Title
Essays on Information Economics: Information Markets, Social Learning, and Information Design

Author
Ma, Yingju

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Essays on Information Economics:
Information Markets, Social Learning, and Information Design

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Economics

by

Yingju Ma

2019

ABSTRACT OF THE DISSERTATION

Essays on Information Economics:
Information Markets, Social Learning, and Information Design

by

Yingju Ma
Doctor of Philosophy in Economics
University of California, Los Angeles, 2019
Professor Ichiro Obara, Chair

This dissertation studies three topics in information economics. In Chapter One, “Monopoly and Competition in the Markets for Information”, I analyze the generation and provision of information products, and the implications of competition in these markets. In the model, buyers face a decision problem with uncertainty about the states of the world. A buyer can purchase experiments to augment his private information, therefore the value of an experiment depends on his private information. To generate these experiments, sellers have to make an investment, which determines the most informative experiment a seller can provide. Sellers then post menus of experiments and prices. I first characterize the optimal menu given any investment level and derive the optimal investment. When two sellers compete with investment, I find an equilibrium in which two sellers split the market: one seller only serves to high belief buyers and the other serves to low beliefs buyers. Each seller specializes in generating a more informative signal about one state. Monopoly seller always provides more informative experiments, and to more buyers, than the case of duopoly competition.

In Chapter Two, “Preferential Attachment as an Information Cascade in Emerging Networks”, I study the preferential attachment observed in real-world social networks as a social learning problem. Networks grown via preferential attachment exhibit the ”rich-get-richer” phenomena; nodes with higher connectivity degree are more likely to acquire more con-

nections. This chapter develops a Bayesian social learning model in which agents arriving sequentially to a network judge the qualities of predecessor agents based on their own private signals, and on public signals inferred from the observed network structure. It shows that preferential attachment emerges endogenously as a sequentially equilibrium of the social learning process, where agents may engage in a rational herd behavior. The condensed preferential attachment, in which one agent gets all the future links, emerges with probability one when the private signals are bounded.

In Chapter Three, "Information Design in Contests", I consider the information disclosure problem in contests. The designer of a contest has an informational advantage over agents' ability. There is a strong agent (res. weak agent) who will has a higher probability of being a high ability (res. low ability) player. In the optimal information disclosure policy, the designer discriminates two types of agents. When the weak agent has a disadvantage in abilities, the designer will partially disclose the state to him privately. Compared with the no-disclosure benchmark, the optimal policy increases the total effort level. On the other hand, committing to a public message disclosure can not improve the equilibrium of the no-disclosure benchmark.

The dissertation of Yingju Ma is approved.

Sushil Bikhchandani

Marek G. Pycia

Joseph M. Ostroy

Simon Adrian Board

Ichiro Obara, Committee Chair

University of California, Los Angeles

2019

To Jia.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am fortunate to have Ichiro Obara as my advisor. He has taught me more than I could ever hope to learn. He is always available for insightful discussions about my work at all stages. I am very grateful to Joseph Ostroy for his constant encouragement and support. I owe them a huge debt, which I only hope to pay back by continuing to contribute to economic theory by my own.

As many thanks to Simon Board, Marek Pycia, and Moritz Meyer-ter-Vehn, from whom I learned how to craft a meaningful model, how to combine mathematical reasoning with economic intuition, and how to appreciate the beauty of economic theory.

I also benefited from other faculty members and fellow students at UCLA. In particular, I would like to thank Sushil Bikhchandani, Tomasz Sadzik, Bill Zame, Jernej Copic, Jay Lu, Kyle Woodward, Kenneth Mirkin, Jen-Wen Chang, Cong Xie, Daehyun Kim, Xinyu Fan, and Zhuoran Lu for their comments and suggestions about my work.

I thank my co-authors, Daehyun Kim, Mihaela van der Schaar, and Ahmed M. Alaa, for allowing me to use our joint works in this dissertation. It was a great experience to work with them.

I thank my cohorts, UCLA Econ 2012 and CCER 2009, and many other friends, for the encouragement and sharing the unique experience together.

My previous advisors in Peking University, Ho-Mou Wu and Dingding Wang, also contribute to my dissertation in their continuous guidance and critical suggestions in the last ten years. They are both the most outstanding Chinese scholars and it is my honor to be a student of them.

I want to thank my parents and parents-in-law for their support. Finally, none of this would be possible if not for the love and partnership from Jia Yu, my dearest friend of life.

VITA

2009	B.S. (Mathematics), Department of Mathematics, Zhejiang University.
2012	M.A. (Economics), National School of Development, Peking University.
2014	M.A. (Economics), Department of Economics, UCLA.

2013–2017	Teaching Assistant, Department of Economics, UCLA.
2015	Instructor, UCLA Extension.
2017–2018	*leave of absence.*
2018–2019	Teaching Fellow, Department of Economics, UCLA.

2012–2017	Graduate Fellowship, Department of Economics, UCLA
2018–2019	Graduate Fellowship, Department of Economics, UCLA.

CHAPTER 1

Monopoly and Competition in the Markets for Information

1.1 Introduction

The markets for information play an increasing role in the economy with the increasing frequency of online activities[1]. The value of information can only be exploited when it helps people making better decisions. Information about borrowers helps banks setting proper credit limits; information about consumers' browsing preferences guides online advertisers. The volume of data generated, however, is so large that the cost of collecting and managing databases, developing and carrying out algorithms, and giving meaningful statistical results, becomes be very high. When more than one sellers compete in such a market for information, it is not clear whether competition leads to a better informed outcome or not. In traditional markets, competition among sellers usually yields higher quality and lower price, therefore makes buyers better off. However, when it requires non-negligible cost to generate informative products, competition might weaken sellers' incentive to invest, due to the concern of a decrease in market share.

In this paper, we develop a framework to analyze the generation and competition in the markets of information products. Buyers face a decision problem with uncertainty of two states. They can purchase additional information (experiments) to augment their private information. A buyer's willingness to pay for an experiment depends on his private information. If a buyer is very certain about one state, he does not value additional information very

[1] Bergemann and Bonatti (2018) summarize a number of markets for information products and current research about these markets. They also provide a framework to categorize different information products.

much. On the other hand, if a buyer is completely unsure about the states, any information has some value to him.

To generate these experiments, sellers have to make an investment, which determines the most informative experiment a seller can provide. Once the investment is made, the seller can provide the most informative experiment to any number of buyers with zero marginal cost. And seller can also babbling on this experiment and provide any less informative experiment with no cost. Since buyers' value for an information product is private, we assume the seller post a menu of experiments and prices. Buyers then choose one experiment from the menu. When there are two sellers, they compete with investment.

The contribution of this paper is that: (i) we characterize the optimal menu given any investment level and derive the optimal investment; (ii) we find an asymmetric equilibrium of market segmentation when two sellers compete with investment; (iii) by comparing the information generated and provided by monopoly and duopoly sellers, we find that monopoly always provides more informative experiments and serves more buyers than the case of duopoly competition.

One special feature of information products is that the value of information depends on both the quality of information and the buyers' beliefs. Consider an experiment sending two signals (bad news and good news) about two states (good state and bad state). The more accurate both signals are, the higher the value is for all buyers. Thus it has a vertical element (quality). On the other hand, for any two experiments, different buyers may have opposite preferences. Pessimistic buyers prefers the experiment giving more accurate good news, since it is more valuable to acquiring knowledge about the state he is less confident at. While optimistic buyers prefers more accurate bad news. So the value of information also has a horizontal element (position).

Since the buyers' beliefs are unobservable, sellers need to screen among buyers buy providing an incentive compatible menu, which includes some experiments and the corresponding prices. Our first result (Proposition 1) gives a characterization of the optimal menu. Given any investment level, the optimal menu includes at most two experiments. It always includes the most informative experiment feasible to the seller. When there are more pessimistic buyers, the seller might provide additional experiment targeting to some of the pessimistic buyers whom otherwise are excluded from the market.

When two sellers compete in the market with investments, out second main result (Proposition 3) is that there exists an equilibrium where one seller specializes in producing good news, and only serves to the pessimistic buyers. While the other seller serves optimistic buyers with more accurate bad news. Pessimistic buyers strictly prefer the good news seller since it gives them higher improvement in decision making problem. Even if two sellers have the same investment level, we have this market segmentation.

Our answer to the question whether competition lead to better informed outcome is the third result (Theorem 1). Monopoly make higher investment and provide more informative experiment than duopoly. First because duopoly both sellers have to invest to generate informative products and they split the market. So the benefit from investment is less than the case of monopoly. However, since duopoly seller faces a submarket with more concentrated beliefs (either pessimistic buyers or optimistic buyers), they can specialize in providing the signals which favorite by the submarket. We show that the first factor becomes dominate in equilibrium.

In Section 2 we present the main settings of the model. Then we analyze the optimal menu and optimal investment of a monopoly seller in Section 3. In Section 4 we consider case of duopoly competition, and compare the optimal investment between monopoly and duopoly sellers.

1.1.1 Related Literature

This paper is part of the literature of selling information to imperfectly informed decision makers. The most related work is Bergemann et al. (2018), which analyzes the problem of a monopoly seller selling information. In their model, the seller can access the complete information. They characterize the revenue-maximizing menu of experiments. Their main results show that the optimal menu includes at most two experiments. The fully informative experiment is always provided. The seller might provide a second partial informative experiment by adding noises on one signal. Based on their framework, we examine the design and pricing problem of a seller whose most informative experiment is limited. Then we consider the optimal investment which determines the most informative experiment, and study the case of competing sellers.

A recent paper Bimpikis et al. (2019) studies the information market where the buyers of the information products competition in a downstream market. They find that seller will provide accurate information when buyers actions in the downstream are strategic complements. When the competition in the downstream market is via strategic substitute actions, the seller will either decrease the supply or the quality of the information products. It differs from our model that we directly analyze the competition among information sellers. In seminal papers on selling information, Admati and Pfleiderer (1986, 1990) studies the case when information buyers are ex ante homogenous, and trade an asset after they get supplemental information. They focus on the interactions among data buyers while the heterogeneous buyers face their own decision problems. Kown (2018) considers a selling mechanism when seller first sells an information structure and then sells a product. The information structure and the product are complementary goods. It is beneficial for both sides if buyers and sellers have private information about the products.

This paper also contributes to the literature on information disclosure with competition. Gentzkow and Kamenica (2016) show that when senders sends coordinated signals, additional senders increase the amount of information. Li and Norman (2018) generalize their setting by allowing sequential senders, and show that with independent signals the result might fail. In these settings senders send information to persuade receivers to take an action, while in our model the sender (seller) is providing information for revenue rather than persuasion.

Board and Lu (2018) considers a model where buyers search a product to fit their needs best, and sellers of a product use information disclosure policies to manage the buyers' search incentives. Buyers need to pay a search cost to acquire more information. They show when buyers' beliefs are observable to sellers, sellers provide monopoly level of information. When buyers' beliefs are observable, full information is provided as the search cost vanishes. In our model information itself is the product the seller sells. Another recent paper Boleslavsky et al. (2018) studies an oligopoly market where each seller chooses both the price of the product and how much information he will disclose to buyers. Then find that when more firms competition in the market, firms reveal more information of their products to buyers.

Compare to these papers, we provide a case where competition leads to a less informative outcome. In those cases of persuasion and information revelation, the payoff of information senders are either through their preference over agents' actions or the sale of a product. Our model the only payoff to information sellers are the monetary transfers. We also model that how information senders (sellers) get the informational advantage by making investments.

We use the cost of generating information.Gentzkow and Kamenica (2014) address the cost of information in a persuasion problem. They consider the cost of signals as a function of the reduction in the receiver's entropy. In general, lower cost leads to a Blackwell more informative experiment. Frankel and Kamenica (2018) provides a framework to link the measure of information in an experiment with the measure of uncertainty in agents' beliefs after receiving signals from the experiment. They propose a class of measures, for example, entropy or variance function of the posterior beliefs. Our analysis allows a general class of cost functions. The cost functions considered in both papers, entropy and variance function, are compatible with our model.

The production technology of information has a feature of large fixed cost and small marginal cost. In a seminar paper, Spence (1976) analyzes the impact of fixed cost on the selection of products. Ronnen (1991) considers the influence of a minimum quality standard on the product qualities. He shows that imposing such a minimum standard will increase the overall quality level. In our model the seller choose quality of a multidimensional product and provide it to essentially two subgroups of buyers, who have different preferences over different dimensions. We show that competition may not lead to an improvement in the quality in the markets for information.

In our model, the competition among information sellers is through investment level. It is essentially a directed search model. Like in Kim and Kircher (2015), sellers make no commitment when buyers choose which seller to visit. They shows that when sellers use cheap talk messages to convey their private reservation values, efficiency can be achieved if sellers run a first price auction after buyers coming. Mirkin and Pycia (2016) explores the capacity of cheap talk communication in matching settings.

1.2 The Model

Buyers' Decision Problem

There are data buyers of measure 1 facing a decision problem under uncertainty. The uncertain state is drawn from set $\Omega = \{\omega_1, \omega_2\}$. A data buyer needs to choose an action a from action set $A = \{a_1, a_2\}$. Without loss of generality, we can normalize the payoff of action a_2 under both states into zero. Therefore the following matrix can capture the payoffs of buyers' decision problem.

$$\begin{array}{cc} & \begin{array}{cc} a_1 & a_2 \end{array} \\ \begin{array}{c} \omega_1 \\ \omega_2 \end{array} & \begin{pmatrix} u_1 & 0 \\ u_2 & 0 \end{pmatrix} \end{array}$$

We can think the two actions as "risky" and "safe" actions. The safe action results in zero payoffs under both states. The risky action yields a payoff $u_1 > 0$ under "good" state ω_1, and $u_2 < 0$ under "bad" state ω_2. The buyer will take action a_1 if $\theta u_1 + (1-\theta) u_2 \geq 0$, which implies $\theta \geq \frac{-u_2}{u_1 - u_2}$. Note θ^* as the belief type that is indifferent between two actions.

$$\theta^* u_1 + (1-\theta^*) u_2 = 0 \;\Leftrightarrow\; \theta^* = \frac{-u_2}{u_1 - u_2} \tag{1.1}$$

To simplify the notation, normalize the difference between risky action's payoffs in two states into 1. Denote $u_1 = 1 - r$ and $u_2 = -r$, where $r > 0$. Then $r = \theta^*$ is the cutoff belief. The buyer's payoff matrix becomes:

$$\begin{array}{cc} & \begin{array}{cc} a_1 & a_2 \end{array} \\ \begin{array}{c} \omega_1 \\ \omega_2 \end{array} & \begin{pmatrix} (1-r) & 0 \\ -r & 0 \end{pmatrix} \end{array}$$

Buyers arrive with a common prior about the state. We assume that each buyer has access to some private information unknown to other buyers and the sellers. After receiving private information, buyers update their prior belief to a interim belief $\theta \in [0,1]$. Assume the interim belief θ satisfies a distribution function $G(\theta)$ and associated density function $g(\theta)$, which are common knowledge. We call belief θ as the type of buyer. The expected utility of the buyer with type θ is,

$$u(\theta) = \max\{\theta - r, 0\}. \tag{1.2}$$

Information Products

Buyers can buy an information product (or an experiment) E from data sellers, to improve his expected utility from the decision problem. An information product is a likelihood function $\pi : \Omega \to \Delta S$, where S is the signal space. We can restrict our attention to the signal space consists of only two signals $S = \{s_1, s_2\}$, without loss of generality[2]. Denote π_i as the conditional probability of signal s_i in state ω_i,

$$\pi_i \triangleq \Pr[s_i \mid \omega_i], \ i = 1, 2.$$

Write an experiment as $E = (\pi_1, \pi_2)$ with the following probabilistic matrix.

E	s_1	s_2
ω_1	π_1	$(1-\pi_1)$
ω_2	$(1-\pi_2)$	π_2

[2]For any signal sets consisting of more signals than the number of actions, we can construct an experiment with two signals achieve the same outcome distribution as the initial experiment. This is an implication of the Prop. 1 in Bergemann et al. (2018).

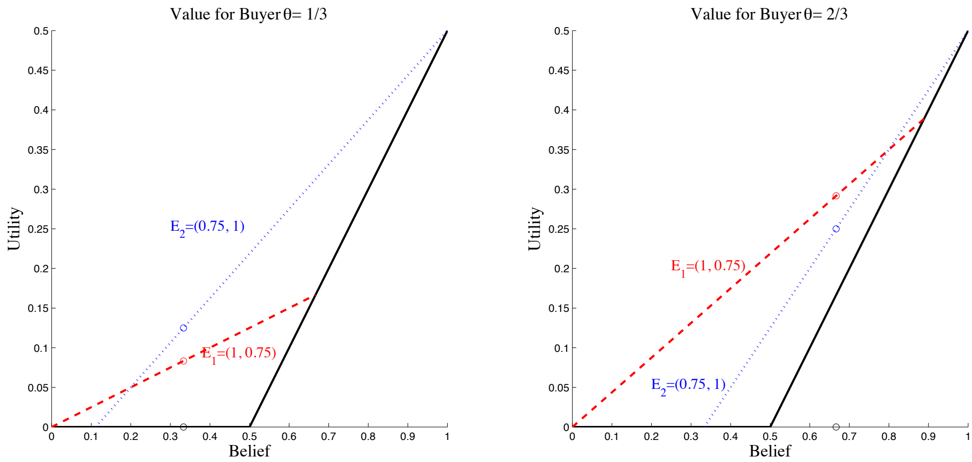


Figure 1.1: Value of Information for Different Buyers

Without loss of generality, we can assume $\pi_1 + \pi_2 \geq 1$. It means that signal s_1 is more likely to happen under state ω_1 than under state ω_2, $\pi_1 \geq (1-\pi_2)$. Some observations are noteworthy introduced here. (i) The experiment $\bar{E} = (1,1)$ is the fully informative experiment. (ii) When the sum of π_1 and π_2 equals to 1, we have $\pi_1 = 1 - \pi_2$. Thus each signal is sent with equal probabilities under any state hence is uninformative. Any experiment $E = (\pi, 1-\pi)$ is completely uninformative for any π. (iii) Fixing π_1, a higher π_2 implies that the probability of sending s_1 in state ω_2 is smaller. Therefore the signal s_1, or the good news, is more accurate. Similarly, a higher π_1 implies more accurate bad news in the experiment.

