Most of learning literature assume that agents can observe the ordered history of the past actions. Agents precisely know how many agents made a certain decision at each stage and this conveys the private information of the previous agents. However, this assumption could be restrictive. If there are infinitely many people in the world, some proportion of decisions might be observable, but not the entire decisions. For example, if others are not their friends or family, they cannot directly observe others’ decisions. In this case, it is reasonable to introduce a noisy observation in process of learning from others’ behavior.

Adoption of new technology is considered in this paper. According to Rogers (1962),
adopters can be categorized into four different types; early adopters, early majority, late majority and laggards. Early adopters are customers who adopt the new technology first. The proportion of early adopters is relatively smaller than others and most of majority make their decisions after early adopters’ decisions. For simplicity, we consider late majority and laggards as the same group in our model. Hence, if there are three stage for adoption decision and majority adopt the new innovation, the number of early adopters is finite and they make their decisions at the first stage. People who adopt the new innovation at the second stage are early majority and the remainders are late majority. Rogers suggested four main elements that influence the spread of new technology; actual quality of the innovation, how people communicate with each other, how long the spread takes, and a social system. Our model focuses how communication channels affect the diffusion of innovations.

This paper explores a model in which there is a noise in diffusion of innovations. If the observation is not perfect, some of agents cannot observe others’ decisions while other agents can. The extent to noisy observation in diffusion of innovations depends on how accurately they observe others’ decisions. In this paper, we assume that with a certain probability, individual decision is observed and thus, they cannot observe more than the actual number of adopters. The probability of observations follows the binomial distribution. Hence, they believe that the number of actual adopters is always equal to or more than the number they observed.

The lack of perfect observation could create the heterogeneous posterior beliefs over agents. Agents observe the different amount of adoptions which affect their posterior beliefs, and thus they will have the heterogeneous posterior beliefs on quality of the
new technology and asymmetry in information. Moreover, the dynamics in the frac-
tion of adoption varies over the different degree of noise.

The fraction of adoption might depend on the noise in observation. When the inno-
vation has the high quality, it is better for majority to adopt earlier. In this paper,
the efficiency is represented by the expected fraction of early majority, conditional
on that the quality of the innovation is high. The expected fraction tends to increase
when the noise is lower. However, in some regions it could be locally decreasing,
which means that the efficiency might be decreasing in some regions.

Our starting point is the endogenous timing model, Chamley and Gale(1994). In
the paper, they considered N-player investment game with a pure informational ex-
ternality. Each player’s payoff depends only on his own action and the state of
nature. Because a player’s action perfectly reveals his private information, players
wait to collect more information by observing what other players do. They found
that there is a unique symmetric Perfect Bayesian Equilibrium in mixed strategies
and this equilibrium is inefficient because delay is costly. Chamley (2004) introduced
a heterogeneous prior beliefs distribution about the return from the investment. He
characterized a pure strategy PBE and possibility of multiple equilibria. Moreover,
the strategic complementarities may arise based on information externalities and this
generates different amount of information. He also considered large economy case
which is related with our model. Acemoglu, Dahleh, Lobel and Ozdaglar (2011) in-
vestigated a model of learning with social network. In the paper, they considered that
each individual observes others’ decisions in a stochastically generated neighborhood,
social network. Their model established the information structure where the ordered
samples are observed by a social network.
The remainder of the paper is organized as follows: Section 3.2 introduces the model setup. Section 3.3 discusses the equilibrium in the adoption game. Section 3.4 provides an example and concluding remarks are offered in Section 3.5.

3.2 Model Setup

**Agents and Endowment:** There is a measure one of risk-neutral agents, \( i \in [0, 1] \) and a new technology is introduced. An agent takes a binary action, \( x_i \in \{0, 1\} \) at each stage. We can consider \( x_i = 1 \) as adopting the new technology and \( x_i = 0 \) as keeping the old technology. If \( x_i = 1 \) is chosen, it is not reversible while \( x_i = 0 \) is reversible. Hence, before the final stage, \( x_i = 0 \) is considered as delaying a decision.

\[
x_i = \begin{cases} 
0 & \text{if the agent } i \text{ keeps the old technology} \\
1 & \text{if the agent } i \text{ adopts the new innovation}
\end{cases}
\]

**States of Nature:** There are two possible states of nature, \( S \in \{H, L\} \) which are equally likely. Both of old and new technology provide dividends every stage, but the dividend from the new technology depends on states of nature. In state \( H \), the new technology provides \( R_H = 1 \) and in state \( L \), the return is \( R_L = 0 \). On the other hand, the old technology provides \( r \in (0, 1) \) for sure, which is independent of state of nature.

**Private signal:** There are three stages, indexed by \( t = 1, 2, 3 \). The first stage is an information spreading stage where only finite number of early adopters receive private signals, \( s \in \{h, l\} \). \( N \) is the number of initially informed early adopters and there is no further information spread at stage 2 and 3. \( q \) is the quality of the private signal, \( q = \text{prob}(h|H) = \text{prob}(l|L) \in (1/2, 1) \). Since the accuracy is greater than \( \frac{1}{2} \), the private signal is informative.
Timing of the game:

The timeline of the adoption game can be summarized as follows:

Stage 1: Finite number of early adopters receive private signals about states of nature. The early adopters decide whether to adopt the new technology or not.

Stage 2 and 3: Remaining majority observe the amount of adoption and decide whether to adopt it or not, but the observation is noisy.

The finite number of early adopters have informational advantages over the uninformed majority. In sum, there are three different types of agents, the informed agents with high signal and low signal and the uninformed agents, γ ∈ \{H,L,U\}.

To simplify the notation, let $p_{i,t}^\gamma$ denote the agent $i$’s posterior belief for state $H$ at stage $t$ when the agent’s type is $\gamma$. An agent takes a binary action, $x_{i,t}^\gamma \in \{0,1\}$. Let $x_{i,t}^\gamma = 0$ represent the decision to keep the old technology and let $x_{i,t}^\gamma = 1$ represent the decision to adopt the new technology.

The equilibrium concept here is Perfect Bayesian equilibrium. In particular, we consider a symmetric pure strategy perfect Bayesian equilibrium in which agents with the same information and observation take the same action at each stage. We can find the cutoff-strategy that agents who observe more than the cutoff adopt the new technology. There is a pure strategy Perfect Bayesian Equilibrium where informed agents with high signals adopt the new technology and informed agents with low signals and uninformed agents keep the old technology at stage 1. Hence, only finite number of agents can be observed at stage 1. Moreover, after stage 2, a fraction of agents makes their decisions. We will figure out these specific variables.

Signals and Observations: At stage 1, $N$ agents receive their private signals. Among the informed agents, $y$ is the number of agents who receive the signal $h$. 

52
Hence, $N - y$ is the number of agents with the signal $l$. Because agents are ex-ante identical, in particular, each agent has the same probability of receiving the signal $h$. In the state $H$, an informed agent receives the signal $h$ with probability $q$. Hence, the probability that $y$ agents receive the signal $h$ in the state $H$ follows the binomial distribution where $N$ is the number of trial and $q$ is the probability of success.

$$h(y; S = H) = \binom{N}{y} q^y (1 - q)^{N-y}, \text{ where } y = 0, 1, \cdots, N$$

Similarly, in the state $L$, the informed agent receives the signal $h$ with probability $1 - q$. Hence, the probability that $y$ agents receive the signal $h$ in the state $L$ follows the binomial distribution where $1 - q$ is the probability of success.

$$h(y; S = L) = \binom{N}{y} (1 - q)^y q^{N-y}, \text{ where } y = 0, 1, \cdots, N$$