Value of Information

Consider a buyer with type θ buys the above experiment E. The buyer will update his belief after receiving each signal. Denote $\tilde{\theta}_j$ as the posterior belief after receiving signal s_i, $i \in \{1,2\}$. We define the value of information product E for buyer θ as the improvement of *expected* utility after updating his beliefs.

$$V(E,\theta) = \mathbb{E}_s\left[u\left(\tilde{\theta}_i\right)\right] - u(\theta) \tag{1.3}$$

The expectation is taken over all signals. Since $\tilde{\theta_1} = \frac{\theta\pi_1}{P(s_1)}$ and $\tilde{\theta_2} = \frac{\theta(1-\pi_1)}{P(s_2)}$, we have

$$V_0(E,\theta) = \max\left\{(1-r)\,\theta\pi_1 - r\,(1-\theta)\,(1-\pi_2)\,,0\right\} + \max\left\{\theta\,(1-r)\,(1-\pi_1) - r\,(1-\theta)\,\pi_2, 0\right\} - u(\theta) \tag{1.4}$$

The value of information $V(E,\theta)$ is always nonnegative. Note that an uninformative experiment always gives zero value. Furthermore, for a buyer has an extreme belief, if the experiment fails to send an accurate engough contrary information, he will ignore the signal and take his initial action. Therefore he gets zero value from the experiment.

Figure (1.1) gives an illustration about how the value of an experiment is determined. Let $r = \frac{1}{2}$ and consider two experiments $E_1 = (1, 0.75)$ and $E_2 = (0.75, 1)$. Two buyers $\theta_A = \frac{1}{3}$ and $\theta_B = \frac{2}{3}$. Then we have $V(E_1, \frac{1}{3}) = \frac{1}{12}$ and $V(E_2, \frac{1}{3}) = \frac{1}{8}$. The pessimistic buyer θ_A prefers E_2 which sends more accurate good news. Since $V(E_1, \frac{2}{3}) = \frac{7}{24}$ and $V(E_2, \frac{2}{3}) = \frac{1}{4}$. The optimistic buyer θ_B prefers E_1 which sends more accurate bad news. Therefore buyers prefer those experiments providing contrary information which disproves their private information.

The Production of Information

To be able to provide information products with certain quality, the data seller needs to make an investment I. Investment determines the most informative experiment the seller can ever provide. Specifically, assume a cost function $C(\pi_1, \pi_2)$ that measures the cost incurring in collecting data, developing technologies, and building statistical models to provide an experiment $E = (\pi_1, \pi_2)$. The most informative experiment $E^* = (\pi_1^*, \pi_2^*)$ which the seller can provide to buyers should satisfy $C\left(\pi_1^*, \pi_2^*\right) \leq I$. Once the investment is made, the seller need to choose the most informative experiment E^*. Despite selling E^*, the seller can add some noise and produce an experiment less informative than E^*. This is natural given the nature of information products. Assume downgrading an experiment is free. Furthermore, we assume the uninformative experiments are also available. Then the feasible set of experiments of a seller choosing E^* is

$$Y(\pi_1^*, \pi_2^*) \triangleq \{(\pi_1, \pi_2) \mid \pi_1 + \pi_2 > 1, \text{ and } \pi_1 \le \pi_1^*, \pi_2 \le \pi_2^*, \text{ where } C(\pi_1^*, \pi_2^*) \le I\} \tag{1.5}$$

Figure (1.2) illustrates the feasible set under (π_1^*, π_2^*). In Bergemann et al. (2018), fully informative experiment is available, $(\pi_1^*, \pi_2^*) = (1, 1)$. Therefore the feasible set becomes the largest upper triangle. Here we endogenize the choice of (π_1^*, π_2^*) and analyze what experiments will be provided given any feasible set.

Note that the straight line $\pi_1 + \pi_2 = 1$ represents all uninformative experiments. However, an experiments may not be responsive. For example, consider $(\pi_1^*, 1 - \pi_1^*)$ for a buyer with belief $\theta > r$. The buyer always chooses action a_1 and does not follow the recommendation of signal s_2. For technical convenience, we replace $(\pi_1^*, 1 - \pi_1^*)$ with $(1, 0)$. The latter one is responsive since the buyer only receive signal s_1. Then for the following analysis we consider the feasible set $Y(\pi_1^*, \pi_2^*) \cup \{\pi_1 + \pi_2 = 1, \pi_i \in [0, 1]\}$.

Menu of Experiments

A buyer's value of an experiment depends on his private belief θ, which is unknown to the seller. To maximize her profit, the seller will screen buyers by providing a menu of products. A menu of experiments, $\mathcal{M} = \{(E, t_E)\}$, consists of a collection of experiments E and their associated prices t_E. The seller posts her menu to the buyers. Then buyers can choose to buy one experiment E from $\mathcal{M}$ and pay the price t_E, or walk away with no additional information.

1.3 Monopoly

We now analyze the model with a monopoly seller. The timing is as follows:

(i) The seller chooses an investment I and the most informative experiment E^*;

(ii) The seller posts a menu $\mathcal{M}$;

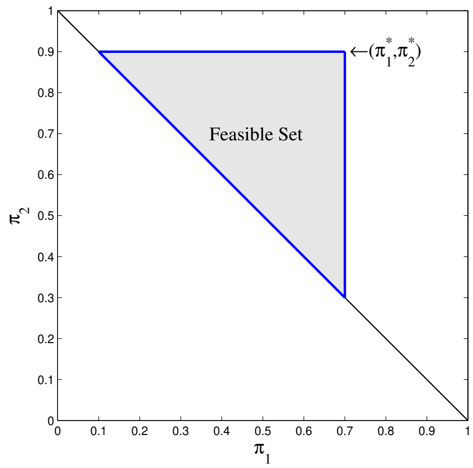


Figure 1.2: Feasible Set of Experiments

(iii) The state ω is realized while unobserved. Buyers receive private information and type θ is realized;

(iv) Every buyer chooses a product E from the menu and pays t_E ;

(v) Every buyer observes a signal in experiment E and chooses an action.

The seller's problem is first to choose an investment level I and the most informative experiment E^*, then to design a incentive compatible menu $\mathcal{M}$. To find the profit maximizing investment level, we need to characterize the optimal menu under all the possible E^*.

1.3.1 Revenue Maximized Menu

Fixing the most informative experiment $E^* = (\pi_1^*, \pi_2^*)$, in principle, the optimal menu $\mathcal{M}$ should include one experiment for each buyer. In this subsection we follow the section III.C of Bergemann et al. (2018), where $E^* = (1, 1)$. We apply their solution method to the general case of (π_1^*, π_2^*). We simplify the seller's problem in three steps. First, exclude those buyers who never buy an informative experiment. Second, restrict our attention on responsive experiments. Third, argue that the seller will not add noises to both two signals, therefore reduce the problem into one dimensional.

Since the informativeness of E^* is limited, a buyer with extreme belief might choose not to follow the recommendations for all feasible experiments. Therefore we can restrict our attention on buyers who have a positive value of the most informative experiment E^*. It is not difficult to get that buyers with belief $\theta \notin [\theta_1, \theta_2]$ gives zero value over all experiments in the feasible set, where $\theta_1 = \frac{(1-\pi_2^*)r}{\pi_1^*(1-r)+(1-\pi_2^*)r}$ and $\theta_2 = \frac{\pi_2^* r}{(1-\pi_1^*)(1-r)+\pi_2^* r}$. These buyers have too strong beliefs and have zero value over all experiments in the feasible set, therefore they can only get uninformative experiment in the market.

Since $\pi_1^* + \pi_2^* \geq 1$, we always have $\theta_1 \leq \theta^*$ and $\theta_2 \geq \theta^*$. If fully informative experiment is available, then $\theta_1 = 0$ and $\theta_2 = 1$. Now we can restrict our attention on those potential buyers with belief $\theta \in [\theta_1, \theta_2]$. Note the tructated distribution of potential buyers as

$$f(\theta) = \frac{g(\theta)}{G(\theta_2) - G(\theta_1)}, \ F(\theta) = \int_{\theta_1}^{\theta} f(x)dx, \ \forall \theta \in [\theta_1, \theta_2].$$

Since a buyer will only pay for an experiment if it has positive value. We can show that the value of an experiment is positive is equivalent with that the buyer will follow the recommendation of both signals. This fits the Blackwell insight that the information is valuable only if it changes decision maker's optimal actions.

We call an experiment E a *responsive experiment* for buyer θ if the buyer will follow the recommendations of both signals. Then we can eliminate the first max operator in the value function for responsive experiments.

$$V_1(E, \theta) = (1-r)\,\theta\pi_1 - r\,(1-\theta)\,(1-\pi_2) - \max\{\theta - r, 0\}. \tag{1.6}$$

Lemma 1.1 *The value of an information product E for buyer θ is positive, if and only if the buyer strictly prefer to following the recommendations, i.e., choosing a_i after receiving s_i, $i = 1, 2$. The outcome of every menu can be attained by a responsive menu, in which $E(\theta)$ is a responsive experiment for type θ, for all $\theta \in [\theta_1, \theta_2]$.*

Lemma 1.1 implies that an experiment is responsive for a buyer if and only if the value of information after following both recommendations, $V_1(E, \theta)$, is nonnegative. Now we can focus on responsive menus, hence the value of information $V_0(E, \theta)$ can be replaced by with

$V_1(E, \theta)$.

The seller's problem is on a two dimensional set. we now give the following characterization of optimal menu which eliminates buyer's uncertainty as much as possible at least along one dimension.

Lemma 1.2 *If an experiment $E = (\pi_1, \pi_2)$ is in the optimal menu $\mathcal{M}$, then either one of the following two conditions holds:*

(i) E is uninformative.

(ii) $\pi_1 = \pi_1^*$ or $\pi_2 = \pi_2^*$.

Therefore any experiment in the optimal menu either contains no information, or achieves maximal accuracy on one signal. Now we can replace the experiment (π_1, π_2) with a one dimensional measure. Rewrite equation (1.6) as

$$V(E, \theta) = \max\left\{\theta\left[\pi_1(1-r) + (1-\pi_2)r\right] - (1-\pi_2)r - \max\left\{\theta - r, 0\right\}, 0\right\} \qquad (1.7)$$

For type θ, define the following one dimensional value

$$q(\theta) \triangleq \pi_1(\theta)(1-r) + (1-\pi_2(\theta))r \ \in \ [0,1].$$

Notice that the endpoints of the interval correspond to two uninformative experiments $(0,1)$ and $(1,0)$. The most informative experiment E^* gives $q_0 = \pi_1^*(1-r) + (1-\pi_2^*)r = \pi_1^* + r(1-\pi_1^* - \pi_2^*)$. Lemma 1.2 implies that if $q \neq 0$ and $q \neq 1$, then we have either $\pi_1 = \pi_1^*$ or $\pi_2 = \pi_2^*$. Therefore $0 < q < q_0$ implies $\pi_2 = \pi_2^*$, and $q_0 < q < 1$ implies $\pi_1 = \pi_1^*$. Now we can rewrite the term $-(1-\pi_2)r$ in equation (1.7) as

$$-(1-\pi_2)r = 1_{\{q\neq 0\}} \cdot \left[\min\left\{q_0 - q, 0\right\} - (1-\pi_2^*)r\right].$$

Therefore the value function can be reduced to a function of q:

$$V(q, \theta) = \max\Big\{\theta q + 1_{\{q\neq 0\}} \cdot \left[\min\left\{q_0 - q, 0\right\} - (1-\pi_2^*)r\right] - \max\left\{\theta - r, 0\right\}, 0\Big\}. \qquad (1.8)$$

From equation (1.8) we can see the main features of a seller's screening problem. First, the most informative experiment $q = q_0$ is the most valuable for all buyers θ. Second, the indifferent buyer type $\theta = r$ has highest value for any experiment q. Third, $V(q, \theta)$ has the single crossing property in (q, θ), which essentially implies that the incentive compatible menu has an increasing $q(\theta)$.

From now on we focus on responsive menus. Lemma 1.1 implies the value of following both signals, $V_1(q(\theta), \theta)$, should be nonnegative. Then with equation (1.8) we can derive the necessary and sufficient conditions of a responsive menu.

$$q(\theta) \in Q(\theta) = \begin{cases} \left[\frac{(1-\pi_2^*)r}{\theta},\ \frac{\pi_1^*(1-r)}{1-\theta}\right] \cup \{0\} & \text{if } \theta \in [\theta_1, r] \\ \left[1 - \frac{\pi_2^* r}{\theta},\ 1 - \frac{(1-\pi_1^*)(1-r)}{1-\theta}\right] \cup \{1\} & \text{if } \theta \in [r, \theta_2] \end{cases} \tag{1.9}$$

Now the seller's problem is choosing $q(\theta)$ for $\theta \in [\theta_1, \theta_2]$ to maximize her revenue such that all buyers are incentive compatible. Then we can apply the method of Bergemann et al. (2018), where $E^* = (1, 1)$. We generalize their solution and derive the following characterizations of the optimal menu. The formal proof is left in the appendix.

Proposition 1.1 *Given $E^* = (\pi_1^*, \pi_2^*)$, the optimal menu includes at most two informative experiments. Let $\phi_1 = \theta f(\theta) + F(\theta)$ and $\phi_2 = (\theta - 1) f(\theta) + F(\theta)$ be the virtual values when $q \leq q_0$ and $q > q_0$ respectively. Let $\bar{\phi}_1, \bar{\phi}_2$ be the ironed virtual values. The necessary and sufficient conditions of the optimal menu $\mathcal{M} = \left\{(q^*(\theta), t^*(\theta))_{\theta \in [0,1]}\right\}$ are:*

(i) there exists $\lambda^* > 0$ such that $q^*(\theta)$ maximizes function $H(q, \theta)$ for all θ, where

$$H(q, \theta) = \begin{cases} \bar{\phi}_1(\theta) q(\theta) - \lambda^* q(\theta) & \text{if } q \leq q_0 \\ \bar{\phi}_2(\theta) q(\theta) + q_0 f(\theta) - \lambda^* q(\theta) & \text{if } q > q_0 \end{cases}; \tag{1.10}$$

(ii) $q^*(\theta)$ is non-decreasing;

(iii) $\int_{\theta_1}^{\theta_2} q^*(\theta)\, d\theta = \theta_2 - r$;

(iv) $q^*(\theta) \in Q(\theta)$;

(v) if the ironed virtual value is constant on an interval, $q^*(\theta)$ implies the same experiment on the interval.

The payment of type θ is

$$t^*(\theta) = \theta q^*(\theta) + 1_{\{q^* \neq 0\}} \cdot [\min\{q_0 - q^*, 0\} - (1 - \pi_2^*) r] - \int_{\theta_1}^{\theta} q^*(s)\, ds. \tag{1.11}$$

Here we discuss some intuitions of the optimal menu. The cardinality of optimal menu is from the linear property of value function. The single crossing properties of value function in q and θ implies that q should increase in θ. For any experiment, the value is continuous and differentiable in the intervals $[\theta_1, r]$ and $[r, \theta_2]$ separately. Then we have condition (iii) by applying the Maximal Theorem. In Bergemann et al. (2018), condition (iv) is implied by condition (ii) and (iii), but it is not the case here.

In general, the optimal menu gives $q(\theta) = 0$ for buyers with extreme pessimistic belief and $q(\theta) = 1$ for those with relatively optimistic beliefs. Since buyers with an extreme prior belief have low values for any experiment, they are excluded from the market. For buyers with beliefs around r, $q(\theta) = q_0$ and they get the most informative experiment. Depending on the shape of virtual values, the optimal menu may include a second informative experiment. Technically, only if the two virtual values need to be ironed, the optimal menu include two experiments. Since two sides of buyers have different preferences, the second experiment will target to those buyers with larger size. For example, if the density function has two peaks, one around a pessimistic belief θ_A and the other around the cutoff belief r. The second experiment is (π_1, π_2^*) where $\pi_1 < \pi_1^*$. The seller downgrade the experiment E^* to attract those buyers around θ_A, who prefer more accurate good news, i.e., a higher π_2. In the following we restrict our attention to the case that optimal menu only includes one product. The following proposition summarizes the sufficient conditions of a single-item optimal menu.

Corollary 1.1 *(BBS, Corollary 1) The optimal menu contains a single item if any of the following conditions (single-item conditions) holds:*

(i) (one-sided distribution) Almost all types have congruent beliefs, $F(r) \in \{0, 1\}$;

(ii) (regularity) Both virtual values $\phi_1(\theta)$ and $\phi_2(\theta)$ are strictly increasing;

(iii) (symmetricity) The monopoly price for experiment E^* is equal on $[\theta_1, \theta_0]$ and $[\theta_0, \theta_2]$.

Condition (i) means that all buyers are optimistic or pessimistic, hence buyers have similar preferences over the experiment. Then the seller does not need to provide an extra experiment to satisfy a subgroup of buyers with different preferences. Condition (ii) is a regularity conditions on the distribution functions. Density functions with one peak satisfy this condition. Condition (iii) states that if the optimistic and pessimistic buyers with binding a participation constraint have the same willingness to pay, the seller will not provide an extra experiment to attract more buyers.

We say the distribution function satisfies *single-item conditions*, if it satisfies one of the conditions in Corollary 1.1. In this case we only need to determine the cutoff beliefs of buyers who are provided with experiment E^*. Notice condition (1.10) implies that if two ironed values have a nonempty intersection of their image, the two cutoff beliefs will make two ironed values equal, i.e., $\bar{\phi}_1(\theta_1^*) = \bar{\phi}_2(\theta_2^*)$. If no such cutoff beliefs can make the two ironed value equal, all buyers on $[\theta_1, \theta_2]$ will be served.

Corollary 1.2 *Suppose the distribution function satisfies the single-item conditions (ii) or (iii), if $\bar{\phi}_1([\theta_1, \theta_2]) \cap \bar{\phi}_2([\theta_1, \theta_2]) \neq \varnothing$, the seller provides E^* to buyers on $[\theta_1^*, \theta_2^*]$. The endpoints can be solved with*

$$\begin{cases} f(\theta_1^*)\theta_1^* + F(\theta_1^*) = f(\theta_2^*)(\theta_2^* - 1) + F(\theta_2^*) \\ q_0(\theta_2^* - \theta_1^*) = \theta_2^* - r \end{cases} \tag{1.12}$$

The price of E^* is given by $t^* = \theta_1 \pi_1^*(1-r) - (1-\theta_1^*)(1-\pi_2^*)r$. The seller's revenue is $R(\pi_1^*, \pi_2^*) = t^* \cdot (\theta_2^* - \theta_1^*)$.

1.3.2 Optimal Investment

Once we have the distribution of buyers belief $G(\cdot)$, we can solve the cutoff beliefs of buyers served in the market, $[\theta_1^*, \theta_2^*]$, and the revenue function $R(\pi_1^*, \pi_2^*)$ with equations (1.12)..

Together with the cost function, we can solve for the monopoly seller's optimal investment choice and the most informative experiment (π_1^*, π_2^*). In the following we give a closed solution when the types follow a uniform distribution and discuss the optimal investment.

When $G(\cdot)$ is uniform distribution, on both sides of the cutoff belief r, half of the potential buyers will be served by the monopoly.

Corollary 1.3 *If $G(\cdot)$ is a uniform distribution on $[0,1]$, then given any $E^* = (\pi_1^*, \pi_2^*)$, the seller provides E^* to buyers with belief $\theta \in \left[\frac{\theta_1+r}{2}, \frac{\theta_2+r}{2}\right]$.*

The price of experiment E^* is given by $t^* = \theta_1^* q_0 - (1-\pi_2^*)r = \frac{1}{2}r(1-r)(\pi_1^*+\pi_2^*-1)$. Note that the value of E^* for uniformed buyer $\theta = r$ is $V(E^*, r) = r(1-r)(\pi_1^* + \pi_2^* - 1)$, which is the highest valuation among all buyers. The price equal half of this highest valuation, which is a standard solution when the demand curve is linear. Then we can obtain the seller's total revenue

$$R^*(\pi_1^*, \pi_2^*) = t^* \cdot \frac{1}{2}(\theta_2 - \theta_1) = \frac{(t^*)^2}{q_0(1-q_0)} = \frac{[r(1-r)(\pi_1^* + \pi_2^* - 1)]^2}{4q_0(1-q_0)} \tag{1.13}$$

Where $q_0 = \pi_1^* + r(1 - \pi_1^* - \pi_2^*)$. Then for any cost function $C(\pi_1, \pi_2)$, we can find the optimal investment level and most informative experiment of the seller by maximizing $R^*(\pi_1^*, \pi_2^*) - C(\pi_1^*, \pi_2^*)$. Intuitively the more accurate the signals are, the higher cost it should require. One way is to assume $C(\pi_1, \pi_2)$, is a function of $(\pi_1 + \pi_2 - 1)$. In this cost cost function two signals are perfect substitutes in production. We can also assume it related with how far the signals deviate from an uninformative signal, $\left[\left(\pi_1^* - \frac{1}{2}\right)^2 + \left(\pi_2^* - \frac{1}{2}\right)^2\right] \cdot c$. In the following proposition we gives a closed form solution for linear cost function, while the properties holds for other meaningful cost functions.

Example 1.1 *Given a cost function $C(\pi_1, \pi_2) = c \cdot (\pi_1 + \pi_2 - 1)$ with $c \in (0, \frac{1}{16}]$, then the optimal investment of seller is: (i) choosing $E^* = (1,1)$ when $r \in [\frac{1}{2} - \sqrt{\frac{1}{4} - 4c}, \frac{1}{2} + \sqrt{\frac{1}{4} - 4c}]$; (ii) choosing no investment otherwise .*

Proposition 1.2 *Given a symmetric cost function such that seller does not provide complete information at $r = \frac{1}{2}$, then as r increases, π_2^* increases and π_1^* decreases. Then both decreases after some cutoff $\bar{r} > \frac{1}{2}$. If $\frac{\partial^2 C}{\partial \pi_i^2} > 0$, investment first increases then decreases.*

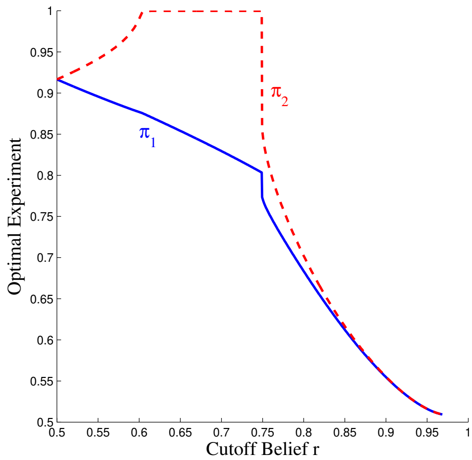


Figure 1.3: Optimal Experiment for Different Cutoff Belief r

Cost function is $C(\pi_1, \pi_2) = \frac{1}{5}\left[\left(\pi_1 - \frac{1}{2}\right)^3 + \left(\pi_2 - \frac{1}{2}\right)^3\right]$. The solid blue line is π_1^* and the dashed red line is π_2^*. The seller always will specializes in one signal when the two sides of buyers is not evenly distributed. When $r > \frac{1}{2}$, there are more pessimistic buyers who prefer higher π_2. Then optimal experiment is to invest more in π_2 until $\pi_2^* = 1$. As r increases from $\frac{1}{2}$, the investment level first increases since the benefit from specializing is the dominant factor. It then decreases because of the decrease in the size of potential buyers and the decrease in the value of information for all buyers.

As r increases, the investment level first increases and then decreases. Two economic forces at work as r increases. First, the value of E^* for the least informed buyer $\theta = r$ decreases since the payoff of good states decreases. Also note $V(E^*, r) = r(1-r)(\pi_1^* + \pi_2^* - 1)$, then the price of E^* decreases. Second, since the seller specializes in providing the experiment preferred by the pessimistic buyers, therefore among those served in the market, the increase in the size of pessimistic buyers is higher than the decrease in the size of optimistic buyers. Hence the sellers market share increases. When r is not very large, the benefit from specialization dominates hence investment increases. Seller benefits from specialization. Then the second effect, the decrease in the value of information for all buyers, becomes larger and the investment decreases. Figure (1.3) shows the optimal menu when we have a cubic cost function. The respective investment and profit are shown in Figure (1.5) of appendix.

1.4 Duopoly Competition

Now we analyze duopoly competition. A continuum types buyers with measure 1 buys information products to improve their decision making. Two sellers competing with the investment level and the highest quality experiments they choose. There are two stages in this model. In the first stage, each seller $k \in \{1, 2\}$ chooses an investment level I_k and the most informative experiment E_k. After observing two sellers' choices, buyers chooses which firm to go to. In the second stage, each seller choose a menu of experiments the provide and the prices, after buyers arriving. The buyers then choose to purchase an experiment from the menu, or exit the market with no additional information.

The timing is as follows:

(i) Each seller chooses an investment I_k and the most informative experiment E_k;

(ii) The state ω is realized while unobserved. Buyers receive private information and type θ is realized;

(iii) Every buyer chooses which seller to go to;

(iv) The seller posts a menu $\mathcal{M}_k$. Each buyer chooses a product E from the menu and pay t_E ;

(v) Every buyer observes a signal in experiment the purchased and chooses an action.

Note that we assume that each buyer can only visit one seller. Once a buyer selects a seller, he cannot buy product from the other seller. Therefore sellers will post the revenue maximizing menus and charge monopoly prices in the second stage. The competition between sellers lies in their choices of investment and the highest quality experiments, which essentially determine their market share and hence their profits in the second stage.

Formally, Each seller first choose investment I_k and her most informative experiment E_k. Then design a menu based on her belief about the distribution of buyers who choose to visit her. Suppose seller k holds a belief G_k about the measure of coming buyers, where

$G_k(\theta) \leq G(\theta)$. The seller's strategy is $\sigma_k = (I_k, E_k, \mathcal{M}_k(I_k, E_k))$. Each buyer first selects a seller and then chooses a product from the menu. Note the strategy of selecting a seller as $K(\theta) \in \Delta\{1,2\}$ and choosing an experiment $E(\theta, \mathcal{M}_{K(\theta)})$ from menu $\mathcal{M}_{K(\theta)}$. Then the strategy of buyer θ is $\sigma_\theta = (K(\theta), E(\theta, \mathcal{M}_{K(\theta)}))$.

Definition 1.1 *We focus on pure strategy* Perfect Bayesian Equilibrium $\Sigma^* = (\sigma_k^*, \sigma_\theta^*)_{k,\theta}$, *which satisfies:*

(a) Given a seller's belief $G_k^*(\theta)$, the investment I_k^* and her most informative experiment E_k^* maximize her profit, the menu she provides $\mathcal{M}_k^*$ is the revenue-maximizing menu under $G_k^*(\theta)$ and E_k^*;

(b) Given sellers' choices of menus $\mathcal{M}_1^*$ and $\mathcal{M}_2^*$, each buyer's choice of seller $K^*(\theta)$ and his choice of experiment $E^*(\theta, \mathcal{M}_{K(\theta)})$ maximize the utility of buyer θ;

(c) Each seller's belief about the types of buyers visiting to her is consistent with buyers' choices, $G_k^*(\theta) = \int_0^\theta 1_{\{K(\theta)=k\}} dG(x)$, $\forall \theta, k$.

1.4.1 Splitting Equilibrium

We focus on a class of equilibrium where buyers splitting their choices about sellers according to a common threshold. Call this type of equilibrium as *splitting equilibrium.*

When the distribution of types is symmetric, there is an equilibrium where two sellers split the market. Those pessimistic buyers all go to one seller, who specializes in generating more accurate good news, i.e. higher π_2. The optimistic buyers select the other seller who specializes in generating more accurate bad news. We can check this is indeed an equilibrium. (a) If all pessimistic buyers visit the same seller, say seller 1, she will specialize in providing experiment with a higher π_2. (b) The cutoff buyer $\theta = r$ is indifferent with two sellers, while for any pessimistic buyer $\theta < r$, he strictly prefers the experiment providing by seller 1. (c) There is market segmentation and both sellers have correct beliefs.

Proposition 1.3 *Suppose the cutoff type $r = \frac{1}{2}$, the density function $g(\cdot)$ and the cost function are symmetric. There exists a Perfect Bayesian Equilibrium where $I_1^* = I_2^*$, buyers* *with belief $\theta \leq \frac{1}{2}$ visit seller 1 and buyers with belief $\theta > \frac{1}{2}$ visit seller 2. Both sellers provide only one informative experiments in the menu. The experiments of two sellers, $E^{*1} = (\pi_1^{*1}, \pi_2^{*1})$ and $E^{*2} = (\pi_1^{*2}, \pi_2^{*2})$, are symmetric. And we have $\pi_1^{*1} \leq \pi_1^{*2}$.*

The formal proof is in Appendix. Note that in this equilibrium each seller only serves buyers with beliefs at one side of r, i.e., either $\theta \leq r$ or $\theta \geq r$. There the one-sided condition in Corollary (1.1) is satisfied. As long as we have function forms of $g(\cdot)$ and $C(\pi_1, \pi_2)$, we can solve the optimal experiment provided by duopoly by applying Corollary (1.2). When the cost function is linear, we have the following bang-bang solution, where sellers either provide the fully informative experiment or uninformed experiments.

Example 1.2 *Suppose the cutoff type $r = \frac{1}{2}$, $G(\cdot)$ is uniform distribution and cost function is $C(\pi_1, \pi_2) = c \cdot (\pi_1 + \pi_2 - 1)$, (i) if $c < \frac{1}{32}$, we have $E_1^* = E_2^* = (1, 1)$; (ii) if $c \geq \frac{1}{32}$, we have $E_1^* = (0, 1)$ and $E_2^* = (1, 0)$.*

Compare the above equilibrium in duopoly competition with the optimal investment of monopoly case in Example 1.1. We can see that when the marginal cost is in a moderate range $c \in (\frac{1}{32}, \frac{1}{16}]$, monopoly seller will make investment and provide a partial informative experiment. However, the duopoly sellers make no investment and provide no information. To see more insights we can work on a quadratic cost function $C(\pi_1, \pi_2) = c \cdot (\pi_1 + \pi_2 - 1)^2$ with $c \leq \frac{27}{256}$. When the marginal cost $c = \frac{1}{16}$. The monopoly seller provides experiment $(0.38, 1)$ (or $(1, 0.38)$), and the duopoly sellers provide $(0.25, 1)$ and $(1, 0.25)$ respectively.

Moreover, the splitting equilibrium exists with more general settings. Since two sides of the buyers have the opposite preferences over π_1 and π_2. Once two sellers choose to differentiate their experiments by specialize in generating different signals, each buyer will visit the seller specializing in his favorite signal. As long as the difference in profit from two sides of buyers are not too large, the splitting equilibrium exists.

Uniqueness. Splitting equilibrium may not be unique. Consider two sellers with different investments. Suppose the high investments specializing in π_2 and low belief buyers come to her. She can attract some high belief buyers with moderate belief, since they value of high investment seller's product is higher although they have opposite preference. Then the threshold in the splitting equilibrium is no longer the cutoff belief r. The new threshold θ^* should make the buyer with belief $\theta = \theta^*$ indifferent between buying from two sellers. We will give more detailed illustration in section 4.3.

Although the equilibrium may not be unique, we show that the above equilibrium maximizes two sellers' joint profit.

Proposition 1.4 *Suppose the cutoff type $r = \frac{1}{2}$, the density function $g(\cdot)$ and the cost function are symmetric. If two sellers split the market with a threshold belief θ, then the joint profit is maximized at $\theta = \frac{1}{2}$.*

1.4.2 Monopoly versus Duopoly

We are interested in the comparison between monopoly and duopoly. The following theorem is our main results that duopoly competition leads to a less informative outcome.

Theorem 1.1 *If the cutoff belief is $\frac{1}{2}$, the type of buyers follows a symmetric distribution, the cost function is symmetric, separable and the marginal cost is increasing, then the monopoly seller provides more informative experiments and serves more buyers than the case of duopoly.*

Proof. Given the most informative experiment (π_1, π_2), assume the monopoly seller's maximized revenue is $R^M(\pi_1, \pi_2)$ and the profit maximizing solution is (m, m). Assume The duopoly seller who serves low belief buyers has revenue function $R^D(\pi_1, \pi_2)$, note as seller 1. Seller 1'a profit maximizing solution is (d_1, d_2). Seller 1 specializes in generating higher π_2, therefore $d_1 < d_2$. So we only need to show that $d_2 < m$ when $m < 1$.

Now we consider constrained problem for seller 1, such that she has to choose $\pi_1 = \pi_2$. Then the pricing strategy of seller 1 is the same as monopoly seller for buyers with low belief $\theta < r$. Therefore we have $R^M(\pi, \pi) = 2R^D(\pi, \pi)$.

Note the separable and symmetric cost function as $C(\pi_1) + C(\pi_2)$. Then we know (m, m) maximized monopoly seller's profit $R^M(\pi_1, \pi_2) - C(\pi_1) - C(\pi_2)$. Therefore m is also the solution of the constraint maximization problem $\max_\pi \left[R^M(\pi,\pi) - 2C(\pi)\right]$. Thus we have

$$m \ \in \text{argmax}_\pi R^M(\pi,\pi) - 2C(\pi) = 2R^D(\pi,\pi) - 2C(\pi)$$

Since $m < 1$, we have that for all $\pi > m$, $\frac{dR^D(\pi,\pi)}{d\pi} \mid_\pi < C'(\pi)$. The monopoly seller does not choose a higher π because of the marginal revenue is lower than the marginal cost.

On the other hand, seller 1's optimal choice (d_1, d_2) maximized $R^D(\pi_1,\pi_2) - C(\pi_1) - C(\pi_2)$. Then we have

$$C'(d_2) = \frac{\partial R^D}{\partial \pi_2}(d_1,d_2) < \frac{\partial R^D}{\partial \pi_2}(d_2,d_2) < \frac{\partial R^D}{\partial \pi_2}(d_2,d_2) + \frac{\partial R^D}{\partial \pi_1}(d_2,d_2) = \frac{dR^D(\pi,\pi)}{d\pi} \mid_{\pi=d_2}$$

The first inequality is because $\frac{\partial R^D}{\partial \pi_2}(\pi_1,\pi_2)$ is increasing in π_1. The second inequality is due to the positive marginal revenue of π_1. The proof of these two properties is left in appendix. Then we have $m > d_2 > d_1$. The monopoly seller provide more informative experiment than the duopoly sellers. □

Therefore we provide a case where competition leads to a less informative outcome. The result holds for a broad class of distributions of types and cost functions. It can be interpreted in three parts. First, the nature of the generation and production of information products has the feature of almost zero marginal cost. Therefore from the viewpoint of maximizing social surplus, it is less efficient to have more than one unity to make investment. Second, duopoly sellers have less incentive to invest since they have less market share than the monopoly. Although duopoly sellers benefit from specialization when facing buyers with more aligned preference. We show the benefit is not large enough to out-weight the loss in market share, compared with monopoly. Third, the distortions in market becomes severe in the case of duopoly. Not only duopoly sellers provide less informative experiments, they also serves to less buyers.

1.4.3 Asymmetric Distributions

Net we check the splitting equilibrium when the distributions on the two sides of cutoff belief r becomes asymmetric. We carry out the analysis by changing r with a uniform distribution of buyers' types.