Let $X_t$ denote the number or fraction of agents who adopt the new technology at stage $t$ and $k_{i,t}$ denote agent $i$'s observation on $X_{t-1}$. Since the observation can vary over agents, we need index $i$ on the individual’s observation. Assume that each adoption decision is observable with probability $\pi$. Each agent can observe an individual’s decision with probability $\pi$ and cannot with probability $1 - \pi$. Then, the probability of observing $k_{i,t}$ adoptions also follows the binomial distribution, where $k_{i,t}$ is the total observation of agent $i$ at stage $t$. The probability that agent $i$ observes $k_{i,2}$ adoptions at stage 2 when $X_1$ early adopters adopted the new innovation is,

$$f(k_{i,2}; X_1) = \binom{X_1}{k_{i,2}} \pi^{k_{i,2}} (1 - \pi)^{X_1 - k_{i,2}}, \text{ where } k_{i,2} = 0, 1, \cdots, X_1$$

On the other hand, if the agent observes a certain amount of decisions, then he can infer the actual amount of adoption using his observation and known distributions.
The probability of the actual adoptions will be negative distribution conditional on the observation, since the actual number of adoptions cannot be lower than the observation. In addition, the upper bound is truncated by the number of informed agents because only the informed early adopters with signal \( h \) have incentive to adopt the new technology. Hence, the probability that the actual adoption is \( X_1 \), given that \( k_{i,2} \) observed is,

\[
f(X_1; k_{i,2}) = \frac{\binom{X_1 - 1}{k_{i,2} - 1} \pi^{k_{i,2}} (1 - \pi)^{X_1 - k_{i,2}}}{\sum_{X=N}^{X_1} \binom{X - 1}{k_{i,2} - 1} \pi^{k_{i,2}} (1 - \pi)^{X - k_{i,2}}}, \text{ where } X_1 = k_{i,2} = k_{i,2}, \cdots, N
\]

**Lemma 3.1.** (Conditional Binomial) If \( X_1 \) follows Bin\((N, q)\), and conditional on \( X_1 \), \( k_{i,2} \) follows Bin\((X_1, \pi)\), then \( k_{i,2} \) follows Bin\((N, \pi q)\).

By lemma 3.1, in state \( H \), \( k_{i,2} \) follows Bin\((N, \pi q)\) because \( \pi \) is the probability of observing each action and \( q \) is the probability of receiving the signal \( h \) in the state \( H \). Hence,

\[
f(k_{i,2}; H) = \binom{N}{k_{i,2}} (\pi q)^{k_{i,2}} (1 - \pi q)^{X_1 - k_{i,2}}, \text{ where } k_{i,2} = 0, 1, \cdots, N
\]

On the other hand, because in the state \( L \), the probability of receiving the signal \( h \) is \( 1 - q \), \( k_{i,2} \) follows Bin\((N, \pi (1 - q))\). Hence,

\[
f(k_{i,2}; L) = \binom{N}{k_{i,2}} (\pi (1 - q))^{k_{i,2}} (1 - \pi (1 - q))^{X_1 - k_{i,2}}, \text{ where } k_{i,2} = 0, 1, \cdots, N
\]
3.3 Adoption Decision

3.3.1 Bayesian Updates

Information set of early adopters at stage 1 depends on which signals they received. An informed agent updates his beliefs about the return being high by the signal he receives. Let $P^H_1$ and $P^L_1$ be the posterior probabilities that the return from adoption is high if a signal $h$ or $l$ is received, respectively, given the prior beliefs, $1/2$. From Bayes’ rule, we have

$$P^H_1 = \frac{1}{2}q + \frac{1}{2}(1 - q) = q \quad (3.1)$$

and

$$P^L_1 = 1 - q \quad (3.2)$$

Since $q > 1/2$, we have $1/2 \leq P^H_1 \leq 1$ and $0 \leq P^L_1 \leq 1/2$. On the other hand, uninformed agents receive no signal at stage 1 and thus, their posterior beliefs are simply the same with the prior beliefs, $P_1 = 1/2$.

The remaining agents update their beliefs about the return from adoption being high by observing the actions taken by the informed agent. Suppose that the remaining agents believe that informed agents follow the signal they received even though the observation is not perfect. The remaining agent observes $k_{i,2}$ which could be different from the actual amount of adoption and they update their beliefs according to their observations.

$$P_{i,2}(k_{i,2}) = \frac{\frac{1}{2}f(k_{i,2}; H)}{\frac{1}{2}f(k_{i,2}; H) + \frac{1}{2}f(k_{i,2}; L)}$$

$$= \frac{(\pi q)^{k_{i,2}}(1 - \pi q)^{N-k_{i,2}}}{(\pi q)^{k_{i,2}}(1 - \pi q)^{N-k_{i,2}} + (\pi (1 - q))^{k_{i,2}}(1 - \pi (1 - q))^{N-k_{i,2}}} \quad (3.3)$$
This is the probability of observing $k_{i,2}$ in high state divided by the probability of observing $k_{i,2}$ in high or low state. Because their beliefs at stage 1 depend on signals they received, the different types of agents have the different prior beliefs at stage 2.

**Lemma 3.2.** *(Monotone Likelihood Ratio)* Because $q$ is higher than $1/2$ and $f(k_{i,2}; X_1)$ satisfies monotone likelihood ratio in $k_{i,2}$, $P_{i,2}$ is increasing in $k_{i,2}$.

This lemma means that the higher observation indicates the more favorable posterior belief on the return from the new technology and thus, we can get cutoff-observed adoption depending on cutoff-belief.

Now, at stage 3, agents can observe the fraction of adoption at stage 2. Since agents can observe each adoption with probability $\pi$ and if a continuum of agents make decisions at stage 2, $k_{i,3} = \pi X_2$ for all $i$ by law of large number. Hence, they can infer the exact fraction of adoption at stage 2 and if we consider the cutoff-belief equilibrium, then the adopters who observe more than the cutoff will adopt the new technology at stage 2. Therefore, all the remaining agents at stage 3 can infer $X_1$ from their observations, $X_2 = 1 - F(\bar{K}_2; X_1)$, where $\bar{K}_2$ is the cutoff-observed adoption at stage 2. At stage 3, all the agents can infer the same $X_1$ and thus the same type of agents will take the same actions.

### 3.3.2 Expected Returns

The construction of equilibrium relies on the expected returns from adoption decisions. The equilibrium strategies of adopters can be constructed accordingly. Because the decision stages are finite, we can solve the equilibrium with backward induction.
Expected returns from adopting the new technology are based on the posterior beliefs. Hence, if they adopt the new technology at stage 3, then the expected return is simply equal to the posterior belief.

\[
EV(x_{i,3} = 1; h_3) = P_3 = \frac{1}{1 + (\frac{1-q}{q})^{2X_1-N}} \tag{3.4}
\]

On the other hand, if they keep the old technology at stage 3, then the expected return \( r \) is certainly provided.

\[
EV(x_{i,3} = 0; h_3) = r \tag{3.5}
\]

We will show that there exists an equilibrium in which the uninformed majority make their decisions according to the following simple rule:

\[
x_{i,3}^U = \begin{cases} 
1 & \text{if } X_1 \geq \bar{K}_3 \\
0 & \text{otherwise}
\end{cases}, \text{ where } \bar{K}_3 = \frac{1}{2} \left[ \frac{ln(\frac{r}{1-r})}{ln(\frac{q}{1-q})} + N \right] \tag{3.6}
\]

\( \bar{K}_3 \) is the cutoff in terms of observation at which uninformed agents are indifferent between adopting the new technology and keeping the old one at stage 3. Since their beliefs at stage 3 depend on their own types, the cutoff varies over their types.

\[
x_{i,3}^H = \begin{cases} 
1 & \text{if } X_1 \geq \bar{K}_3^H \\
0 & \text{otherwise}
\end{cases}, \text{ where } \bar{K}_3^H = \frac{1}{2} \left[ \frac{ln(\frac{r}{1-r})}{ln(\frac{q}{1-q})} + N \right] - 1 \tag{3.7}
\]

\[
x_{i,3}^L = \begin{cases} 
1 & \text{if } X_1 \geq \bar{K}_3^L \\
0 & \text{otherwise}
\end{cases}, \text{ where } \bar{K}_3^L = \frac{1}{2} \left[ \frac{ln(\frac{r}{1-r})}{ln(\frac{q}{1-q})} + N \right] \tag{3.8}
\]

Since the high signal agents already have their own signals \( h \), they need one less high signal to adopt the new technology than the uninformed agents. On the other
hand, the low signal agents do not have the signal \( h \) and thus, they need the same amount of observation with the uninformed agents.