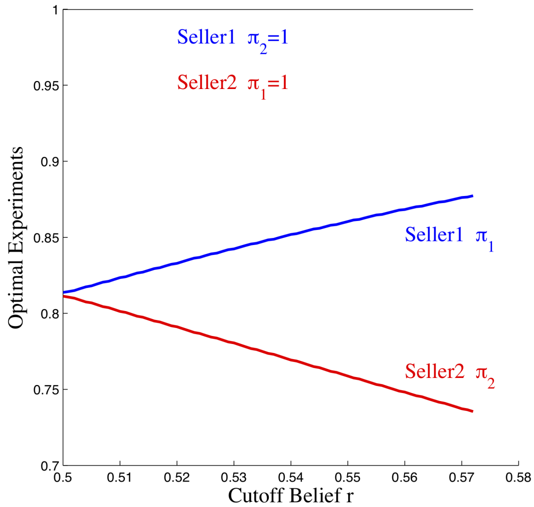
Suppose the cutoff belief r is larger than $\frac{1}{2}$. Then the seller serves the low belief buyers, say seller 1, has a higher benefit from specialization. Therefore she will increase the investment. Then the experiment provided by seller 1 is more attractive even for some buyers with belief $\theta > r$, while θ is near r. Then seller 1 can steal additional buyers from seller 2's market share. We show that there exists a new threshold $\theta^* > r$. Buyers splitting according to threshold θ^* is an equilibrium.

Proposition 1.5 *For any $r \in (0,1)$ and $r \neq \frac{1}{2}$, buyers' types follow a uniform distribution $G(\cdot)$, cost function is symmetric. If $r > \frac{1}{2}$, there exists a separating equilibrium with threshold $\theta^* > r$ such that buyers on the two sides of θ^* go to different sellers. The seller 1 serves buyers $\theta < \theta^*$ chooses $\pi_1^{*1} \leq \pi_2^{*1}$, and the seller serves buyers $\theta > \theta^*$ chooses $\pi_1^{*1} \leq \pi_2^{*1}$. Seller 1 has a larger market share.*

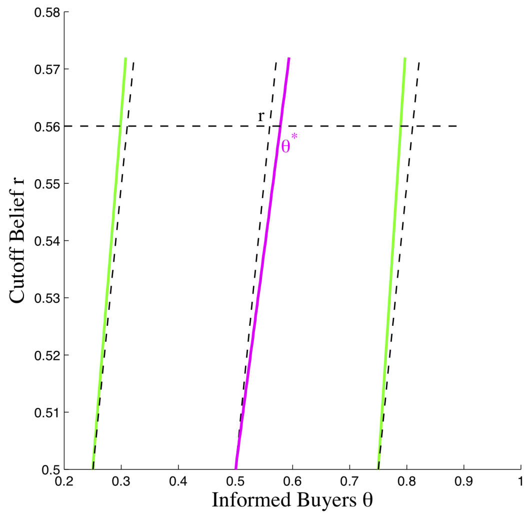
Figure 1.4 gives an illustration of the equilibrium of duopoly as r increases from $\frac{1}{2}$. Note that the optimal experiment of monopoly is $(1,1)$, which is more informative than duopoly sellers. Each duopoly seller specializes in generating one perfectly informative signal. While $r$increases, seller 1 increases her investment. For buyers with belief θ slightly higher than r, seller 1's experiment is more valuable. So the new threshold θ^* of splitting equilibrium lies on the right side of r.

1.5 Conclusion

In this paper we provide a framework to analyze the information provision with costly production. In the market for information, we find that sellers specializes in generating one signal, and buyers separating according to their private beliefs. In equilibrium there could be asymmericity in investments and profits, caused by skewed distribution of buyer's beliefs.



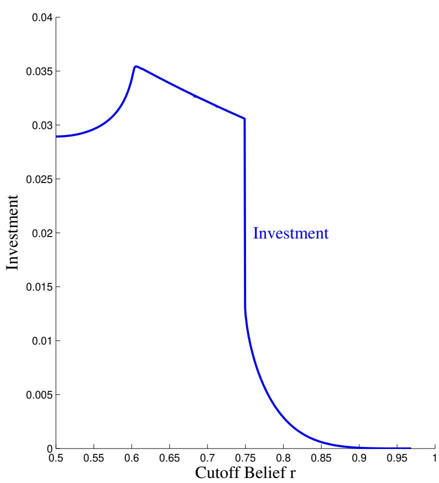
[Optimal Experiments]



[Informed Buyers]

Figure 1.4: Oligopoly Sellers for Different Cutoff Belief r

Cost function is $C(\pi_1, \pi_2) = \frac{1}{20}\left[\left(\pi_1 - \frac{1}{2}\right)^2 + \left(\pi_2 - \frac{1}{2}\right)^2\right]$. Panel (a) shows that how the optimal experiments chosen by duopoly sellers change as the cutoff belief r increases from $\frac{1}{2}$. Seller 1 serves buyers with low belief. Panel (b) shows how the measure of informed buyers changes. For any given r, seller 1 serves the buyers with belief between the left green line and the purple line. Seller 2 serves the buyers with belief between the purple line and the right green line. There middle dash line indicates $\theta = r$. The three dash lines are parallel. Seller 1 has a larger market share as r increases.



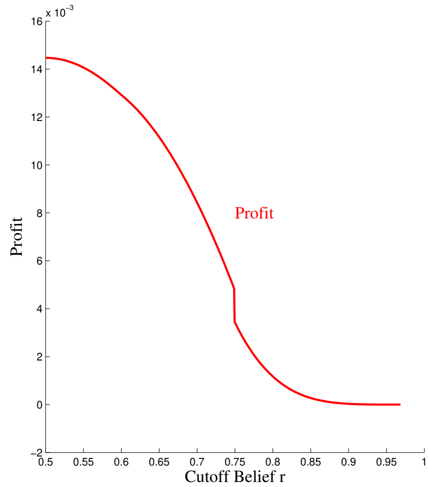


Figure 1.5: Monopoly Seller's Investment and Profit for Different Cutoff Belief r
Cost function is $C(\pi_1, \pi_2) = \frac{1}{5}\left[\left(\pi_1 - \frac{1}{2}\right)^3 + \left(\pi_2 - \frac{1}{2}\right)^3\right]$. Buyers' types follow a uniform distribution. Here shows how the total investment and profit of the monopoly seller change as the cutoff belief r increases from $\frac{1}{2}$. Investment first increases due to the benefit from specialization and then decreases as a reason of decrease in the value of information. However, the monopoly seller's profit decreases. It shows that the seller can extract the most consumer surplus when the distribution is symmetric on two sides of the cutoff belief.

The main finding is that when the generation of information is costly, competition leads to a less informed outcome.

Appendices

Proof of Lemma 1.1:

First we show that an experiment has positive value is equivalent with that it is responsive.

For any experiment E, note $U_1(\theta) = (1-r)\,\theta\pi_1 - r\,(1-\theta)\,(1-\pi_2)$ and $U_2(\theta) = (1-r)\,\theta\,(1-\pi_1) - r\,(1-\theta)\,\pi_2$, notice that $U_1 + U_2 = \theta\,(1-r) - (1-\theta)\,r$. Then we can rewrite equation (1.4) as

$$V(E,\theta) = \max\{U_1, 0\} + \max\{U_2, 0\} - \max\{U_1 + U_2, 0\}. \tag{1.14}$$

It's obvious that $V(E,\theta) \geq 0$. Now consider the case that $V(E,\theta) > 0$.

When $\theta \leq r$, then $U_1 + U_2 \leq 0$. Therefore $V(E,\theta) = \max\{U_1, 0\} + \max\{U_2, 0\} > 0$.

That means there must be one positive value in U_1 and U_2. Hence the other one must be negative. Since $U_2(\theta)$ is increasing in θ, then $U_2(\theta) \leq U_2(r) = r\,(1-r)\,(1-\pi_1-\pi_2) \leq 0$. Therefore $U_1 > 0$ and $U_2 < 0$.

When $\theta \geq \theta^*$, since $U_1 + U_2 \geq 0$, then $V(E,\theta) = \max\{U_1, 0\} + \max\{U_2, 0\} - U_1 + U_2$, which means there must be one negative value in U_1 and U_2. Hence the other one must be positive. Since $U_1(\theta)$ is increasing in θ, then $U_1(\theta) \geq U_1(r) = r\,(1-r)\,(\pi_1+\pi_2-1) \geq 0$. Therefore $U_1 > 0$ and $U_2 < 0$.

Therefore all buyers will follow the recommendation actions, i.e., chooses action a_1 under signal s_1 and chooses action a_2 under signal s_2.

The necessity is straightforward. If $U_1 > 0$ and $U_2 < 0$, we must have $V(E,\theta) > 0$ whenever $\theta \leq r$or $\theta \geq r$.

Next we show that we can focus on the responsive menu.

Without loss of generality, let type θ chooses the same action a_1 after each signal in $E(\theta) = (\pi_1(\theta), \pi_2(\theta))$. Replace $E(\theta)$ with $E_1 = (1, 0)$. Basically we increase the probability of sending signal s_1 in both states to one. Therefore E_1 gives the same outcome distribution as $E(\theta)$ for θ. Then $V(E_1,\theta) = V(E(\theta),\theta)$. Since E_1 is a garbling of $E(\theta)$. By Blackwell's theorem, $V(E_1,\theta') \leq V(E(\theta),\theta')$ for all θ'. Therefore the change will not affect all other types' IC constraints.

Proof of Lemma 1.2:

For any experiment $E = (\pi_1, \pi_2)$ in the optimal menu $\mathcal{M}$, if $\pi_1 < \pi_1^*$ or $\pi_2 < \pi_2^*$. Let $\delta = \min\{u_1\,(\pi_1^* - \pi_1), -u_2\,(\pi_2^* - \pi_2)\}$, and $\epsilon_1 = \frac{\delta}{u_1}$, $\epsilon_2 = \frac{\delta}{-u_2}$. Construct a new experiment E' where $\pi_1^{'} = \pi_1 + \epsilon_1$ and $\pi_2^{'} = \pi_2 + \epsilon_2$. Then $\pi_1^{'} \leq \pi_1^*$ and $\pi_2^{'} \leq \pi_2^*$, and there must be one hold with equality by construction. Furthermore $V(E^{'},\theta) - V(E,\theta) = \theta(\pi_1^{'} - \pi_1)u_1 + (1-\theta)(\pi_2^{'} - \pi_2)(-u_2) = \delta$ is independent of buyer's type. So the seller can replace (E, t_E) with $(E^{'}, t_E + \delta)$ without change all buyers' constraints and increase her revenue.

Proof for Proposition 1.1:

We generalize the approach used in the proof of Proposition 4 and 5 in Bergemann et al. (2018). Basically one can think BBS as a special case when $(\pi_1^*, \pi_2^*) = (1,1)$ and here we generalize it for any (π_1^*, π_2^*). This proof works for non-congruent beliefs, where $F(r) \notin \{0,1\}$. For congruent belief the solution is easier and we have it as part of the proof for Theorem 1.1.

Step 1. Claim that a menu is implementable and responsive if and only if: (i) $q(\theta) \in Q(\theta)$ is increasing and (ii) $\int_{\theta_1}^{\theta_2} q(\theta) = \theta_2 - r$.

Necessity.

Consider a responsive menu $\mathcal{M} = \{q(\theta), t(\theta)\}_\theta$. For any experiment $q(\theta)$, only those obedient type θ' of $q(\theta)$ have an incentive to mimic type θ. Therefore we can focus on those experiment q and types θ' such that $V(q,\theta') > 0$. For these cases we the write the value of information in equation (1.8) as

$$V(q,\theta) = \theta q + 1_{\{q \neq 0\}} \cdot [\min\{q_0 - q, 0\} - (1-\pi_2^*) r] - \max\{\theta - r, 0\} \qquad (1.15)$$

Consider any two types θ^1, θ^2 with $\theta^1 < \theta^2$ and any responsive experiments q^1 for type θ^1 and q^2 for type θ^2. Since the menu is implementable, the IC constraints implies $V(q^2,\theta^2) - t^2 \geq V(q^1,\theta^2) - t^1$ and $V(q^1,\theta^1) - t^1 \geq V(q^2,\theta^1) - t^2$. Therefore we have

$$V(q^2,\theta^2) - V(q^1,\theta^2) \geq t^2 - t^1 \geq V(q^2,\theta^1) - V(q^1,\theta^1)$$

The strict single crossing property of $V(q,\theta)$ implies that $q^2 \geq q^1$. Thus $q(\theta)$ in increasing. The value of experiment $V(q,\theta)$ in differentiable in θ on $[\theta_1, r]$ and $[r, \theta_2]$ respectively. For any q, the value of experiment is also continuous in θ. Therefore the rent function $V(\theta)$ of the optimal menu is also continuous by the Maximum Theorem. Then we have

$$V(r) = V(\theta_1) + \int_{\theta_1}^{r} V_\theta(q,\theta)d\theta = V(\theta_2) - \int_{r}^{\theta_2} V_\theta(q,\theta)d\theta.$$

Because the two types θ_1 and θ_2 have zero value on all experiments, we have $V(\theta_1) = V(\theta_2) = 0$. Therefore $\int_{\theta_1}^{r} V_\theta(q,\theta)d\theta + \int_{r}^{\theta_2} V_\theta(q,\theta)d\theta = 0$. Apply envelope theorem to equation (1.15) and we have $\int_{\theta_1}^{r} qd\theta + \int_{r}^{\theta_2}(q-1)d\theta = 0$. Thus we obtain:

$$\int_{\theta_1}^{\theta_2} qd\theta = \theta_2 - r.$$

Sufficiency.

If $q(\theta)$ satisfies the two conditions. Construct the following prices:

$$t(\theta) = \begin{cases} \theta q(\theta) + 1_{\{q\neq 0\}} \cdot [\min\{q_0 - q(\theta), 0\} - (1-\pi_2^*)\, r] - \int_{\theta_1}^{\theta} q(x)dx & \text{if } \theta \leq r \\ \theta q(\theta) + 1_{\{q\neq 0\}} \cdot [\min\{q_0 - q(\theta), 0\} - (1-\pi_2^*)\, r] + \int_{\theta}^{\theta_2} (q(x)-1)\, dx - (\theta - r) & \text{if } \theta > r \end{cases} \tag{1.16}$$

With $\int_{\theta_1}^{\theta_2} q(\theta) = \theta_2 - r$ we can simplify the price function as

$$t(\theta) = \theta q(\theta) + 1_{\{q\neq 0\}} \cdot [\min\{q_0 - q(\theta), 0\} - (1-\pi_2^*)\, r] - \int_{\theta_1}^{\theta} q(x)dx \quad \forall \theta \in [\theta_1, \theta_2]\,. \tag{1.17}$$

For any type θ, the net value from reporting θ' is

$$V(q(\theta'),\theta) - t(\theta') = q(\theta')(\theta - \theta') + \int_{\theta_1}^{\theta'} q(x)dx - \max\{\theta - r, 0\}$$

Since $q(\cdot)$ is increasing on $[\theta_1, \theta_2]$, the net value reaches its maximal at $\theta' = \theta$. Then the menu including $\{q(\theta), t(\theta)\}_\theta$ is implementable and incentive compatible.

Furthermore, Figure 1.6 gives an illustration about the constraints on $q(\theta)$ for responsive experiments.

Step 2. At most two informative experiment in the optimal menu.

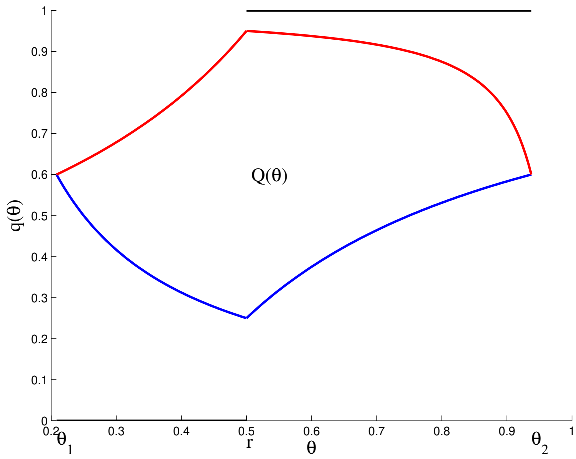


Figure 1.6: Constraint Set $Q(\theta)$ on Responsive Experiments
($r = 0.5$, $\pi_1^* = 0.95$ $\pi_2^* = 0.75$, which gives $\theta_1 = 0.24$, $\theta_2 = 0.93$, and $q_0 = 0.625$.)

First we write down the seller's optimization problem formally. With the price function (1.17) we can derive the seller's expected revenue

$$\int_{\theta_1}^{\theta_2} t(\theta) dF(\theta) = \int_{\theta_1}^{\theta_2} \left\{ \left[\theta f(\theta) + F(\theta)\right] q(\theta) + 1_{\{q \neq 0\}} \cdot \left[\min\left\{q_0 - q(\theta), 0\right\} - \left(1 - \pi_2^*\right) r\right] f(\theta) \right\} d\theta$$

The linearity of value of informations implies that the optimal menu is a step functions. To apply the Fundamental Theorem of Linear Programming to this maximization problem, we need first discretize the buyers type into finite grids $[\theta_1, \theta + \epsilon]$, $[\theta + 2\epsilon, \theta + 3\epsilon]$, ..., $[\theta_2 - \epsilon, \theta_2]$. Then given the fact that the optimal menu must include the most informative experiment q_0, we can construct a linear programing problem similar with the one in the Proposition 4 of BBS. Here we omit the details.

Step 3. Characterize the optimal menu.

The seller's optimization problem has an objective function concave in q. The constraints on $q(\cdot)$ is given by the two conditions in Step 1. Lagrangian method is valid. To deal with the integral constraint, we can use the method in Toikka (2011). Together with BBS, we can use the ironed virtual value to solve the Lagrangian problem.

Proof for Proposition 1.2:

We first give the monopoly seller's revenue function given a uniform distribution.

$$R^M(\pi_1,\pi_2,r) = \frac{r(1-r)(\pi_1+\pi_2-1)}{r(1-r)\left(\pi_1\pi_2+(1-\pi_1)(1-\pi_2)\right)+r^2\pi_2(1-\pi_2)+(1-r)^2\pi_1(1-\pi_1)}.$$

Suppose monopoly seller provide (π,π) when $r=\frac{1}{2}$. Starting from $\pi_1=\pi_2=\pi$, we examine the change of revenue if there is a small change of π_1 and π_2 when r is increasing from $\frac{1}{2}$. Since the cost function is symmetric, we can let $\pi_1=\pi-\epsilon$ and $\pi_2=\pi+\epsilon$ without changing the investment. The term $(\pi_1+\pi_2-1)$ and $(\pi_1\pi_2+(1-\pi_1)(1-\pi_2))$ do not have a first order change. Let $K=r^2\pi_2(1-\pi_2)+(1-r)^2\pi_1(1-\pi_1)$ and its change determines whether the revenue is increasing or decreasing. $\Delta K = r^2(1-2\pi)\epsilon-(1-r)^2(1-2\pi)\epsilon = \left[r^2-(1-r)^2\right](1-2\pi)\epsilon$. If $\epsilon>0$, we have $\Delta K<0$. Therefore the revenue R^M is decreasing. Since we fixed the cost, then the profit is increasing. Then the monopoly seller will increases π_2 and decreases π_1.

Proof for Proposition 1.3:

Suppose two sellers choose $I_1^*=I_2^*$ and buyers that splitting at threshold $r=\frac{1}{2}$, we show that it is buyer's best response to choose a seller according to the splitting technology. Notice that the value of information for low belief buyers $(\theta<r)$ and high belief buyers $(\theta>r)$ are symmetric in π_1 and π_2. When cost function is symmetric, suppose seller 1 choose the most informative experiment as (π,π'), we know that $\pi<\pi'$. Since the distribution of buyers faced by seller 1 is congruent, $F_1(r)=1$, then the seller 1 only provide one item in its menu. Suppose the price is t_1 and buyer $\underline{\theta}_1$ is the lowest informed buyer from seller 1. Then $t_1=V(\underline{\theta}_1,\pi,\pi')$. Due to symmetricity, seller 2 will provide a single item menu with price $t_2=t_1$. Therefore for the cutoff type buyer $\theta=r$, it is indifferent from buying seller 1 and seller 2. For all other buyers, stick to current choices is strictly better than visiting the other seller. Intuitively, because of the difference preferences on two sides of the cutoff belief, the seller can specialize their investment in the way to better meet their customers' needs. Then the buyers will splitting as a result. When investments are not observable, deviation is not optimal for each seller. Since $I_1^* = I_2^*$ is the optimal investment when both sellers acting as a monopoly in the second stage, any deviation is not profitable.

Proof for Theorem 1.1:

Consider the duopoly seller 1 who serves buyers with belief $\theta < r = \frac{1}{2}$. Given any most informative experiment (π_1, π_2), since the value function is strictly increasing in θ for these buyers, the seller 1 only needs to choose the buyers $\underline{\theta}$ which is the lowest type she serves. The price of the experiment is then $V(\underline{\theta}, \pi_1, \pi_2)$. Then seller 1's revenue function is

$$\begin{aligned} R^D(\pi_1, \pi_2) &= \max_{\theta} V(\theta) \cdot \left[\frac{1}{2} - G(\theta)\right] \\ &= \max_{\theta} \frac{1}{2}\left[\theta\pi_1 - (1-\theta)(1-\pi_2)\right] \cdot \left[\frac{1}{2} - G(\theta)\right] \end{aligned} \tag{1.18}$$

Apply the envelope theorem for the problem we have that $\frac{\partial R^D}{\partial \pi_2} = \frac{1}{2}(1-\theta^*) \cdot \left[\frac{1}{2} - G(\theta^*)\right]$ and $\frac{\partial R^D}{\partial \pi_2} = \frac{1}{2}\theta^* \cdot \left[\frac{1}{2} - G(\theta^*)\right]$, where θ^* is the optimal solution. Then both $\frac{\partial R^D}{\partial \pi_2}$ and $\frac{\partial R^D}{\partial \pi_1}$ is positive, the duopoly seller's revenue is increasing in both π_1 and π_2. To show $\frac{\partial R^D}{\partial \pi_2}$ is strictly increasing in π_1, it is sufficient if θ^* is decreasing in π_1, $\frac{\partial \theta^*}{\partial \pi_1} < 0$.

The first order condition of the maximization problem is

$$(\pi_1 - \pi_2 + 1)\left[\frac{1}{2} - G(\theta^*) - g(\theta^*)\theta^*\right] = -(1-\pi_2)\, g(\theta^*) \tag{1.19}$$

Take derivative w.r.t. π_1, we have

$$\frac{\partial \theta^*}{\partial \pi_1} \cdot \left\{-g'(\theta^*)\left[\theta^*\pi_1 - (1-\theta^*)(1-\pi_2)\right] + 2g(\theta^*)\left(\pi_1 - \pi_2 + 1\right)\right\} = \frac{1}{2} - G(\theta^*) - g(\theta^*)\theta^* \tag{1.20}$$

From the FOC we know that $\left[\frac{1}{2} - G(\theta^*) - g(\theta^*)\theta^*\right] < 0$. Let the term on the left hand side be $LHS = -g' \cdot [\theta^*\pi_1 - (1-\theta^*)(1-\pi_2)] + 2g \cdot (\pi_1 - \pi_2 + 1)$, we only need to show $LHS > 0$.

Case 1, $g'(\theta^*) \geq 0$. Since $V(\theta^*) > 0$, then $\theta^* \pi_1 > (1-\theta^*)(1-\pi_2)$. We can rewrite LHS as

$$\begin{aligned}
LHS &= \pi_1 (2g - \theta^* \cdot g') + (1-\pi_2)(g' + \theta^* \cdot g' - 2g) \\
&> (1-\pi_2)\left[\frac{1-\theta^*}{\theta^*}(2g - \theta^* \cdot g') + (g' + \theta^* \cdot g' - 2g)\right] \\
&= 2(1-\pi_2)\left[\left(\frac{1}{\theta^*} - 2\right) g + \theta^* \cdot g'\right] > 0
\end{aligned}$$

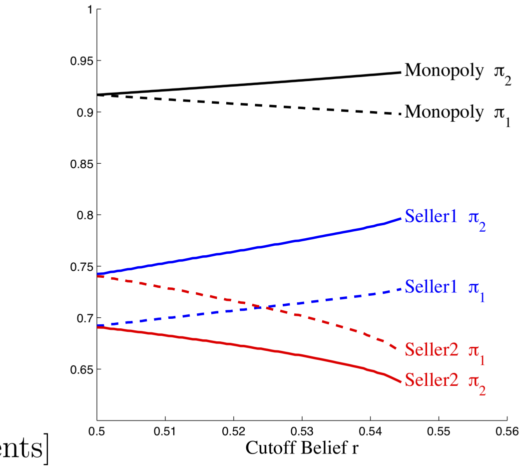
Case 2, $g'(\theta^*) < 0$. FOC gives that $(1-\pi_2) = \frac{1}{g}(G + \theta g - \frac{1}{2})(\pi_1 - \pi_2 + 1)$. We can rewrite LHS as

$$\begin{aligned}
LHS &= (\pi_1 - \pi_2 + 1)(2g - \theta^* \cdot g') + (1-\pi_2) g' \\
&= (\pi_1 - \pi_2 + 1)\left[(2g - \theta^* \cdot g') + \frac{1}{g}(G + \theta g - \frac{1}{2}) \cdot g'\right] \\
&= \frac{1}{g}(\pi_1 - \pi_2 + 1)\left[2g^2 + \left(G - \frac{1}{2}\right) g'\right] > 0
\end{aligned}$$

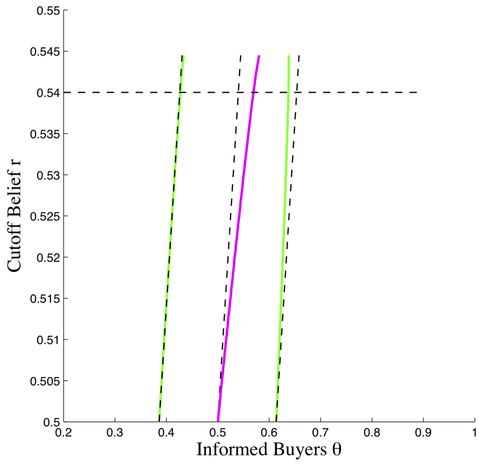
Therefore we finish the proof that $\frac{\partial R^D}{\partial \pi_2}$ is strictly increasing in π_1.

Proof for Proposition 1.5:

When buyers' types follow a uniform distribution. Sellers always provide single item menu in any splitting equilibrium. When the cutoff belief r increases from $\frac{1}{2}$, we can use the argument from the poof for Proposition **1.2** to claim that seller 1 will increase both π_1^{*1} and π_2^{*1}. First suppose buyers still splitting according to r, then seller 1's revenue from the market when providing (π_1, π_2) is $R^1 = V(\frac{r+\theta_1}{2}, \pi_1, \pi_2) \cdot (\frac{r-\theta_1}{2})$, where θ_1 is given by Corollary 1.3. The function form is similar with the monopoly's revenue, so apply the same argument we can get that seller 1 will increase π_1^{*1} and π_2^{*1}. The intuition is that seller 1 has a larger market share therefore the investment will increase. Due to the monopoly pricing, the increase in price is less than the increase in the value of the cutoff buyer. As a result, seller 1 will increase its price. On the other hand, by examining seller 2's revenue function we can get that seller 2 will decrease both π_1^{*2} and π_2^{*2} and decrease its price. Now the net value for cutoff buyer



[Optimal Experiments]



[Informed Buyers]

Figure 1.7: Duopoly Sellers for Different Cutoff Belief r

Cost function is $C(\pi_1, \pi_2) = \frac{1}{5}\left[\left(\pi_1 - \frac{1}{2}\right)^3 + \left(\pi_2 - \frac{1}{2}\right)^3\right]$. Buyers' types follow a uniform distribution. Panel (a) shows that how the optimal experiments chosen by monopoly and duopoly sellers change as the cutoff belief r increases from $\frac{1}{2}$. Seller 1 serves buyers with low belief. Note that seller 1 increases investment in both signals, and seller 2 decreases investment in both signals. The optimal experiment of duopoly sellers are always less informative than that of monopoly. Panel (b) shows how the informed buyers change as the cutoff belief r increases. For any given r, seller 1 serves the buyers with belief between the left green line and the purple line. Seller 2 serves the buyers with belief between the purple line and the right green line. All other buyers get uninformative experiment.

from seller 1 is strictly higher than that from seller 2 Consider the buyer with type $\theta = r + \epsilon$. Although he has a different preference over two signals from low belief buyers, the net value from seller 1's experiment is higher than the value from seller 2's experiment. However, due to the effect of specialization, the value function of high belief buyers for seller 1's experiment decreases faster than the value function fro seller 2's experiment. So by moving the threshold from r to $r + \epsilon$ we can decrease the difference of threshold buyer's net payoff from two sellers. Then there exist a $\theta^* > r$ such that he is indifferent again from two sellers.

CHAPTER 2

Preferential Attachment as an Information Cascade in Emerging Networks[1]

2.1 Introduction

Preferential attachment is one of the central concepts in network science[2], and is the underlying mechanism that generates most of the commonly observed social network properties, such as power laws and scale-free degree distributions. The preferential attachment mechanism, which is the networked version of the ubiquitous Pólya urn process [Barabási and Albert (1999)], generates random networks by sequentially adding nodes to the network, where each node is connected to the predecessor nodes with a probability that is proportional to their current degrees, giving rise to the rich-gets-richer phenomena in terms of the agents' connectivity degrees. Since its introduction to the network science community by Barabási and Albert (1999), the preferential attachment mechanism has become a key component in most current generative network models[3]. Since evidence for preferential attachment has been found in empirical data [Perc (2014); Zhou et al. (2007)], and since preferential attachment generates networks with properties that are observed in real-world data (such as scale-free degree distributions), almost all current generative network models use preferential attachment as an exogenously imposed probabilistic rule for link formation, without much

[1]This chapter is a joint work with Professor Mihaela van der Schaar, and Ahmed M. Alaa. (van der Schaar: mihaela@ee.ucla.edu, UCLA; Alla, ahmedmalaa@ucla.edu, UCLA.)

[2]See Barabási (2012); Papadopoulos et al. (2012); Perc (2014); Vázquez (2003); D'souza et al. (2007); Kakade et al. (2005); Freno et al. (2012); Gopalan et al. (2013).

[3]See Papadopoulos et al. (2012); Jackson and Rogers (2007); Vázquez (2003); Leskovec et al. (2008); D'souza et al. (2007); Kakade et al. (2005); Freno et al. (2012); Gopalan et al. (2013); Even-Dar and Kearns (2007); Gabel and Redner (2013); Zhou et al. (2007); Krapivsky and Krioukov (2008).

attention to the reasons behind its emergence[4].

In this paper, we aim at providing an informational interpretation of preferential attachment in networks of rational agents, such as citation networks [Golosovsky and Solomon (2012)], online social networks [Leskovec et al. (2008)], etc, where human agency and decision-making drive network evolution. We view preferential attachment as a herd behavior in which newly arriving nodes are influenced by the actions of their predecessors', i.e., network agents, confronted with lack of information, are more likely to form links with the most connected predecessor agent because they (rationally) believe that those agents are of a good quality. Our view of preferential attachment is inspired by the Bandwagon cognitive bias in perceptual decision-making [Bardone (2011)], and by the social learning models in microeconomics[5].

Summary of contributions. In Section 2.2, we develop a social learning model for network evolution, where agents of random qualities join the network in sequence, and each agent aims at linking with a high quality predecessor agent. Qualities are unknown, but can be inferred from two sources of information: private signals that an agent has about the predecessor agents (and are correlated with their true qualities), and public information represented as the current network structure, which implicitly encodes the actions of predecessor agents. We formulate the social learning problem as a network formation game, and characterize its Perfect Bayesian Equilibria in Section 2.3. We show that various forms of preferential attachment can emerge as sequentially rational equilibria of the network formation game. For instance, we show that when the agents' private beliefs are bounded, *condensed preferential attachment* emerges at equilibrium, in which all successor agents after some point of time will follow the herd and connect to the same predecessor agent. If the private beliefs are unbounded, then nonlinear preferential attachment emerges at equilibrium, where the probability of an agent getting a link found to be an exponential function

[4]See Vázquez (2003); Kakade et al. (2005); Freno et al. (2012); Gopalan et al. (2013); Even-Dar and Kearns (2007); Gabel and Redner (2013); Zhou et al. (2007); Krapivsky and Krioukov (2008).

[5]See seminal papers Bikhchandani et al. (1992); Banerjee (1992), and also later papers in Vázquez (2003); Smith and Sørensen (2000); Acemoglu et al. (2014, 2011).

of her in-degree.

Related works. Previous generative network models have imposed preferential attachment as an exogenous network formation mechanism, without attempting to understand its micro-foundations[6], whereas other works have focused on empirically validating the preferential attachment hypothesis in real-world networks [see Perc (2014)]. Heuristic interpretations for preferential attachment in terms of random walks on graphs can be found in Vázquez (2003). However, these interpretations ignore the decision-making and informational aspects of network formation, and the role of human agency, and view network growth as a pure random process. Other models interpret preferential attachment as a result of optimization problems that each agent solves in order to decide which link to form; however, these models assume that popularity (or the in-degree) are part of the optimization objective function of the agent, and hence preferential attachment straightforwardly emerges in these models without providing much insight on its micro-foundations.

Social learning has been recently studied by Acemoglu et al. (2011) and Acemoglu et al. (2014) in the context of social networks. Our work differs fundamentally from those models, in addition to classical social learning models by Bikhchandani et al. (1992), Banerjee (1992), and Smith and Sørensen (2000), in that agents in our model are learning about all predecessors' types, rather than learning one constant underlying state-of-the-world. Our focus is not on asymptotic learning as in Acemoglu et al. (2011) and Acemoglu et al. (2014), but rather on the implications of incomplete information and social learning on the network structures (i.e. emergence of preferential attachment). To the best of our knowledge, our model is the first decision-making network formation model for which preferential attachment emerges endogenously rather than being imposed as an exogenous probabilistic mechanism.

---

[6]See Papadopoulos et al. (2012); Jackson and Rogers (2007); Vázquez (2003); Leskovec et al. (2008); D'souza et al. (2007); Kakade et al. (2005); Freno et al. (2012); Gopalan et al. (2013); Even-Dar and Kearns (2007); Gabel and Redner (2013); Zhou et al. (2007); Krapivsky and Krioukov (2008).

2.2 Model

2.2.1 Network Evolution as a Sequential Decision Process

Agents. Consider a countably infinite set of agents $\mathcal{N} = \{0, 1, 2, \ldots\}$, making link formation decisions sequentially. Each agent $n \in \mathcal{N}$ is characterized by a *type* attribute $q_n \in Q$, where Q is the space of all possible types. We assume that agents are of two possible types, i.e. $Q = \{L, H\}$, where L denotes *low quality* agents, whereas H denotes *high quality* agents. We assume that the types of the agents are drawn randomly and independently from a *Bernoulli* distribution, with a prior $\mathbb{P}(q_n = H) = p$, and $\mathbb{P}(q_n = L) = 1 - p, \forall n \in \mathcal{N}$.

Agents make link formation decisions in sequence; the order of the agents is exogenous and is common knowledge. This is a common assumption in classical social learning models; in the context of social networks, this captures the sequentiality of arrivals and decision-making in a growing network such as a citation network or Twitter [see Leskovec et al. (2008); Golosovsky and Solomon (2012)]. The payoff of an agent n forming a link with agent j depends only on the type of agent j, i.e. agent n benefits from linking to agent j only if agent j is a high quality agent (e.g. citing a high quality paper).