At stage 2, the agent observes adoption decisions of the informed agents at stage 1, but the observation is always smaller than the actual amount of adoption, \( k_{i,2} \leq X_1 \) for all \( i \). Therefore \( \bar{K}_2 \leq \bar{K}_3 \), the cutoff at stage 2 cannot be greater than that at stage 3. If an agent observes more than \( \bar{K}_3 \) at stage 2, then \( X_1 \) is definitely greater than \( \bar{K}_3 \). Hence, he has no incentive to wait.

If an agent adopts the new technology at stage 2, the expected return is twice of posterior beliefs since he will obtain dividends at stage 2 and 3.

\[
EV(x_{i,2} = 1; h_2) = 2P_{i,2} = \frac{2}{1 + (1-q)^{k_{i,2}}(\frac{1-(1-q)\pi N-k_{i,2}}{1-q\pi X_1})} \tag{3.9}
\]

On the other hand, if he waits, then the agent receives the dividend from the old technology at stage 2. At stage 3, he will keep the old technology at stage 3 when \( X_1 \) is smaller than \( \bar{K}_3 \) and he will obtain \( r \). However, when \( X_1 \) is higher than \( \bar{K}_3 \), he will adopt the new technology and receive the posterior belief depending on \( X_1 \).

\[
EV(x_{i,2} = 0; h_2) = r(1 + \sum_{X_1 = k_{i,2}}^{K_3} \left( \frac{X_1 - 1}{k_{i,2} - 1} \right) \pi^{k_{i,2}} (1 - \pi)^{X_1-k_{i,2}})
= \sum_{X_1 = k_{i,2}}^{K_3} \left( \frac{X_1 - 1}{k_{i,2} - 1} \right) \pi^{k_{i,2}} (1 - \pi)^{X_1-k_{i,2}}
+ \sum_{X = k_{i,2}}^{N} \left( \frac{X - 1}{k_{i,2} - 1} \right) \pi^{k_{i,2}} (1 - \pi)^{X-k_{i,2}}
\]

\[
1 + (1-q)^K_2(\frac{1-(1-q)\pi N-K_2}{1-q\pi}) \tag{3.10}
\]

Hence, we can find the cutoff for stage 2 and if the observation is greater than \( \bar{K}_2 \), then he will adopt the new technology while he will wait otherwise.

\[
x_{i,2} = \begin{cases} 
1 & \text{if } k_{i,2} \geq \bar{K}_2 \\
0 & \text{otherwise}
\end{cases}
\]

where \( \bar{K}_2 \) solves \( 2 \left[ 1 + (1-q)^K_2(\frac{1-(1-q)\pi N-K_2}{1-q\pi}) \right] \)
\[ r(1 + \sum_{X_1 = \bar{K}_2}^{X_1 = \bar{K}_3} \left( \frac{X_1 - 1}{k_{i,2} - 1} \right) \pi^{k_{i,2}}(1 - \pi)^{X_1 - k_{i,2}} \) \\
= r \left( 1 + \sum_{X_1 = \bar{K}_2}^{X_1 = \bar{K}_3} \left( \frac{X_1 - 1}{k_{i,2} - 1} \right) \pi^{k_{i,2}}(1 - \pi)^{X_1 - k_{i,2}} \right) \]

\[ + \sum_{X_1 = \bar{K}_2}^{X_1 = \bar{K}_3} \left( \frac{X_1 - 1}{k_{i,2} - 1} \right) \pi^{K_2}(1 - \pi)^{X_1 - K_2} \frac{1}{1 + \left( \frac{1-q}{q} \right)^{2X - N}} \]  

(3.11)

At stage 1, if an agent adopts the new technology, he will obtain the dividend from the new technology during three stages, and thus expected return is three times of the posterior belief at stage 1 depending on the signal he received.

\[ EV(x_{i,1} = 1; h_1) = 3P_1 \]  

(3.12)

However, if he waits, then one of the three mutually exclusive situations arises. If the number of early adopters is below the cutoff for stage 2, then the observation at stage 2 is definitely lower than the cutoff for stage 3, no one adopts the new technology at stage 2 and 3. If the number of early adopters is above the cutoff for stage 3, then the majority adopt the new innovation at stage 2 or 3. However, if the number of early adopters is between cutoffs for stage 2 and 3, the majority adopt the new innovation at stage 2 or do not adopt.

1. If \( X_1 \in [0, \bar{K}_2) \), then \( k_{i,2} < \bar{K}_2 < \bar{K}_3 \Rightarrow x_{i,2} = 0 \) and \( x_{i,3} = 0 \)

2. If \( X_1 \in [\bar{K}_2, \bar{K}_3) \)

   (a) and \( k_{i,2} \in [0, \bar{K}_2) \), then \( x_{i,2} = 0 \) and \( x_{i,3} = 0 \)

   (b) and \( k_{i,2} \in [\bar{K}_2, X_1) \), then \( x_{i,2} = 1 \)

3. If \( X_1 \in [\bar{K}_3, N] \)
(a) and \( k_{i,2} \in [0, \bar{K}_2] \), then \( x_{i,2} = 0 \) and \( x_{i,3} = 1 \)

(b) and \( k_{i,2} \in [\bar{K}_2, X_1] \), then \( x_{i,2} = 1 \)

The expected return from keeping the old technology at stage 1 should be considered all these possible situations. Hence, it is the following.

\[
EV(x_1 = 0; h_1) = r + 2r \sum_{X_1 = 0}^{K_2} \sum_{k_{i,2} = 0}^{K_2} [P_1 h(X_1; S = H) + (1 - P_1) h(X_1; S = L)] f(k_{i,2}; X_1)
\]

\[
+ 2r \sum_{X_1 = K_2}^{K_3} \sum_{k_{i,2} = 0}^{K_2} [P_1 h(X_1; S = H) + (1 - P_1) h(X_1; S = L)] f(k_{i,2}; X_1) P_{i,2}(k_{i,2})
\]

\[
+ \sum_{X_1 = K_3}^{N} \sum_{k_{i,2} = 0}^{K_2} [P_1 h(X_1; S = H) + (1 - P_1) h(X_1; S = L)] f(k_{i,2}; X_1) [r + P_3(X_1)]
\]

\[
+ 2 \sum_{X_1 = K_3}^{N} \sum_{k_{i,2} = K_2}^{K_2} [P_1 h(X_1; S = H) + (1 - P_1) h(X_1; S = L)] f(k_{i,2}; X_1) P_{i,2}(k_{i,2})
\]  

(3.13)

3.3.3 A Perfect Bayesian Equilibrium

Since it is increasing in \( r \), we can find the upper bound \( \bar{r} \) s.t. \( EV(x_1^H = 0; h_1, \bar{r}) \leq 3P_1^H \) and the lower bound \( \underline{r} \) s.t. \( EV(x_1^0 = 0; h_1, \underline{r}) > 3P_1^0 \). For \( r \in (\underline{r}, \bar{r}) \), \( x_1^H = 1, x_1^L = 0, \) and \( x_1^0 = 0 \).

**Proposition 3.1.** For \( r \in (\underline{r}, \bar{r}) \), the beliefs and strategies in (3.1) - (3.13) constitute a pure strategy perfect Bayesian Equilibrium.

3.4 Discussion of Equilibrium

The equilibrium in this paper is determined by the four variables; the return from adopting the good quality innovation, the return from the old technology, the number of informed agents, and the probability of observing the individual adoption decision. A simple example can help us to understand the feature of equilibrium. With the cutoff belief condition, variables comprise the threshold beliefs which the informed agents adopt the new technology and below which they keep the old technology.
3.4.1 Example

The parameters related with the equilibrium are as follows: $N = 10$ and $r = 0.5$. Figure 3.1 illustrates how the equilibrium strategy of early adopters with signal $h$ is changed by changes in $\pi$ and $q$. In this figure, in the area below the line adopting the new technology at stage 1 is optimal for high type agents. On the other hand, in the area above the line, waiting is better for them at stage 1. The area of pure strategy equilibrium expands as private signals are more accurate or they are able to observe others’ actions less accurately. When early adopters are confident with their private signals, they are more likely to follow their signals. Moreover, when they can observe others’ actions with higher noise, they are pessimistic about the observation and value their private signals above the possible observation.