As agents form links with other agents, a network is formed and agents in such a network would gain different levels of popularity. We model the network evolution as a discrete-time graph process $\{G_n\}_{n \in \mathbb{N}}$, where G_n is a *directed* graph. In every time period n, agent n is required to take an irreversible link formation decision, i.e. agent n must select *one* agent to link with[7]. We assume that agent 0, who arrives first, does not make a link. At time n, agent n can only link to one of the agents in the set $\{0, 1, 2, \ldots, n-1\}$. The qualities of the agents are not publicly known, thus agents are uncertain about the types of each other[8]. However, each agent n holds a set of *private signals* that indicate the qualities of all agents

[7]Assuming that agents form multiple links would lead to analytical intractability without adding much to the insights and results.

[8]Note that unlike classical social learning models, where there is a single binary *state-of-the-world* variable [as in Bikhchandani et al. (1992); Banerjee (1992); Smith and Sørensen (2000)], the state-of-the-world in our model at time n is the set of agent qualities $\{q_0, \ldots, q_{n-1}\}$, which is dynamic and grows more complex as the network grows.

in the set $\{0, 1, 2, \ldots, n-1\}$. Such signals convey information about the agents' qualities, yet they do not perfectly reveal the true qualities of the agents.

Private signal structure. Let $s_n = (s_n^0, \ldots, s_n^{n-1})$ be the set of private signals of agent n, where s_n^j represents agent n's signal on agent j's quality. The private signal s_n^j is drawn from a distribution $\mathbb{F}^H(s)$ if $q_j = H$, and is drawn from a distribution $\mathbb{F}^L(s)$ if $q_j = L$. Conditional on the true quality of all agents $\{q_1, q_2, \ldots\}$, all private signals are independently and identically distributed. We assume that the probability distribution of every private signal s_n^j satisfies the following regularity conditions:

- *(Common support)* $\mathbb{F}^H(s)$ and $\mathbb{F}^L(s)$ are mutually absolutely continuous with a common support supp($\mathbb{F}$) (That is, a private signal that any agent has for the quality of every agent j cannot perfectly reveal the true quality of that agent).

- *(Monotone likelihood)* The conditional pdfs $f^H(s) = f(s \mid q_j = H)$ and $f^L(s) = f(s \mid q_j = L)$ satisfy the *monotone likelihood ratio property* (MLRP) in the sense that $\frac{f^H(s)}{f^L(s)}$ is strictly increasing in s. It implies that a high quality agent is more likely to send large signals than a low quality agent.

Observations. Upon making a decision, agent n does not observe the private information of other agents, nor does she observe their true types. That is, agent n only observes her private signal s_n and the graph formed in the previous time step G_{n-1}. Note that the graph G_{n-1} encodes the actions of all the predecessors of n. Let the action of every agent j be denoted by a_j, where $a_j \in \{0, 1, 2, \ldots, j-1\}$. Then we can denote the graph as $G_{n-1} = (a_1, a_2, \ldots, a_{n-1})$ for simplicity. Thus, agent n observes the private signal s_n, and the sequence of actions of her predecessors, G_{n-1}.

In order to make a rational decision, an agent forms a belief about the types of the predecessors having observed her private signals (her own assessment of the qualities of the predecessors), and the actions of the predecessors (the public assessment, or the reputation of those predecessors). The actual payoff that an agent n realizes after taking her linking

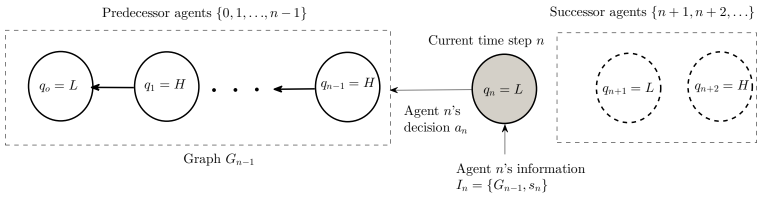


Figure 2.1: Depiction for the sequential decision-making process.

action is given by

$$u_n(a_n) = 1_{\{q_{a_n}=H\}}, \tag{2.1}$$

where 1 is the indicator function. This means that agent n gets a payoff of 1 if she links to a high quality agent, whereas she gets a 0 payoff otherwise. Agents gain *popularity* (in-degree) by having other agents linking to them. The *in-degree* of agent j at the beginning of time step n is given by $d_j(G_{n-1})$, and reflects the number of agents that formed links with agent j up to agent $n-1$. Figure 1 depicts pictorially the network growth process and the decision-making sequence via a snapshot captured at time step n.

2.2.2 Solution Concept

Since each agent n observes its private signal s_n and the actions of all predecessors G_{n-1}, then the *information set* of agent n is given by $I_n = \{s_n, G_{n-1}\}$. We denote the set of all possible information sets of agent n as $\mathcal{I}_n$. A pure strategy for agent n is a mapping $\sigma_n : \mathcal{I}_n \to \{0, 1, 2, \ldots, n-1\}$, whereas a strategy profile is defined as $\sigma = \{\sigma_n\}_{n\in\mathbb{N}}$. Since each agent acts sequentially and aims at maximizing its expected (Bayesian) payoff given the available information, it becomes natural to formulate the network formation process as a dynamic game with incomplete information, and adopt the Perfect Bayesian Equilibrium (PBE) as a solution concept, which is defined as follows.

Definition 2.1 *A strategy profile σ^* is a PBE if for every $n \in \mathcal{N}$, σ_n^* maximizes the expected payoff of agent n given the strategies of all other agents $\sigma_{-n}^* = \{\sigma_j^*\}_{j\in\mathcal{N}/\{n\}}$. That is, we have that*

$$\sigma_n^*(I_n) = \arg\max_{0\le j<n} \mathbb{E}\left[1_{\{q_j=H\}} \,|\, I_n\right] = \arg\max_{0\le j<n} \mathbb{P}\left(q_j = H \,|\, I_n\right).$$

Thus, in every equilibrium network, each agent links to the predecessor agent that has the maximum posterior probability of being of a high quality. Note that agent n's belief about the qualities of the predecessors depend on two pieces of information: her own private belief captured by the information in s_n, and the public belief that has formed by virtue of the actions of predecessors, and captured by the graph structure G_{n-1}. If agents lean towards following the public belief, then a *rational herd* behavior would emerge, which would affect both the network structure and its evolution. Since a herd behavior, or following the crowd with respect to link formation actions means that each agent will tend to rely on public beliefs to judge the predecessors' qualities, we expect that a form of preferential attachment would emerge in the network formation process as a direct consequence of the herd behavior.

2.2.3 Preferential Attachment and Information Cascades

The notions of preferential attachment and information cascades collapse into a single conceptualization in our model, since information cascades have to do with actions, which are essentially linking actions in our model, and preferential attachment has to do with the evolution of network structure, which is governed by these actions. In the following, we provide a unified general definition for preferential attachment and information cascades.

Definition 2.2 *We say that preferential attachment (or equivalently, an information cascade) governs the evolution of a network's sample path under a PBE σ^* of the network formation game, if there exists an agent $N < \infty$ after which all successors adopt a (behavioral) strategy that satisfies*

$$\mathbb{P}_{\sigma^*}\left(a_j = m \,|I_n\right) > \mathbb{P}_{\sigma^*}\left(a_j = k \,|I_n\right),$$

if $d_m(j) > d_k(j)$, $\forall j > N$. We say that condensed preferential attachment (or a condensed information cascade) occurs in σ^* if there exists an agent $N < \infty$ after which $a_j = a_k, \forall k, j > N$.

The definition above represents the central message of this paper: emergence of preferential attachment in a network with incomplete information is a structural manifestation for the emergence of an information cascade in the underlying decision-making process. Note that in the definition above, we provided a general definition for preferential attachment without specifying a functional form (e.g. linear [Barabási and Albert (1999); Vázquez (2003)], sub-linear [Gabel and Redner (2013)], or super-linear preferential attachment [Zhou et al. (2007); Krapivsky and Krioukov (2008)]). Since preferential attachment is driven by information (i.e. a probability of forming a link conditional on an *information set*), we need a characterization of the network's informational structure based on which we can study different regimes for the emergence of different forms of preferential attachment. This is achieved by describing the behavior of the agents' private beliefs, i.e. how private signals map to posterior beliefs on the predecessors' qualities.

2.2.4 Private Beliefs

An agent's belief comes from the observed public information, G_{n-1}, and the private signals s_n. The private belief only depends on the private signals, and is not a function of the strategy profile. Therefore, we can represent the *private belief* of agent n on agent j's quality as $\mathbb{P}\left(q_j = H \left| s_n^j \right.\right)$. Given the signal structure (F^H, F^L), we have that $\mathbb{P}\left(q_j = H \left| s_n^j \right.\right) = \left(1 + \frac{d\mathbb{F}^L}{d\mathbb{F}^H}(s_n^j)\right)^{-1}$, where $\frac{d\mathbb{F}^L}{d\mathbb{F}^H}(s)$ is the Radon-Nikodym derivative of $\mathbb{F}^L$ with respect to $\mathbb{F}^H$. The support of private beliefs is the interval $[b_l, b_h]$ such that $b_l = \inf\left\{v \in [0,1] : \mathbb{P}\left(p_s \leq v\right) > 0\right\}$, and $b_h = \sup\left\{v \in [0,1] : \mathbb{P}\left(p_s \leq v\right) < 1\right\}$, where $p_s = \mathbb{P}\left(q_j = H \left| s_n^j \right.\right)$. We identify two types of private beliefs that corresponds to two different classes of the network's informational structure:

- **Bounded private beliefs**: if $b_l > 0$ and $b_h < 1$.
- **Unbounded private beliefs**: $b_l = 0$ and $b_h = 1$.

Bounded private beliefs describe networks in which agents cannot shout out too loud, i.e. the private signals cannot be arbitrarily large (or small), whereas unbounded beliefs corresponds to networks where private beliefs can take arbitrary, unbounded values[9]. The two information

[9]One can think of citation networks as a network with bounded private beliefs, where papers that are not cited early on are forgotten, whereas Twitter can represent an unbounded private belief network, where users have the opportunity to post content and gain an arbitrarily large attention at any point in time.

structures can lead to different forms of preferential attachments, and different likelihoods of pathological outcomes.

In the rest of this paper, we will investigate the informational conditions (e.g. structure of private signals) under which preferential attachment will emerge in an equilibrium path.

2.3 Equilibrium Networks and the Emergence of Preferential Attachment

2.3.1 Posterior Beliefs

In this section, we characterize the equilibrium strategies of the agents in the network formation game. We start by formulating the *posterior beliefs* of the agents about the types of other agents given their information sets. Let the posterior belief of agent n about agent j's type in a strategy profile σ given her information set be denoted by $\mu_n^j(I_n)$. Using Bayes' rule, we have that

$$\mu_n^j(I_n) = \mathbb{P}(q_j = H \mid I_n) = \frac{d\mathbb{P}(I_n \mid q_j = H)\,\mathbb{P}(q_j = H)}{\sum_{q \in Q} d\mathbb{P}(I_n \mid q_j = q)\,\mathbb{P}(q_j = q)}. \tag{2.2}$$

We know that conditioned on the true quality, the private signal of n is independent of the observed decisions of the predecessors, which implies that

$$d\mathbb{P}\left(I_n \left| \omega_j^H\right.\right) = d\mathbb{P}\left(s_n \left| \omega_j^H\right.\right) \cdot \mathbb{P}\left(G_{n-1} \left| \omega_j^H\right.\right)$$

where ω_j^H represents the event $q_j = H$. Therefore, in order to construct her posterior belief, agent n needs to evaluate the probabilities that the observed graph and signals were generated by a high quality agent j. Since the type space Q comprises only two types, we can write the posterior belief $\mu_n^j(I_n)$ in terms of the likelihood ratios $\frac{\mathbb{P}(G_{n-1}|\omega_j^L)}{\mathbb{P}(G_{n-1}|\omega_j^H)}$ and $\frac{d\mathbb{P}(s_n|\omega_j^L)}{d\mathbb{P}(s_n|\omega_j^H)}$

as follows

$$\mu_n^j(I_n) = \mu_n^j(G_{n-1}, s_n) = \left[1 + \frac{1-p}{p} \cdot \frac{\mathbb{P}(G_{n-1} \mid \omega_j^L)}{\mathbb{P}(G_{n-1} \mid \omega_j^H)} \cdot \frac{d\mathbb{P}(s_n \mid \omega_j^L)}{d\mathbb{P}(s_n \mid \omega_j^H)}\right]^{-1} \tag{2.3}$$

In a PBE strategy profile, every agent forms a link with the one for whom she has the highest posterior belief. Therefore, after observing G_{n-1}, which includes all previous agents' actions, agent n can infer her predecessors' qualities by considering all the possible private signal profiles that they might receive.

2.3.2 Public and Private Likelihood

We define the *public likelihood of agent* j given a graph G_{n-1} as

$$\pi_n^j(G_{n-1}) = \frac{\mathbb{P}(G_{n-1} \mid \omega_j^H)}{\mathbb{P}(G_{n-1} \mid \omega_j^L)}.$$

The public likelihood of agent j given the graph G_{n-1} measures the likelihood of the graph G_{n-1} being generated under the case that agent j is of a high quality, as compared to the case that agent j is of a low quality. The evolution of the public likelihood process for the network agents affect the evolution of the successors' posterior beliefs (see equation (2.3)).

Since agent j's quality is the only determinant of the generating process for the signal s_n^j, and is independent of s_n^{-j}, then we have that $\frac{\mathbb{P}(s_n|\omega_j^H)}{\mathbb{P}(s_n|\omega_j^L)} = \frac{\mathbb{P}(s_n^j|\omega_j^H)}{\mathbb{P}(s_n^j|\omega_j^L)}$. We call this ratio a *private likelihood function*, and denote as $l(s_n^j) = \frac{\mathbb{P}(s_n^j|\omega_j^H)}{\mathbb{P}(s_n^j|\omega_j^L)} = \frac{d\mathbb{F}^H(s_n^j)}{d\mathbb{F}^L(s_n^j)}$. The private likelihood function $l(\cdot)$ only depends on the private signal structure, and is same for all agent j.

2.3.3 Equilibrium Strategies

Given the constructions for the public and private likelihood functions in Section 3.2, we conclude that after observing the graph G_{n-1} and the signals s_n, agent n's strategy is to evaluate the public likelihood based on G_{n-1}, and evaluate the private likelihood based on s_n, and then choose to link to the agent with the highest likelihood of being a high quality agent, since then it complies with the PBE equilibrium condition defined in Section 2.2. More precisely, we have the following proposition which provides the cornerstone for our analysis.[10]

Proposition 2.1 ***(Equilibrium strategies in terms of public and private likelihood).*** *The network formation game admits a generically unique PBE in pure strategies*[11]. *The posterior beliefs of every agent n depend only on the observed graph G_{n-1} and private signals s_n, $\mu_n^j(I_n, \sigma) = \mu_n^j(G_{n-1}, s_n)$. Moreover, the posterior beliefs and the public signals are related in the following way:*
(a) $\mu_n^j(G_{n-1}, s_n) \geq \mu_n^{j'}(G_{n-1}, s_n)$, if and only if
(b) $\pi_n^j(G_{n-1}) \cdot l(s_n^j) \geq \pi_n^{j'}(G_{n-1}) \cdot l(s_n^{j'})$, for any $j, j' \in \{0, 1, ..., n-1\}$.

Proposition 2.1 provides a decomposition of posteriors and a practical choice rule in equilibrium. Every agent n adopts the strategy that to link with the predecessor agent with the highest public-private likelihood product. With this decomposition we can further study the dynamics of beliefs and how it depends on the public and private information respectively.

2.3.4 Belief Dynamics

To characterize the posterior belief dynamics we need to track the evolution of the public likelihood [12]. The following Lemma describes the public likelihood dynamics in a recursive form.

Lemma 2.1 ***(Public likelihood dynamics).*** *The public likelihood of agent j evolves*

[10] All proofs are provided in the supplementary material.

[11] In knife-edge situations, there might be multiple PBEs in pure strategies that differ only in the tie-breaking rule.

[12] The private likelihood depends on the private signals, which are independent on the network structure.

according to the following recursive dynamics

$$\pi_{n+1}^{j}(G_n) = \begin{cases} \pi_n^j(G_{n-1}) \cdot R_n^j(G_{n-1}), & \textit{if } a_n = j \\ \pi_n^j(G_{n-1}) \cdot D_n^j(G_{n-1}), & \textit{if } a_n \neq j \end{cases} \tag{2.4}$$

with initial conditions $\pi_{j+1}^j(G_j) = 1$*, for any* $j \geq 0$*. The function* $R_n^j(G_{n-1})$ *(or* $D_n^j(G_{n-1})$*) indicate the change in the public likelihood after* j *being linked (or not being linked) by agent* n*.*

The function $R_n^j(G_{n-1})$, which we call the *reputation growth function*, measures the rate of increase in the public likelihood of agent j when it receives a link from agent n, whereas the function $D_n^j(G_{n-1})$, which we call the *reputation decay function*, measures the rate of decrease in the public likelihood of agent j when it misses the link formed by agent n. When the private beliefs are unbounded, we have that $R_n^j(G_{n-1}) = \frac{E_{s_n^{-j}}\{\Pi_{j' \neq j}[1-F^H(z)]\}}{E_{s_n^{-j}}\{\Pi_{j' \neq j}[1-F^L(z)]\}}$, and $R_n^j(G_{n-1}) = \frac{1-E_{s_n^{-j}}\Pi_{j' \neq j}(1-F^H(z))}{1-E_{s_n^{-j}}\Pi_{j' \neq j}(1-F^L(z))}$, where $z = l^{-1}(l(s_n^{j'})\pi_n^{j'}/\pi_n^j)$. Since the MLRP property implies the *first-order stochastic dominance*, we have that $R_n^j(G_{n-1}) > 1$ and $D_n^j(G_{n-1}) < 1$.