If we fix the accuracy of private signals, then we can draw the cutoff in terms of observation for uninformed majority at stage 2. The cutoff is the threshold in which uninformed majority adopt above while they wait below the cutoff. Figure 3.2 shows how the cutoff is changed by $\pi$. This cutoff is weakly increasing in the probability of
observing each adoption decision, $\pi$, which means that the uninformed agents should observe the larger amount of observation to adopt when they are able to the adoption decisions more accurately.

Figure 3.3 describes the expected fraction of uninformed agents who adopt the new technology at stage 2. Expected fraction at stage 2 is the fraction of uninformed agents who observe more than $\bar{K}_2$ depending on the state of nature. It tends to increase in $\pi$, which means that the more agents adopt the new technology when they can observe others’ decisions more correctly. However, there are some regions where the expected fraction is locally decreasing. This implies that the efficiency might be reduced even when the noise is less in observation. The formula for the expected return is the following:

$$EX_2(S = H) = 1 - F(\bar{K}_2; S = H) = \sum_{k_{i,2} = \bar{K}_2}^{N} \binom{N}{k_{i,2}} (\pi q)^{k_{i,2}}(1-\pi q)^{N-k_{i,2}}, \text{ where } k_{i,2} \sim Bin(N, \pi q)$$
As $\pi$ increases, agents are more likely to observe the correct amount of adoption. Figure 3.4 shows that the expected fraction of agents adopting the new technology at stage 2 is increasing from when $\pi = 0.4$ to when $\pi = 0.5$. In this change, $\bar{K}_2$ remains at 3 and thus the expected fraction is simply increasing. Since they can observe others’ actions more accurately and the cutoff remains the same, they easily satisfy the cutoff and adopt the new technology when they observe the same amount of adoption. However in some regions, the higher $\pi$ requires the higher cutoff. $\bar{K}_2$ is weakly increasing in $\pi$. Figure 3.5 describes this case. The cutoff increases to 4 when $\pi = 0.6$. In order for the uninformed agents to adopt the new technology, they have to observe the greater adoption. For this reason, the expected fraction could be locally decreasing. If the higher $\pi$ represents the lower noise, the efficiency might locally decrease even if the noise is reduced.

### 3.5 Conclusion

This paper has provided a model for noisy observation in diffusion of innovations. When there is an uncertainty in quality of new technology, the agents decide whether
Figure 3.4: Binomial Distributions when $\pi = 0.4$ and $\pi = 0.5$ and the Corresponding Cutoffs

Figure 3.5: Binomial Distributions when $\pi = 0.5$ and $\pi = 0.6$ and the Corresponding Cutoffs
to adopt the new technology or not based on their private information. Moreover, they can accumulate further information by observing others’ actions. In this model, the observation is not perfect and thus, the agents cannot observe others’ actions accurately. Agents might observe the less amount of adoption than the actual one. The noisy observation can create the inefficiency which can be represented as the expected fraction of agents who adopt the new technology at stage 2 in high state. When the quality of the new technology is high, it is desirable that all agents adopt the new technology as soon as possible. Moreover, the uninformed majority make their decisions at stage 2. Therefore, the fraction of agents who adopt the new technology at stage 2 can represent the efficiency. In this model, the expected fraction could be locally decreasing even when the noise is reduced and agents can observe others’ decisions more accurately. This is because if they believe that they can observe more correctly, then they should observe more amount of adoption in order to trigger their adoption. In some regions, the cutoff is increasing faster than they observe more correctly and thus, the expected fraction is decreasing.

In this paper, we could characterize the pure strategy Perfect Bayesian Equilibrium where agents adopt the new technology when they observe more than the cutoff at stage 2. On the other hand, they keep the old technology when they observe less than the cutoff. It is worthwhile to introduce a noise in observation and characterize the pure strategy equilibrium. If we introduce a noisy observation, then it is possible that there is a pure strategy Perfect Bayesian Equilibrium.

We have assumed that there are three stages for agents to decide whether to adopt the new technology or keep the old technology. Early adopters can make their decision at the first stage while other majority decide whether to adopt it or not at
the second and the last stages. However, it seems appropriate to introduce the more
general timing to the adoption game. Future work will extend the model to the more
general timing model of adoption game.
Appendix A: Why Are Information Senders Biased?

Proof of Lemma 1.3

Suppose $v \in [\bar{v}_k, \bar{v}_k]$. From lemma 2, the equilibrium action for the decision maker is following a signal from $k$, $x^*(v, k, s) = s$. Then, the expected utility from subscribing to the information sender $k$ is,

$$Eu(y = k|v) = \sum_{\omega} \sum_{s} Pr(\omega) Pr_k(s|\omega) Eu(x = s|v, k, s)$$

$$= \frac{1}{2}[\mu_k + (1 - \lambda_k)]Eu(x = 1|v, k, s = 1) + \frac{1}{2}[(1 - \mu_k) + \lambda_k]Eu(x = -1|v, k, s = -1).$$

From the equation (5), $Eu(x = 1|v, k, s = 1) = Pr_k(\omega = 1|s = 1)(u + v) + Pr_k(\omega = -1|s = -1) = \mu_k(u + v) + (1 - \lambda_k)(-u + v)$. Similarly, $Eu(x = -1|v, k, s = -1) = (1 - \mu_k)(-u - v) + \lambda_k(u - v)$. Therefore, $Eu(y = k|v) = (\mu_k + \lambda_k - 1)u + (\mu_k - \lambda_k)v = \alpha_k u + \beta_k v$. □

Proof of Lemma 1.6

Proof for lemma 1.6 (1). Suppose that they have the same positive $\beta_k$, $\beta_1 = \beta_2 > 0$. If one of information sender has the lower overall accuracy, i.e., $\alpha_i > \alpha_j$, the sender $j$ is simply dominated and it cannot be optimal for the sender $j$. Now, consider the case in which two senders provide the same overall accuracy, $\alpha_1 = \alpha_2 \equiv \alpha$. In this case, both senders provide exactly the same information structure and thus, each sender’s
share is \( S_1 = S_2 = \frac{1}{2} \left( \frac{1}{\beta_1} + \frac{1}{\beta_2} \right) \). Once one of senders deviates to be biased towards the opposite, \( \beta_2 = -\beta_1 \), her share will be \( S'_2 = \frac{1}{2} \left[ \frac{\alpha}{1+\beta_2} \right] = \frac{1}{2} \left[ \frac{\alpha}{1-\beta_1} \right] \).

Since \( S'_2 - S_2 = \frac{\alpha \beta_1}{2(1-\beta_1)(1+\beta_2)} \) is strictly positive, she has an incentive to move to the opposite. Hence, it is not equilibrium for them. It is similar to the case where \( \beta_1 = \beta_2 < 0 \).

Now, suppose that \( \beta_1 = \beta_2 = 0 \) and \( \alpha_1 = \alpha_2 \equiv \alpha \). In this case, \( S_1 = \frac{\alpha}{2} \). If one of senders slightly increases accuracy for one state and decreases for another one, \( \mu'_i = \mu_i + \epsilon \) and \( \lambda'_i = \lambda_i - \epsilon, \epsilon > 0 \), then \( S'_i = \frac{1}{2} \left[ \frac{1}{1-\beta_i} \right] > S_1 \) and thus, it is not equilibrium either.

Proof for lemma 1.6 (3). Suppose that one of senders provide more accurate signals and they are biased towards different directions, \( \alpha_1 > \alpha_2 \) and \( \beta_1 \geq 0 \geq \beta_2 \).