2.3.5 Emergence of Herd Behavior in Equilibrium Networks

Now that we have constructed the equilibrium strategy and characterized the evolution of the posterior belief process, can we figure out scenarios in which these strategies would give rise to a herd behavior in the link formation decisions? Intuitively, one can see that if the public likelihood of agent j given a graph G_{n-1} is high enough so that no private signal realization can push a higher posterior belief of another agent j', then agent n may ignore all her private signals and links to agent j who has the best reputation, even if her own private signal on agent j's quality is low.

The occurrence of a herding scenario in an equilibrium network depends on the network's informational structure (i.e. the private beliefs). For instance, if private beliefs are unbounded, then one expects that no public reputation can take over all the private signals forever, and there will always be a chance for a predecessor agent to express herself via a high private signal realization and gain a link even if her reputation in terms of the public belief is low. In the following Corollary, we link the private signal structure to the potential emergence of a herd behavior on the network evolution path.

Proposition 2.2 ***(Spotlights and Kick-outs).*** *If the private beliefs are unbounded, then any agent j has a positive probability to receive a link from agent n, for any $n > j$.*

(The spotlight condition) If the private beliefs are bounded, then an agent j will absorb all the links formed after time step $n-1$ whenever $\pi_n^j(G_{n-1}) max\{\pi_n^{j'}(G_{n-1})\}_{j' \neq j} > \beta$, for some $n \in \mathbb{N}$ and G_{n-1}, where $\beta = \frac{b_h(1-b_l)}{b_l(1-b_h)} > 1$. In other words, condensed preferential attachment will occur in the equilibrium and all successor agents will link to agent j.

(The kick-out condition) If the private belief is bounded, then an agent j will get new links with probability zero whenever $\pi_n^j(G_{n-1}) max\{\pi_n^{j'}(G_{n-1})\}_{j' \neq j} < \frac{1}{\beta}$, for some n and G_{n-1}.

Proposition 2.2 says that under the condition of unbounded private belief, it is informationally viable for any agent to gain the attention of successor agents and receive links from them, since unbounded private belief means that the private likelihood $l(\cdot)$ could be large enough with positive probability, and every agent (of a high or low quality) may send a high enough private signal to the agent currently forming the link.

In the second part of Proposition 2.2, we see that informationally limited networks (networks with bounded private beliefs) may exhibit serious pathological outcomes: under some conditions, one agent may absorb all the links and become the only hub in the network, only because she attained a large enough reputation, i.e. public likelihood, that makes it impossible for other agents to achieve higher posterior beliefs with respects to the successor agents given their bounded private signal realizations. Hence, condensed preferential attachment occurs as an extreme case of the rich-get-richer phenomena in Perc (2014), in which only one agent rich with links is getting all the new links. We call the condition at which such an equilibrium network emerges as the *spotlight condition.*

Note that the spotlight agent may not be a high quality agent, hence *asymptotic learning*[13] may not occur in a network with bounded private beliefs as all agents may end up

[13]See the definition of asymptotic learning in classical social learning settings in Acemoglu et al. (2011)

getting linked with a low quality agent, which means that due to the emerging herd behavior, the average network payoff will converge to zero as it grows asymptotically large. Similarly, the kick-out condition represent the case in which a -potentially high quality-agent gets forgotten because of her low reputation (low public likelihood), which leads her receiving no links from successor agents.

In the next subsection, we offer a more general characterization for the equilibrium networks, showing the emergence of more general forms of preferential attachment.

2.3.6 Emergence of Preferential Attachment in Equilibrium Networks

To derive the properties of our network formation game, we need to look into the dynamics of public likelihood of agents. Note that the functions $R(\cdot)$ and $D(\cdot)$ are defined on the set of all possible networks. In the proof of Lemma 2.1 we give a recursive algorithm to calculate these two functions for any network G_n. For bounded belief structure we can get the following characterization of the public likelihood dynamics.

Lemma 2.2 *When the private signals are discrete, the reputation growth function, the $R_n^j(G_{n-1})$, satisfies the following equation:*

$$R_n^j(G_{n-1}) = \Pi_{j'\neq j}\left\{\frac{\sum_{s_m\in S} P_m \phi^H(s_m, \pi_n^{j'}, \pi_n^j)}{\sum_{s_m\in S} P_m \phi^L(s_m, \pi_n^{j'}, \pi_n^j)}\right\} \tag{2.5}$$

where $s_m \in S$ represents all the possible signals of agent j' received by agent n; P_m is the prior of sending signal s_m; $\phi^q(s_m, \pi_n^{j'}, \pi_n^j)$ is the probability that the posterior of agent j is no less than that of agent j', given that the private signal about agentj$'$ is s_m, and two agents' public likelihoods are $\pi_n^j(G_{n-1})$ and $\pi_n^{j'}(G_{n-1})$. The reputation decay function $D_n^j(G_{n-1})$ satisfies

$$D_n^j(G_{n-1}) = \frac{1 - \Pi_{j'\neq j}\left[\sum_{s_m\in S} P_m \phi^H(s_m, \pi_n^{j'}, \pi_n^j)\right]}{1 - \Pi_{j'\neq j}\left[\sum_{s_m\in S} P_m \phi^H(s_m, \pi_n^{j'}, \pi_n^j)\right]} \tag{2.6}$$

Note that the reputation growth function is only well defined when agent j has positive

and Acemoglu et al. (2014).

probability of being linked by agent n. Otherwise, if the agent is kicked-out, the denominator in equation 2.5 is zero. Similarly, the reputation decay function is only well defined when an agent has not become the spotlight.

The function $\phi^q(s_m, \pi_n^{j'}, \pi_n^j)$ characterize the relative difficulty of agent j getting a new link given the current reputation and private signals about agent j'. The better signal s_m is, the smaller ϕ^q is: it's more difficult to be linked when a competitor is sending a better signal. The larger $\pi_n^{j'} \pi_n^j$ is, the smaller ϕ^q is: it's more difficult to get a new link when a competitor has relatively good reputation. Given any $(s_m, \pi_n^{j'}, \pi_n^j)$, the value of ϕ^H is weakly larger than ϕ^L, a result of the MLRP assumption. When there are discrete signals, the number of values in the image of function ϕ^q are finite.

Moreover, the reputation growth function is decrease in the public likelihood π_n^j. The explanation is following: when an agent has a low reputation while being linked by agent n, getting this new link will significantly increase this agent's public likelihood. When an agent already has a good reputation, getting a new link will not impact the reputation that much. In the extreme case, when an agent becomes the spotlight, $R_n^j = 1$: no matter how many new links this agent get, the public likelihood will no longer increase.

Similar properties apply for the reputation decay function. When an agent has a good reputation, not being linked by agent n will impose a negative shock on its public likelihood, therefore the reputation decrease largely. On the other hand, having a low reputation and not getting the new link will not hurt the reputation that much.

It's worth to point out that the set of public likelihood might be stable as the network increases. The growth rate of the reputation of getting the new link is decreasing with her reputation, and will be adjusted by the ratios to all other agents' reputation. The decreasing rate of not getting the new link is increasing with her reputation, and will be adjusted by the ratios to all other agents' reputation. Furthermore, at each period, there is a new agent with public likelihood equals to one joining the game. So we conjecture that the set of public likelihood should be uniformly bounded. However, we have not formally prove this.

With this detailed characterization of the dynamics of public likelihood, we can get the following result for the case of bounded private signals.

Theorem 2.1 *When the private beliefs of the signal structure are bounded, the condensed preferential attachment emerges with probability one.*

Theorem 2.1 implies that, in any network evolution process, there will almost always be one agent getting all the future new links. Here we briefly introduce the proof. First, at any period, if an agent get several links successively in the next several period, his public belief will be high enough so that this agent will get all the future links. Note that at any time if an agent get a new link his public belief will increase while all other agents' public belief will decrease (weakly). We can show that there is a upper bound of the number of successive links one agent needs. Second, at any period, the probability of a agent with relatively high public belief getting a fix number of successive links is bounded below. For example, to let the agent with highest public belief get all the future links, it is sufficient that new agents always receive high signal of her successively. The formal proof is in the appendix.

When the private belief is bounded, the preferential attachment always occurs. Theorem 2.1 also means that asymptotic learning may not occur in this social learning game on networks. As the network structure grows, the information conveyed by the number of links one agent has, might outweigh the information of private signals. Therefore, as new agents arrive, they will ignore their private information and believe more in the predecessors' choices.

When the private belief is unbounded, we conjecture that the probability of an agent of getting a new links will be proportional to an exponential function of the number of its links.

2.4 Extensions

In this section we extend the basic model into general settings. We will first allow agents to make more than one links. Then we analyze a special binary private signal structure.

2.4.1 Making a Fixed Number of Links

Now we relax the assumption that each agent links with one predecessors, and allow each agent to make more than one link. Denote agent n's link choice as A_n, where $A_n = (a_n^0, a_n^1, ..., a_n^{n-1})$ is a n dimensional vector. Let $a_n^j = 1$ if liking with agent j, otherwise $a_n^j = 0$. We first fix $\sum_j a_n^j = K \geq 2$ as a constant[14] and relax this assumption in next subsection.

Each agent maximizes his utility from every link, therefore he will link with the K predecessors with highest posteriors, given the information set $I_n = \{s_n, G_{n-1}\}$, where $G_{n-1} = (A_1, A_2, ...A_{n-1})$ is the existing network. Note that in this case the networkG_{n-1} is a $(n-1)$ by $(n-1)$ matrix rather than a vector. In this case, the separation between private and public likelihood still hold, we have the following lemma as the Lemma (2.1). in the basic model.

Lemma 2.3 *There existing an essentially unique PBE in pure strategies , in which agent links with the K predecessors with highest value of $(\pi_n^j(G_{n-1}) \cdot l(s_n^j))$.*

The public likelihood of agent j evolves according to the following equation

$$\pi_{n+1}^j(G_n) = \begin{cases} \pi_n^j(G_{n-1}) \cdot R_{n,j}^K(G_{n-1}) & \text{if } a_n^j = 1 \\ \pi_n^j(G_{n-1}) \cdot D_{n,j}^K(G_{n-1}) & \text{if } a_n^j = 0 \end{cases} \tag{2.7}$$

with initial conditions $\pi_{j+1}^j(G_j) = 1$, for any $j \geq 0$.

The two functions $R_{n,j}^K(\cdot)$ and $D_{n,j}^K(\cdot)$ in the above equation indicates the change in public likelihood after j is linked or not linked by agent n, respectively, given that agent n makes K links. The algorithm of calculating R^K and D^K, is a generalization of the previous algorithm. We basically change the use the Kth order distribution of the signal distribution to represent the agent's interpretation about a previous link, instead of the 1st order distribution in the basic model. We leave it in the proof of Lemma (2.3).

[14]For agent $j \leq K$, we assume that he makes j links.

With the evolution of posteriors, we can get another version of spotlight condition and kick-out condition.

Proposition 2.3 *If the private belief is bounded,*

(1) an agent will get all new links if his public likelihood satisfies a weaker spotlight condition, $\pi_n^j(G_{n-1})\pi_n^{(K)}(G_{n-1}) > \beta$*, where* $\pi_n^{(K)}(G_{n-1})$ *is the* K*th largest value in* $\{\pi_n^{j'}(G_{n-1})\}_{j' \neq j}$*;*

(2) an agent j *will get new links with zero probability whenever* $\pi_n^j(G_{n-1})K th\ max\{\pi_n^{j'}(G_{n-1})\}_{j' \neq j} < \frac{1}{\beta}$*, for some* n *and* G_{n-1}*;*

(3) asymptotic learning will not occur.

Note that the *spotlight condition* here is weaker than that of the basic model, while the *kick-out condition* is stronger. Intuitively, after observing a link with some agent i, the information one can get from this link is less than the basic model. Therefore the public likelihood increases slower and decreases slower than the basic model. To get all the following links, the requirement of public likelihood is weaker. On the other hand, since there are higher chance to get future links, only when an agent has very low public likelihood, he could be kicked out from the network for sure.

2.4.2 Arbitrary Number of Links

Now we assume every agent n can make x_n links $x_n \in \{1, 2, ..., n\}$. Then we have $\sum_j a_n^j = x_n$. When agent n arrives, he can observe the existing network G_n. Then he knows the number of all previous agents' links $x_1, x_2, ..., x_{n-1}$. Thus it doesn't matter whether x_n is predetermined or generated randomly, as long as we assume agent n knows his own x_n and $(x_1, x_2, ..., x_{n-1})$.

First note that similar with the basic model, the separation condition of public and private likelihood still holds.

Lemma 2.4 *Agent* n *links with the* x_n *predecessors with highest value of* $(\pi_n^j(G_{n-1}) \cdot l(s_n^j))$*.*

The public likelihood of agent j evolves according to the following equation

$$\pi^j_{n+1}(G_n) = \begin{cases} \pi^j_n(G_{n-1}) \cdot R^{x_n}_{n,j}(G_{n-1}) & \text{if } a^j_n = 1 \\ \pi^j_n(G_{n-1}) \cdot D^{x_n}_{n,j}(G_{n-1}) & \text{if } a^j_n = 0 \end{cases} \tag{2.8}$$

with initial conditions $\pi^j_{j+1}(G_j) = 1$, for any $j \geq 0$.

The evolution of public likelihood is similar. We leave it in the proof. The idea is that we use $R^{x_n}_{n,j}$ and $D^{x_n}_{n,j}$ instead of $R^K_{n,j}$ and $D^K_{n,j}$ to interpret agent n's linking choices.

In this case, since the number of future links is unknown, there is no condition for an agent to be kicked out forever. However, note that $x_n \geq 1$, the spotlight condition in Proposition 2.2 still holds. Similarly, the asymptotic learning will not occur with bounded private belief.

2.4.3 Bounded Binary Signals

Next we consider a special signal structure. Assume there are only two signals, *good* signal and *bad* signal. High type agent is more likely to send a good signal. Assume a high type agent could send a good signal with probability a, and a low type agent could send a good signal with probability b, and $a > b$. Then we can note the likelihood ratio as $l = \frac{a}{b} > 1$. Note $p_G = pa + (1-p)b$ as the *ex-ante* probability of a good signal. We also assume $x_n = 1$ here, and a general case is given in the appendix.

To get a closed form expression of the public likelihood, we need to calculate the rate R^j_n and D^i_n for any agent j. This change of agent j's public likelihood depends on how many agents have a public likelihood near to this agent. If there are many agents have similar public likelihoods with agent j, then π^j_n will increase a lot after being liked with agent n. And will also decrease a lot after not being link with agent n. More precisely, given any graph G_{n-1} and public likelihood $(\pi^j_n)_j$, define the following two numbers:

$$K_1(G_n, \pi^i_n) = \#\{j' \mid \pi^{j'}_n \in (\frac{1}{l^2}\pi^i_n, \pi^i_n]\} \text{ and } K_2(G_n, \pi^i_n) = \#\{j' \mid \pi^{j'}_n \in (\pi^i_n, \pi^i_n l^2]\}$$

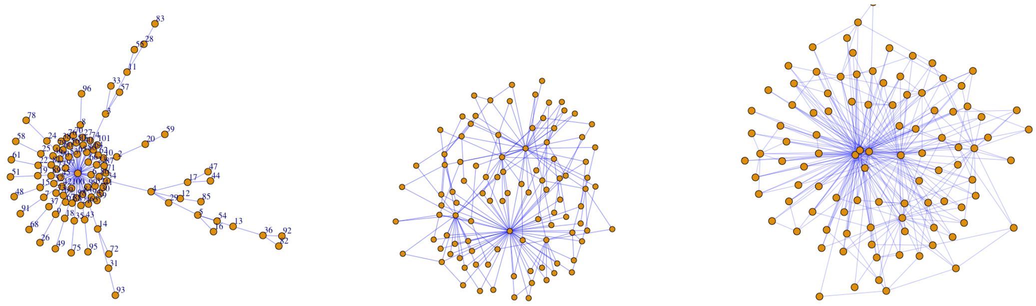


Figure 2.2: Some Networks
Each agent forms one link (left), 2 links (middle), $[log(n)+1]$ links (might) .

Lemma 2.5 *Given any graph G_{n-1} and public likelihood $(\pi_n^j)_j$, the change of agent i's public likelihood is determined by Lemma (2.1). Where R_n^i and D_n^i are given as following.*

$$R_n^i(\pi_n, G_{n-1}) = \left[\frac{p_G a + (1-p_G)}{p_G b + (1-p_G)}\right]^{K_1} \left(\frac{a}{b}\right)^{K_2},$$

.

$$D_n^i(\pi_n, G_{n-1}) = \frac{1 - [p_G a + (1-p_G)]^{K_1} [(1-p_G)\, a]^{K_2}}{1 - [p_G b + (1-p_G)]^{K_1} [(1-p_G)\, b]^{K_2}}$$

R_n^i is increasing in K_1, and K_2, and D_n^i is decreasing in K_1, and K_2.

We use the binary private signal structure to generate simulations. Figure (2.2) gives some sample networks in our model.

2.4.4 Who Will Be the Center?

In the case of bounded belief, eventually some agent will become the center. Figure (2.3) shows that the advantage of arriving earlier. In the left panel first two agents are both high type, while in more than 80% simulations one of them become the center of the network[15]. In the right panel, the first two agents are both low type, but the become the center in more than 80% simulations. By arriving early an agent can get more links in the initial stages, therefore accumulate a large public likelihood. Therefore he will have a larger chance to

[15]Note that since agent 1 will always link with agent 0, the first link contains no information about agent 0's type. The first two agents have exactly the same role in the model.