Then, the senders 1’s share is \( S_1 = \frac{1}{2} \left[ \frac{\alpha_1}{1-\beta_1} - \frac{\alpha_2}{\beta_1-\beta_2} \right] \). If the senders 1 changes her bias to \( \beta'_1 = -\beta_1 \) and keeps her overall accuracy at \( \alpha_1 \), then she dominates the senders 2 and her share will be \( S'_1 = \frac{\alpha_1}{2} \left[ \frac{1}{1-\beta_1} + \frac{1}{1+\beta_1} \right] \) which is higher than \( S_1 \). Hence, the higher accuracy sender always has an incentive to change its bias to dominate another and thus, the lower accuracy sender will be dominated. Therefore, all active senders will provide the same overall accuracy. \( \square \)

**Proof of Proposition 1.3 and Remark 1.1**

Suppose that \( K \) is even. From Proposition 1.3, a half of senders have the same positive \( \beta_k \) while another half of them have the same negative \( \beta_k \); \( \beta^*_K = \cdots = \beta^*_{K+1} < 0 < \beta^*_K = \cdots = \beta^*_1 \). Since all senders have the same overall accuracy and two opposite bias, they should have the same profit. Then, the equilibrium reporting strategy for an information sender, \( \{\mu^*_i, \lambda^*_i\} \), whose \( \alpha_i > 1/2 \) and \( \beta_i > 0 \), is determined when
her marginal benefit of increasing $\lambda_i$ is the same with her marginal cost of it. In this case, if a sender $i$ whose $\beta_i > 0$ slightly increases $\lambda_i$, then she can take $\frac{K-2}{2K}$ from senders who are biased towards the same direction and $\frac{\partial S_i}{\partial \lambda_i}$ from the opposite. Hence, $\frac{K-2}{2K} + \frac{\partial S_i}{\partial \lambda_i} = \frac{\partial c_i}{\partial \lambda_i}$ for $i = 1, \cdots, K$. On the other hand, for an information sender whose $\beta_j < 0$, the equilibrium reporting strategy, $\{\mu_j^*, \lambda_j^*\}$, is determined when $\frac{K-2}{2K} + \frac{\partial S_i}{\partial \mu_j} = \frac{\partial c_j}{\partial \mu_j}$ for $j = K + 1, \cdots, K$. Since all have the same overall accuracy, they have the same profit in equilibrium. As long as each sender’s profit is positive, no senders leave the market. If $\pi_i^* \leq 0$, however, one of senders from each segment leaves the market. Now, there are $K - 2$ senders in the market and $\frac{K-2}{2}$ for each segment.

In general, suppose that $N$ is the equilibrium number of senders in the market, $N \leq K$. As long as $\gamma < \frac{N-1}{4N}$, the equilibrium overall accuracy is greater than $1/2$ and the equilibrium reporting strategy satisfies $\frac{N-2}{2N} + \frac{\partial S_i}{\partial \lambda_i} = \frac{\partial c_i}{\partial \lambda_i}$. Hence, the equilibrium reporting strategy for the sender who is biased towards state $1$ is $\{\mu_1^*, \lambda_1^*\} = \{1, 1 + \frac{N}{4(N-2)} - \sqrt{\frac{2N}{N-2} \gamma + \left[\frac{N}{4(N-2)}\right]^2}\}$, and that for the sender who is biased towards state $-1$ is $\{\mu_1^*, \lambda_1^*\} = \left\{1 + \frac{N}{4(N-2)} - \sqrt{\frac{2N}{N-2} \gamma + \left[\frac{N}{4(N-2)}\right]^2}, 1\right\}$. Since each sender’s profit is $\pi_i^* = \frac{N-2}{2N} - \gamma\left[\frac{1+\frac{N}{4(N-2)} - \sqrt{\frac{2N}{N-2} \gamma + \left[\frac{N}{4(N-2)}\right]^2}}{\sqrt{\frac{2N}{N-2} \gamma + \left[\frac{N}{4(N-2)}\right]^2}}\right]$, $N$ senders are in the market as long as the profit is positive. Once their profits turn out to be negative, one from each segment leaves the market and thus, now $N - 2$ senders remain in the market. This process is repeated until only two senders remain in the market. If $\gamma > \frac{N-1}{4N}$, then the market is in the extreme bias region and senders leave the market once $\gamma > \frac{1}{N}$.

$^{27}$Overall accuracy is $1 + \frac{N}{4(N-2)} - \sqrt{\frac{2N}{N-2} \gamma + \left[\frac{N}{4(N-2)}\right]^2}$, which is increasing in $N$, the equilibrium number of senders in the market.

$^{28}$The extreme bias region exists only when $K = 1, \cdots, 4$. 
When $K$ is odd, there must be one more sender in one segment than the other. Since all senders have the same overall accuracy, their costs are also the same. Hence, in the segment with more senders, their fraction of subscribers and profits are smaller than the other’s. As information cost increases, their profits turn to be negative first and thus, an information sender from the segment with more senders leaves the market. Now $K - 1$ senders are active which is even, and the process after then is the same as above.

**Proof of Corollary 1.2**

Consider that two information senders are biased towards 1 and their overall accuracy is higher than $1/2$, i.e., $\alpha_k > 1/2$ and $\beta_k > 0$ for $k = 1, 2$. In this case, their accuracy for state 1 is perfectly accurate, $\mu_k = 1$, so the senders cannot increase it anymore. If one of senders slightly increases accuracy for the other, $\lambda_k \in (\frac{1}{2}, 1)$, then the marginal benefit from this is taking the market share from the same biased sender, $1/4$, and a part of the opposite sender, $\frac{1}{4(1-\lambda_k)}$. Hence, the marginal benefit from this is, $MB_k = \frac{1}{4} + \frac{1}{4(1-\lambda_k)}$. The marginal cost, on the other hand, is simply $\frac{\gamma}{(1-\lambda_k)^2}$. Hence, it is equilibrium when $\mu_k^* = 1$ and $\lambda_k^* = \frac{3}{2} - \sqrt{\frac{1}{4} + 4\gamma}$ for $k = 1, 2$. From Proposition 3.2, the opposite sender also provides the same overall accuracy and thus, $\mu_3^* = \frac{3}{2} - \sqrt{\frac{1}{3} + 4\gamma}$ and $\lambda_k^* = 1$. When the information cost $\gamma$ is greater than $\frac{3}{16}$, then the market is in the extreme bias region and they leave the market as $\gamma$ increases.

**Proof of Lemma 1.9**

It is easy to show that for $v \in (\bar{v}_{ij}, u)$, the equilibrium action is choosing 1 even after he receives two negative signals, $E u(x = 1|s_i = s_j = -1) > E u(x = -1|s_i =$
\[ s_j = 1, \] where \( \bar{v}_{ij} = \frac{\lambda_i \lambda_j - (1 - \mu_i)(1 - \mu_j)}{\lambda_i \lambda_j + (1 - \mu_i)(1 - \mu_j)} u. \) This implies that the extreme decision makers with \( v \in (\bar{v}_{ij}, u) \) do not follow signals they receive but always choose 1. It is similar for \( v \in (-u, \bar{v}_{ij}) \), \( E_u(x = 1|s_i = s_j = 1) < E_u(x = -1|s_i = s_j = 1) \), where \( \bar{v}_{ij} = \frac{(1 - \lambda_i)(1 - \lambda_j) - \mu_i \mu_j}{(1 - \lambda_i)(1 - \lambda_j) + \mu_i \mu_j} u. \)

The interesting part is what happens if two signals contract with each other. For example, if \( s_i = 1 \) and \( s_j = -1 \), then for \( v \in (\hat{v}_{1,-1}, \bar{v}_{ij}) \), he chooses 1, but below \( \hat{v}_{1,-1} \), the equilibrium action is -1, where \( \hat{v}_{1,-1} = \frac{(1 - \lambda_i)\lambda_j - \mu_i(1 - \mu_j)}{(1 - \lambda_i)\lambda_j + \mu_i(1 - \mu_j)} u \). On the other hand, if \( s_i = 1 \) and \( s_j = -1 \), then we have a different cutoff, \( \hat{v}_{-1,1} = \frac{\lambda_i(1 - \lambda_j) - (1 - \mu_i)\mu_j}{\lambda_i(1 - \lambda_j) + (1 - \mu_i)\mu_j} u \) where below the cutoff, the equilibrium action is -1 and vice versa. Hence, the equilibrium action strategy depends on whether \( \hat{v}_{1,-1} \) or not, and it is determined by which one is greater than another between \( \mu_i \lambda_i \) and \( \mu_j \lambda_j \). When \( \mu_i \lambda_i > \mu_j \lambda_j \), \( \hat{v}_{1,-1} \) is greater than \( \hat{v}_{-1,1} \), and vice versa. For \( v \in (\min\{\hat{v}_{-1,1}, \hat{v}_{1,-1}\}, \max\{\hat{v}_{-1,1}, \hat{v}_{1,-1}\}) \), the decision maker simply follows the signal \( s_k \), where \( s_k = \begin{cases} s_i & \text{if } \hat{v}_{1,-1} > \hat{v}_{-1,1} \\ s_j & \text{if } \hat{v}_{-1,1} > \hat{v}_{1,-1} \end{cases} \). \( \square \)
Appendix B: Social Learning with a Ratings System

Proof for Lemma 2.2

Assume that $\lambda(q_n, H)$ is not weakly decreasing in $q_n$. Then, we can find $q_0$ and $q_1$ s.t. $q_0 > q_1$ and $\lambda(q_0, H) > \lambda(q_1, H)$. Since $q_0 > q_1$, we can find a signal with accuracy $\sigma_1 > \frac{1}{2}$ s.t.

$$q_0 = \frac{q_1\sigma_1}{q_1\sigma_1 + (1-q_1)(1-\sigma_1)}.$$  

Let $q_2 \equiv \frac{q_1(1-\sigma_1)}{q_1(1-\sigma_1) + (1-q_1)\sigma_1} < q_1$. Hence, $\lambda(q_1, H) = \sigma_1\lambda(q_0, H) + (1-\sigma_1)\lambda(q_2, H)$. Since $\sigma_1 > \frac{1}{2}$ and $\lambda(q_0, H) > \lambda(q_1, H)$, $\lambda(q_1, H) > \lambda(q_2, H)$ and $q_1 > q_2$. Continuously, $q_1 > q_2 > q_3 > \cdots$ and $\lambda(q_1, H) > \lambda(q_2, H) > \lambda(q_3, H) > \cdots$. Ultimately, $\lambda(q_0, H) > \lambda(0, H)$. But, when $q = 0$, all agents choose $x = -1$ and thus, $\lambda(0, H) = 1$. There is a contradiction.

Proof for Proposition 2.1

Lemma B.1. Given state $H$, $< t^H_n >$ is a martingale stochastic process in the ratings information environment.
Proof for Lemma 2.3

We want to show $E(\ell_{n+1}^H|\ell_n^H, H) = \ell_n^H$ to prove that $<\ell_n^H>$ is martingale. The transition of the likelihood ratio depends on purchase decisions, ratings and the current likelihood ratio. If the likelihood ratio is in the learning region, $(\underline{\ell}_n^H, \overline{\ell}_n^H)$, both purchase decisions and ratings provide the previous buyers’ private information.

$$\ell_{n+1}^H = \begin{cases} \ell_n^H \frac{1 - \sigma}{\sigma} G(-u) & \text{if } x_n = 1 \text{ and } y_n = 1 \\ \ell_n^H \frac{1 - \sigma}{\sigma} G(u) & \text{if } x_n = 1 \text{ and } y_n = -1 \\ \ell_n^H \frac{\sigma}{1 - \sigma} & \text{if } x_n = 0 \text{ and } y_n = 0. \end{cases}$$

$$E(\ell_{n+1}^H|\ell_n^H \in (\underline{\ell}_n^H, \overline{\ell}_n^H), H) = \{\sigma G(u)\left[\frac{1 - \sigma}{\sigma} G(-u)\right] + \sigma G(-u)\left[\frac{1 - \sigma}{\sigma} G(u)\right] + (1 - \sigma)\left[\frac{\sigma}{1 - \sigma}\right]\} \ell_n^H$$

$$= \ell_n^H.$$

When the likelihood ratio is higher than $\overline{\ell}_n^H$, no one purchases the product and thus, no further information is updated to successors. Hence, the likelihood ratio is not changed no matter what actions and ratings are, i.e., $\ell_{n+1}^H = \ell_n^H$ for any $x_n$ and $y_n$. Therefore, $E(\ell_{n+1}^H|\ell_n^H \in (\overline{\ell}_n^H, \infty), H) = \ell_n^H$.

Below $\underline{\ell}_n^H$, all later buyers always purchase the product, so purchase decisions do not provide any private information. In contrast to the previous case, however, buyers still give ratings, which changes the likelihood ratio.

$$\ell_{n+1}^H = \begin{cases} \ell_n^H \frac{G(-u)}{G(u)} & \text{if } x_n = 1 \text{ and } y_n = 1 \\ \ell_n^H \frac{G(u)}{G(-u)} & \text{if } x_n = 1 \text{ and } y_n = -1 \\ \ell_n^H \frac{G(u)}{G(-u)} & \text{if } x_n = 0 \text{ and } y_n = 0. \end{cases}$$

$$E(\ell_{n+1}^H|\ell_n^H \in (0, \underline{\ell}_n^H), H) = [G(u) \frac{G(-u)}{G(u)} + G(-u) \frac{G(u)}{G(-u)}] \ell_n^H = \ell_n^H.$$
Therefore, \( < \ell_n^H > \) is a martingale stochastic process conditional on state \( H \).

Similarly, we can show that \( < \ell_n^L > \) is also a martingale conditional on state \( L \).

**Lemma B.2.** Given state \( H \), \( < \ell_n^H > \) is convergent.

Fatou’s lemma implies that \( < \ell_n^H > \) is convergent since it is martingale and bounded below.

Since the initial likelihood ratio is 1 and \( < \ell_n^H > \) is martingale, we can see that the sum of all possible convergent points times the probabilities that the likelihood ratio converges to the points should be 1.

Define \( \bar{\ell}_i^H \) be a convergent point of the likelihood ratio for \( i = 1, \ldots, I \). Without loss of generality, we can rearrange that \( \bar{\ell}_0^H < \cdots < \bar{\ell}_I^H \).

**Lemma B.3.** \( \sum_{i=0}^I \text{prob}(\bar{\ell}_i^H)\bar{\ell}_i^H = 1 \).

**Proof for Proposition 2.1**

We need to see proposition 3 first to know proof for proposition 2.1. From proposition 2.3, the likelihood ratio converges to the set, \( \{0\} \cup (\bar{\ell}_H, \infty) \). Since the minimum value is zero, we can rewrite Lemma 4 as \( \sum_{i=1}^I \text{prob}(\bar{\ell}_i^H)\bar{\ell}_i^H = 1 \) and reorder the convergent points by

\[
\sum_{i=1}^I \text{prob}(\bar{\ell}_i^H)\bar{\ell}_i^H < \sum_{i=1}^I \text{prob}(\bar{\ell}_i^H)\bar{\ell}_i^H < \sum_{i=1}^I \text{prob}(\bar{\ell}_i^H)\bar{\ell}_i^H,
\]

which implies that \( \frac{1}{\bar{\ell}_I^H} < \sum_{i=1}^I \text{prob}(\bar{\ell}_i^H) < \frac{1}{\bar{\ell}_I^H} \), i.e., the probability of incorrect herds in state \( H \) is bounded in \( (\frac{1}{\bar{\ell}_I^H}, \frac{1}{\bar{\ell}_I^H}) \).

Suppose that ratings information is more accurate than private signals, \( G(u) > \sigma \).

In this case, when all the previous buyers purchase but give negative ratings, the
current likelihood ratio is higher than the initial value. If it is significantly high, then it is possible that the likelihood ratio is greater than $\bar{\ell}_H$ and no one purchases the product. In order to figure out how many negative ratings can lead people to herd on non-purchase, we need to define the relative accuracy between ratings and private signals.

\[ \frac{\sigma}{1 - \sigma} = \left[ \frac{G(u)}{G(-u)} \right]^\gamma, \] where $\gamma$ is the relative accuracy.

When $G(u) > \sigma$, $\gamma$ is less than 1. Now, we can find the minimum number of negative ratings, $t$ which causes herds on non-purchase decisions even when all the previous buyers purchase the product.

In this case, there are 3 possible scenarios where incorrect herds occur in state $H$.

The first one is that two non-purchase decisions are consequent from the prior and the likelihood ratio is \( \frac{1}{2} \sigma (1 + \sigma) \). In the second scenario, there are $t$ number of purchase decisions, but all buyers give negative ratings, where $t \in \left( \frac{\gamma}{1 - \gamma}, \frac{1}{1 - \gamma} \right)$. $t$ is the number of negative ratings which causes incorrect herds for high quality products when purchase decisions have the same number with non-purchase decisions. Since $t$ is the minimum number of negative ratings which prevents buyers from purchasing the high quality products only by themselves, likelihood ratios should be lower than $\bar{\ell}_H$ until $(t - 1)$ negative ratings. It implies that \[ \frac{1 - \sigma}{\sigma} \frac{G(u)}{G(-u)} \] \[ \left[ \frac{1 - \sigma}{\sigma} \frac{G(u)}{G(-u)} \right]^{t-1} < \frac{\sigma}{1 - \sigma}. \] The last scenario is that there are $i$ purchase decisions and negative ratings, and $(i + 1)th$ non-purchase decision, where $i = 1, \cdots, t - 1$. Then, the likelihood ratio is \[ \frac{1}{\sigma} \left[ \frac{1 - \sigma}{\sigma} \frac{G(u)}{G(-u)} \right]^{i-1}, \] where $i = 1, \cdots, t - 1$. When there are more non-purchase decisions or negative ratings, likelihood ratios are already greater than $\bar{\ell}_H$ and incorrect herds are in progress. Hence, we can simplify
all incorrect herding situations in state $H$ to three scenarios mentioned above. Since
\[
\frac{G(u)}{G(-u)} > \frac{1}{1 - \sigma}, \left[\frac{1 - \sigma}{\sigma}\right]^{t-2}\left[\frac{G(u)}{G(-u)}\right]^{t-1} > \left[\frac{1 - \sigma}{\sigma}\right]^{i-1}\left[\frac{G(u)}{G(-u)}\right]^i \text{ for } i = 1, \ldots, t-2, \text{ and }
\]
\[
\frac{G(u)}{G(-u)} < \left[\frac{1 - \sigma}{\sigma}\right]^{i-1}\left[\frac{G(u)}{G(-u)}\right]^i \text{ for } i = 2, \ldots, t-1.
\]

If $G(u) \in (\sigma, \bar{\sigma})$ where $\bar{\sigma} = \min\left\{\frac{\sigma^2}{\sigma^2 + (1 - \sigma)^2}, \frac{(1 - \sigma)^{(t-1)/t}(1 + \sigma)^{1/t}}{\sigma^{(t-1)/t}(2 - \sigma)^{1/t} + (1 - \sigma)^{(t-1)/t}(1 + \sigma)^{1/t}}\right\}$, then $\bar{\ell}_i^H = \frac{1}{2} \frac{\sigma(1 + \sigma)}{(1 - \sigma)(1 - \frac{1}{2}\sigma)}$ and thus, $\lambda_R(H) > \frac{(1 - \sigma)(1 - \frac{1}{2}\sigma)}{\frac{1}{2} \sigma(1 + \sigma)}$.

Since $\sigma$ is lower than both
\[
\frac{\sigma^2}{\sigma^2 + (1 - \sigma)^2} \text{ and } \frac{(1 - \sigma)^{(t-1)/t}(1 + \sigma)^{1/t}}{\sigma^{(t-1)/t}(2 - \sigma)^{1/t} + (1 - \sigma)^{(t-1)/t}(1 + \sigma)^{1/t}},
\]
the range, $(\sigma, \bar{\sigma})$ always exists.

Now, suppose that $G(u) < \sigma$, i.e., ratings are less accurate than private signals. Even if all agents give negative ratings in this case, the public beliefs are still higher than the initial value no matter how many negative ratings are provided. Moreover, herds on non-purchase cannot be caused only by negative ratings even when there are more purchase decisions than non-purchase decisions. There should be at least one more non-purchase decision for herds on non-purchase. Hence, the possible convergent points are,
\[
\frac{1}{2} \frac{\sigma(1 + \sigma)}{(1 - \sigma)(1 - \frac{1}{2}\sigma)} \text{ and } \left[\frac{1 - \sigma}{\sigma}\right]^{i-2}\left[\frac{G(u)}{G(-u)}\right]^i \text{ for } i = 1, \ldots, k \text{ where } \gamma \in \left(\frac{k + 1}{k}, \frac{k}{k - 1}\right)
\]
and $\gamma < 1$. When $\frac{1 - \sigma}{\sigma} \left[\frac{G(u)}{G(-u)}\right]^2$ and $\frac{\sigma}{1 - \sigma} \frac{G(u)}{G(-u)}$ are smaller than $\frac{1}{2} \frac{\sigma(1 + \sigma)}{(1 - \sigma)(1 - \frac{1}{2}\sigma)}$,
\[
\bar{\ell}_i^H = \frac{1}{2} \frac{\sigma(1 + \sigma)}{(1 - \sigma)(1 - \frac{1}{2}\sigma)}.
\]
Hence, if \( G(u) \in (\sigma, \bar{\sigma}) \) where \( \sigma = \max\{\frac{\sqrt{\sigma}}{\sqrt{\sigma} + \sqrt{1-\sigma}}, \frac{1+\sigma}{3}\} \), then \( \lambda_R(H) < \frac{(1-\sigma)(1-\frac{1}{2}\sigma)}{\frac{1}{2}\sigma(1+\sigma)} \). In sum, \( \lambda_R(H) > \frac{(1-\sigma)(1-\frac{1}{2}\sigma)}{\frac{1}{2}\sigma(1+\sigma)} \) when \( G(u) \in (\sigma, \bar{\sigma}) \), and \( \lambda_R(H) < \frac{(1-\sigma)(1-\frac{1}{2}\sigma)}{\frac{1}{2}\sigma(1+\sigma)} \) when \( G(u) \in (\sigma, \sigma) \). Therefore, the probability of incorrect herds is higher in this case than that in the previous one, so \( \lambda_R(H) \) is not monotonic in \( G(u) \).

**Proof for Proposition 2.2**

The proof for part 1 of proposition 2.2 is trivial. Since ratings can stop consumers purchasing low quality products, incorrect herds do not occur with ratings almost surely and the probability of incorrect herds with ratings is clearly lower than that without ratings.

For proof of part 2, suppose that the state is \( H \) and ratings information is more accurate than private signals, \( G(u) > \sigma \). Our assumption on the prior is simply \( \frac{1}{2} \), \( q_n = \frac{1}{2} \). Then the public beliefs of the next period after observing a purchase decision and a negative rating is lower than the prior, \( q_{n+1}(x_n = 1, y_n = -1) < \frac{1}{2} \). Since we assume that all buyers must give ratings while non-buyers do not, three exclusive situations could occur at period \( n \); the agent \( n \) purchases the product and give a positive rating, he purchases the product but give a negative rating, and he does not purchase the product. The probabilities for the situations are respectively, \( \sigma G(u) \), \( \sigma G(-u) \) and \( (1-\sigma) \). Hence, we can write the probability of incorrect herds in state \( H \) as the following:

\[
\lambda_R(H) = \sigma G(u)\lambda_R(\mu_R(x_n = 1, y_n = 1), H) + \sigma G(-u)\lambda_R(\mu_R(x_n = 1, y_n = -1), H) + (1-\sigma)\lambda_R(\mu_R(x_n = 0, y_n = 0), H).
\]
1. Since $\lambda_R$ is the probability of incorrect herds, it is bounded from below, $\lambda_R(q_{n+1}(x_n = 1, y_n = 1), H) \geq 0$.

2. From Lemma 1, we know that the probability of incorrect herds is weakly decreasing in public beliefs conditional on the state $H$. In addition, $q_{n+1}(x_n = 1, y_n = -1) < \frac{1}{2}$. Hence, $\lambda_R(q_{n+1}(x_n = 1, y_n = -1), H) > \lambda_R(H)$.

3. $\lambda_R(q_{n+1}(x_n = 0, y_n = 0), H) = (1 - \frac{1}{2}\sigma)\lambda_R(q_{n+1}((x_n = 0, y_n = 0), (x_{n+1} = 0, y_{n+1} = 0), H) + \frac{1}{2}\sigma G(u)\lambda_R(q_{n+1}((x_n = 0, y_n = 0), (x_{n+1} = 0, y_{n+1} = 0), H) + G(-u)\lambda_R(q_{n+1}((x_n = 0, y_n = 0), (x_{n+1} = 1, y_{n+1} = -1), H)]$.

4. If the first two agents do not purchase the product, then no ratings information is provided and thus, the likelihood ratio, $\ell_{n+2}$ will be outside of the learning region. Hence, learning stops and incorrect herds occur: $\lambda_R(q_{n+2}((x_n = 0, y_n = 0), (x_{n+1} = 0, y_{n+1} = 0)), H) = 1$.

5. For any actions or ratings observed, the probability of incorrect herds is still non-negative: $\lambda_R(q_{n+2}((x_n = 0, y_n = 0), (x_{n+1} = 1, y_{n+1} = 1)), H), H) \geq 0$.

6. Since $G(u) > \sigma$, $q_{n+2}$ is lower than $q_n$ after observing one non-purchase decision and one purchase decision with negative rating. $\lambda_R(q_{n+2}((x_n = 0, y_n = 0), (x_{n+1} = 1, y_{n+1} = -1)), H), H) \geq \lambda_R(H)$.

Hence in sum, $[1 - \sigma G(-u) - \frac{1}{2}\sigma(1 - \sigma)G(-u)]\lambda_R(H) > (1 - \sigma)(1 - \frac{1}{2}\sigma) + (1 - \sigma)\frac{1}{2}\sigma G(-u)$, which implies

$$\lambda_R(H) > \frac{(1 - \sigma)(1 - \frac{1}{2}\sigma) + (1 - \sigma)\frac{1}{2}\sigma G(-u)}{[1 - \sigma G(-u) - \frac{1}{2}\sigma(1 - \sigma)G(-u)]}.$$
Since \( \lambda_{NR}(H) = \frac{(1 - \sigma)(1 - \frac{1}{2}\sigma)}{1 - \sigma + \sigma^2} \), if \( \frac{(1 - \sigma)(1 - \frac{1}{2}\sigma) + (1 - \sigma)\frac{1}{2}\sigma G(-u)}{[1 - \sigma G(-u) - \frac{1}{2}\sigma(1 - \sigma)G(-u)]} > \frac{(1 - \sigma)(1 - \frac{1}{2}\sigma)}{1 - \sigma + \sigma^2} \), then

\[ \lambda_R(H) > \lambda_{NR}(H). \]

When \( G(u) \) is greater than \( \frac{4 - \sigma + \sigma^2}{8 - 7\sigma + 3\sigma^2} \), the above inequality is satisfied. Therefore, if \( G(u) \in (\sigma, \frac{4 - \sigma + \sigma^2}{8 - 7\sigma + 3\sigma^2}) \), then \( \lambda_R(H) > \lambda_{NR}(H) \).

Since \( \sigma \leq \frac{4 - \sigma + \sigma^2}{8 - 7\sigma + 3\sigma^2} \) for all \( \sigma \in \left(\frac{1}{2}, 1\right) \), the region, \((\sigma, \frac{4 - \sigma + \sigma^2}{8 - 7\sigma + 3\sigma^2})\), always exists.

**Proof for Proposition 2.3**

1. In the state \( H \), we consider the likelihood ratio \( \ell_n^H \).

   (1) Case 1: \( \ell_n^H \in [\underline{\ell}^H, \bar{\ell}^H] \)

   In the learning region, \( \ell_n^H \in (\underline{\ell}^H, \bar{\ell}^H) \) and all actions and ratings by buyers deliver their private signals and ex-post payoffs information. The transition of the likelihood ratio is the following:

   \[
   \ell_{n+1}^H = \begin{cases} 
   \frac{\ell_n^H (1 - \sigma) G(-u)}{\sigma G(u)} & \text{if } x_n = 1 \text{ and } y_n = 1 \\
   \frac{\ell_n^H (1 - \sigma) G(u)}{\sigma G(-u)} & \text{if } x_n = 1 \text{ and } y_n = -1 \\
   \ell_n^H \frac{1 - \sigma}{1 - \sigma} & \text{if } x_n = 0 \text{ and } y_n = 0.
   \end{cases}
   \]

   Suppose that learning stops in the region, \( \ell_\infty^H \in (\underline{\ell}^H, \bar{\ell}^H) \), no matter what the agent \( n \)'s choices and ratings are. However, \( \ell_\infty^H \) is equal to \( \ell_n^H \) only when \( \ell_\infty^H = 0 \) which is out of the learning region, contradicting that the likelihood ratio is changed when the later agents observe any actions or ratings. Therefore, the likelihood ratio does not converge into the learning region. This result is similar to the cases when \( \ell_n^H = \underline{\ell}^H \) or \( \ell_n^H = \bar{\ell}^H \)
(2) Case 2: \( \ell^H_n \in (\bar{\ell}^H, \infty) \)

When the likelihood ratio is outside of the learning region, no one purchases the product and thus, ratings information is no longer provided. Therefore, \( \ell^H_{n+1} = \ell^H_n \) for any \( x_n \) and \( y_n \), and learning stops in this region.

(3) Case 3: \( \ell^H_n \in [0, \bar{\ell}^H) \)

In this region, all agents purchase the product and thus, purchase decisions do not deliver any private information. The transition of the likelihood ratio is the following:

\[
\ell^H_{n+1} = \begin{cases} 
\ell^H_n \frac{G(u)}{G(-u)} & \text{if } x_n = 1 \text{ and } y_n = 1 \\
\ell^H_n \frac{G(-u)}{G(u)} & \text{if } x_n = 1 \text{ and } y_n = -1.
\end{cases}
\]

If \( \ell^H_n \in [0, \bar{\ell}^H) \), \( \ell^H_{n+1} = \ell^H_n \) only when they are zero. It implies that the likelihood ratio converges to zero and learning stops.

In sum, \( \ell^H_n \) converges almost surely to a random variable \( \ell^H_\infty \) with support in the set \( \{0\} \cup (\bar{\ell}^H, \infty) \), given that the state is \( H \).

2. In the state \( L \), we consider the likelihood ratio \( \ell^L_n \). Similarly, the likelihood ratio does not converge in the learning region. However, the difference occurs in \((\bar{\ell}^L, \infty)\) where \( \bar{\ell}^L = \frac{1}{\ell^H} \). This region corresponds to \( \ell^H_n \in [0, \bar{\ell}^H) \) in state \( H \) in which all agents purchase the product. The transition of the likelihood ratio is the following:

\[
\ell^L_{n+1} = \begin{cases} 
\ell^L_n \frac{G(u)}{G(-u)} & \text{if } x_n = 1 \text{ and } y_n = 1 \\
\ell^L_n \frac{G(-u)}{G(u)} & \text{if } x_n = 1 \text{ and } y_n = -1.
\end{cases}
\]

In the transition, \( \ell^L_{n+1} = \ell^L_n \) only when they are zero which is not included in \((\bar{\ell}^L, \infty)\). Hence, the likelihood ratio does not converge into \((\bar{\ell}^L, \infty)\), but only into \([0, \bar{\ell}^L)\) where \( \bar{\ell}^L = \frac{1}{\ell^H} \).
Bibliography


[37] Levin, Dan and Peck, James, “Investment dynamics with common and private values,” *Journal of Economic Theory*, 143, 114-139.