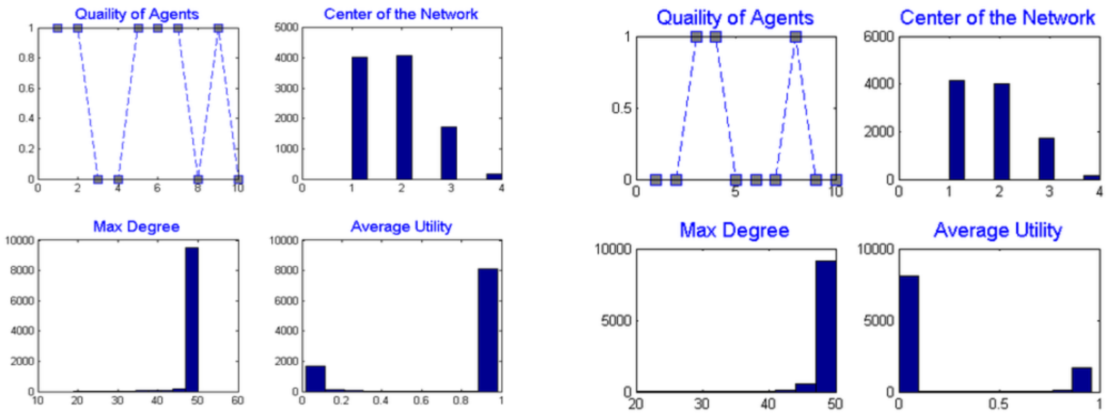


Figure 2.3: The first several agents always become the center
$p_H = 0.8$, $a = 0.6$, $l = 3$. *In each panel, the first graph the first 3 agents' types and we fix these types, the second graph shows the number of simulations in which an agent become the center. The third graph shows the number of simulations of a maximum degree of the center agent; The last graph show the number of simulations of a particular average utility.*

become the one with highest posterior in the future. The latter one agent arrives, the more agent he should compete with to get a link. On the other hand, with fewer links, it become even harder to have a high posterior.

However, the informational effect does exist in extreme case. In Figure (2.4), we let he agent $j \in \{0, 1, ...10, 12, ...20\}$ always send bad signal. Then agent 11 together with agent $21, 22, 23$ become the center of the network. Note that agent 11 is low type and he become the center only because he is the first agent who can send some good signals. This again shows that the above analysis that it's import to generate some good signals earlier than most of others.

2.5 Concluding Remarks

In this chapter, we present a social learning model where agents arriving sequentially to a network evaluate the qualities of predecessor agents based on their own private signals and public signals inferred from the network structure. In the network game, we characterize agents' optimal strategies in the Perfect Bayesian Equilibrium. A key feature of our model is that preferential attachment emerges endogenously as a sequentially rational equilibrium

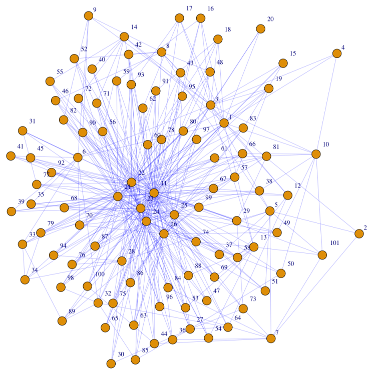


Figure 2.4: Signals matters in an extreme case
$x_n = [log(n)] + 1$, $p_H = 0.8$, $a = 0.6$, $l = 3$, $Q_{11} = Low$. *Agent 11 becomes the center when other agents before agent* 21 *will always send bad signals.*

of the social learning process. If agents' private beliefs are bounded, condensed preferential attachment emerges at equilibrium. We give a sufficient condition of the condensed preferential attachment and a sufficient condition of the case that an agent will be kicked out of the network. If agents' private belief are bounded, condensed preferential attachment emerges at equilibrium with probability one.

The future work lies in three directions. First, we need to generalize the main result of condensed preferential attachment to the case of multi links and random links. Second, we can reduce the information observed by a new agent, for example, with limited observations of the network structure. The assumption of an agents observing the whole network structure becomes unrealistic when the network grows large. Second, we need to study the relationship between the model prediction with the observed real life networks. Our conjecture about exponential preferential attachment is largely motived by these observations and needs to be verified.

Supplementary Material

Proof of Proposition 2.1

Agent n solves the problem

$$\max_{0\leq j<n} \mathbb{E}\left[\mathbb{I}_{\{j=H\}} \mid I_n\right] = \mu_n^j\left(I_n\right).$$

It only depends on the strategies of agent $j < n$. Since the values of posterior beliefs is bounded and the choice set is finite, the maximization problem must have a solution. If we apply the following tie-breaking rule $\sigma_n = \min\left\{\arg\max_{0\leq j<n} \mathbb{E}\left[\mathbb{I}_{\{j=H\}} \mid I_n\right]\right\}$. When the signal space is not discrete, since a tie only happens with zero probability, the equilibrium is still unique without such a tie-breaking rule.

Equation 2.3 gives the expression of posterior $\mu_n^j\left(I_n\right)$. Given the equilibrium strategy σ^*, when agent n forms a belief about agent j's type, she needs to consider every case of the private signal profiles of all previous agents. Therefore the exception in equation 2.3 is taken on $S_{n-1} = (s_1, ..., s_{n-1})$. Note that the probabilities $\mathbb{P}(G_{n-1} \mid \omega_j^q)$ is based on the exception of all $s_k^j, k \leq n$, then we can write the posterior as the following.

$$\begin{aligned}
\mu_n^j\left(I_n\right) &= \left[1 + \frac{1-p}{p} \cdot \frac{\mathbb{P}(G_{n-1} \mid \omega_j^L)}{\mathbb{P}(G_{n-1} \mid \omega_j^H)} \cdot \frac{d\mathbb{P}(s_n \mid \omega_j^L)}{d\mathbb{P}(s_n \mid \omega_j^H)}\right]^{-1} \\
&= \left[1 + \frac{1-p}{p} \cdot \left(\frac{\mathbb{P}(G_{n-1} \mid \omega_j^H)}{\mathbb{P}(G_{n-1} \mid \omega_j^L)} \cdot l(s_n^j)\right)^{-1}\right]^{-1}
\end{aligned} \tag{2.9}$$

Note $\tau = \frac{\mathbb{P}(G_{n-1}|\omega_j^H)}{\mathbb{P}(G_{n-1}|\omega_j^L)} \cdot l(s_n^j)$, and $g(\tau) = \left[1 + \frac{1-p}{p} \cdot (\tau)^{-1}\right]^{-1}$. The function $g(\cdot)$ is strictly increasing and concave.

Proof of Lemma 2.1

Given that $G_n = (G_{n-1}, a_n)$, we can rewrite the public likelihood $\pi_{n+1}^j(G_n)$ as

$$\frac{\mathbb{P}(G_n \mid \omega_j^H)}{\mathbb{P}(G_n \mid \omega_j^L)} = \frac{\mathbb{P}(G_{n-1} \mid \omega_j^H) \cdot \mathbb{P}(a_n \mid G_{n-1}, \omega_j^H)}{\mathbb{P}(G_{n-1} \mid \omega_j^L) \cdot \mathbb{P}(a_n \mid G_{n-1}, \omega_j^L)} = \pi_n^j(G_{n-1}) \cdot \frac{\mathbb{P}(a_n \mid G_{n-1}, \omega_j^H)}{\mathbb{P}(a_n \mid G_{n-1}, \omega_j^L)}. \tag{2.10}$$

Step 1.

When $a_n = j$, we can write the last term in above equation as,

$$\begin{aligned}
\frac{\mathbb{P}(a_n = j \mid G_{n-1}, \omega_j^H)}{\mathbb{P}(a_n = j \mid G_{n-1}, \omega_j^L)} &= \frac{E_{s_n^{-j}}\left[\mathbb{P}(a_n = j \mid G_{n-1}, \omega_j^H, s_n^{-j})\right]}{E_{s_n^{-j}}\left[\mathbb{P}(a_n = j \mid G_{n-1}, \omega_j^L, s_n^{-j})\right]} \\
&= \frac{E_{s_n^{-j}}\left[\mathbb{P}(\mu_n^j \geq \mu_n^{j'}, \forall j' \neq j \mid G_{n-1}, \omega_j^H, s_n^{-j})\right]}{E_{s_n^{-j}}\left[\mathbb{P}(\mu_n^j \geq \mu_n^{j'}, \forall j' \neq j \mid G_{n-1}, \omega_j^L, s_n^{-j})\right]}
\end{aligned}$$

The probability is of the agent n's private signals about agent j, s_n^j. Given G_{n-1}, by Proposition 2.1 the two probabilities can be expressed by the following, where $q \in q^H, q^L$.

$$\begin{aligned}
\mathbb{P}(\mu_n^j \geq \mu_n^{j'}, \forall j' \neq j \mid G_{n-1}, \omega_j^q, s_n^{-j}) &= \mathbb{P}\left[\pi_n^j(G_{n-1}) \cdot l(s_n^j) \geq \pi_n^{j'}(G_{n-1}) \cdot l(s_n^{j'}), \forall j' \neq j \mid G_{n-1}, \omega_j^q, s_n^{-j}\right] \\
&= \mathbb{P}\left[l(s_n^j) \geq \frac{\pi_n^{j'}(G_{n-1})}{\pi_n^j(G_{n-1})} l(s_n^{j'}), \forall j' \neq j \mid G_{n-1}, \omega_j^q, s_n^{-j}\right] \\
&= \Pi_{j' \neq j} \mathbb{P}\left[l(s_n^j) \geq \frac{\pi_n^{j'}(G_{n-1})}{\pi_n^j(G_{n-1})} l(s_n^{j'}) \mid G_{n-1}, \omega_j^q, s_n^{j'}\right]
\end{aligned}$$

The RHS of the inequality does not depend on s_n^j, since the probability is for the realization of s_n^j. Note $\mathbb{P}\left[l(s_n^j) \geq \frac{\pi_n^{j'}(G_{n-1})}{\pi_n^j(G_{n-1})} l(s_n^{j'}) \mid G_{n-1}, \omega_j^q, s_n^{j'}\right] = \phi^q\left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right)$, then we have

$$\frac{\mathbb{P}(a_n = j \mid G_{n-1}, \omega_j^H)}{\mathbb{P}(a_n = j \mid G_{n-1}, \omega_j^L)} = \frac{E_{s_n^{-j}}\left[\Pi_{j' \neq j} \phi^H\left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right)\right]}{E_{s_n^{-j}}\left[\Pi_{j' \neq j} \phi^L\left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right)\right]} \tag{2.11}$$

When $a_n \neq j$, the probability can be expressed as

$$\mathbb{P}\left[l(s_n^j) \leq \frac{\pi_n^{j'}(G_{n-1})}{\pi_n^j(G_{n-1})} l(s_n^{j'}),\ \text{for some } j' \neq j \mid G_{n-1}, \omega_j^q, s_n^{-j}\right] = 1 - \Pi_{j' \neq j} \phi^H \left(s_n^j, \pi_n^{j'}, \pi_n^j\right)$$

then we have

$$\frac{\mathbb{P}(a_n \neq j \mid G_{n-1}, \omega_j^H)}{\mathbb{P}(a_n \neq j \mid G_{n-1}, \omega_j^L)} = \frac{E_{s_n^{-j}}\left[1 - \Pi_{j' \neq j} \phi^H \left(s_n^j, \pi_n^{j'}, \pi_n^j\right)\right]}{E_{s_n^{-j}}\left[1 - \Pi_{j' \neq j} \phi^L \left(s_n^j, \pi_n^{j'}, \pi_n^j\right)\right]} \tag{2.12}$$

Step 2: Derive $\phi^q \left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right)$.

When the private belief is unbounded, $\phi^q \left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right) = F^q \left[l^{-1} \left(\frac{\pi_n^{j'}}{\pi_n^j} l\left(s_n^{j'}\right)\right)\right]$.

When the private belief is bounded, note that $l_1 = \frac{b_l}{1-b_l}$and $l_2 = \frac{b_h}{1-b_h}$, then $l(\cdot) \in [l_1, l_2]$. We also have that $0 < \mathbb{P}\left(l(s_n^j) \leq l_0\right) < 1$ if and only if $l_0 \in (l_1, l_2)$. Then consider the following three cases:

(a) If $\frac{\pi_k^{j'}}{\pi_k^j} l\left(s_n^{j'}\right) \geq l_2$, then $\phi^q \left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right) = \mathbb{P}\left[l(s_n^j) \geq \frac{\pi_n^{j'}}{\pi_n^j} l(s_n^{j'}) \mid G_{n-1}, \omega_j^q, s_n\right] = 0$.

(b) If $l_2 > \frac{\pi_k^{j'}}{\pi_k^j} l\left(s_n^{j'}\right) \geq l_1$, then $\phi^q \left(s_k^{i'}, \pi_k^{j'}, \pi_k^j\right) = 1 - F^q(z)$, where $z = l^{-1}(l(s_k^{j'})\pi_k^{j'}/\pi_k^j)$;

(c) If $l_1 > \frac{\pi_k^{j'}}{\pi_k^j} l\left(s_n^{j'}\right)$, then $\phi^q \left(s_k^{i'}, \pi_k^{j'}, \pi_k^j\right) = 1$.

Step 3.

To show that $R_n^j(G_{n-1}) > 1$ and $D_n^j(G_{n-1}) < 1$, it is sufficient to show $\frac{\phi^H\left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right)}{\phi^L\left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right)} > 1$ and $\frac{1-\Pi_{j' \neq j}\phi^H\left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right)}{1-\Pi_{j' \neq j}\phi^L\left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right)} < 1$ for any $s_n^{j'}, \pi_n^{j'}, \pi_n^j$. When the private belief is unbounded, $\phi^q \left(s_n^{i'}, \pi_n^{j'}, \pi_n^j\right) = 1 - F^q(z)$ where $z = l^{-1}\left(\frac{\pi_n^{j'}}{\pi_n^j} l\left(s_n^{j'}\right)\right)$. By the assumption of strictly MLRP we have F^H is a (strictly) first order stochastic dominance over F^L, i.e., $F^H(z) < F^L(z)$. Then we immediately have the results.

When the private belief is bounded, for case (a) and (c), $\phi^H \left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right) = \phi^L \left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right)$. For case (b), we have $\frac{\phi^H\left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right)}{\phi^L\left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right)} > 1$ and $\frac{1-\Pi_{j' \neq j}\phi^H\left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right)}{1-\Pi_{j' \neq j}\phi^L\left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right)} < 1$. Since the expectation is taken over all signals s_n^{-j}, case (b) must happen with positive probability, therefore the results still hold.

Proof of Proposition 2.2

If the private belief is unbounded, then for any $b \in (0,1)$, $0 < \mathbb{P}(p_s \leq b) < 1$, where $p_s = \mathbb{P}(q_j = H\,|s_n^j)$. Then we have that for any $l_0 \in (0,\infty)$, $0 < \mathbb{P}(l(s_n^j) \leq l_0) < 1$. This implies that $l(s_n^j)$ could be larger than any K with positive probability. From Proposition 2.1 we have that $a_n = j$ as long as $l(s_n^j) > \max_{j' \neq j}\left\{\frac{\pi_n^{j'}(G_{n-1})l(s_n^{j'})}{\pi_n^j(G_{n-1})}\right\}$, which happens with positive probability.

When the private belief is bounded, if $\pi_n^j(G_{n-1})\max\{\pi_n^{j'}(G_{n-1})\}_{j'\neq j} > \frac{b_h(1-b_l)}{b_l(1-b_h)}$, we have that $\pi_n^j(G_{n-1})\cdot\frac{b_l}{(1-b_l)} > \max\{\pi_n^{j'}(G_{n-1})\}_{j'\neq j}\cdot\frac{b_h}{(1-b_h)}$. Then $l_1 > \max\{\frac{\pi_n^{j'}(G_{n-1})\cdot l_2}{\pi_n^j(G_{n-1})\cdot}\}_{j'\neq j}$. Therefore the probability that agent j get a link from agent n is

$$E_{s_n^{-j}}\mathbb{P}\left[l(s_n^j) > \max_{j\neq j}\frac{\pi_n^{j'}(G_{n-1})}{\pi_n^j(G_{n-1})}l(s_n^{j'}) \mid G_{n-1}, s_n^{-j}\right] \geq \mathbb{P}\left[l_1 > \max_{j\neq j}\frac{\pi_n^{j'}(G_{n-1})}{\pi_n^j(G_{n-1})}l_2 \mid G_{n-1}\right] = 1$$

The first inequality follows from that $0 < \mathbb{P}(l(s_n^j) \leq l_0) < 1$ iff $l_0 \in (l_1, l_2)$.

If $\pi_n^j(G_{n-1})\max\{\pi_n^{j'}(G_{n-1})\}_{j'\neq j} < \frac{b_l(1-b_h)}{b_h(1-b_l)}$, we then have that $\pi_n^j(G_{n-1})\cdot\frac{b_h}{(1-b_h)} < \max\{\pi_n^{j'}(G_{n-1})\}_{j'\neq j}\cdot\frac{b_l}{(1-b_l)}$. Then $l_2 < \max\{\frac{\pi_n^{j'}(G_{n-1})\cdot l_1}{\pi_n^j(G_{n-1})\cdot}\}_{j'\neq j}$. Therefore the probability that agent j get a link from agent n is

$$E_{s_n^{-j}}\mathbb{P}\left[l(s_n^j) > \max_{j\neq j}\frac{\pi_n^{j'}(G_{n-1})}{\pi_n^j(G_{n-1})}l(s_n^{j'}) \mid G_{n-1}, s_n^{-j}\right] \leq \mathbb{P}\left[l_2 > \max_{j\neq j}\frac{\pi_n^{j'}(G_{n-1})}{\pi_n^j(G_{n-1})}l_1 \mid G_{n-1}\right] = 0.$$

Proof of Lemma 2.2

When the private signals are discrete, then for each $s_n^{j'} \in s_n^{-j}$, all the possible realization of $s_n^{j'}$ is the set S including discrete signals $s_m \in S$. Then we have

$$
\begin{aligned}
\mathbb{P}(a_n = j \mid G_{n-1}, \omega_j^H) &= E_{s_n^{-j}}\left[\Pi_{j' \neq j}\phi^H\left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right)\right] \\
&= \Sigma_{s_n^{-j} \in S^{N-2}}\left[\mathbb{P}\left(s_n^{-j}\right)\Pi_{j' \neq j}\phi^H\left(s_n^{j'}, \pi_n^{j'}, \pi_n^j\right)\right] \\
&= \Pi_{j' \neq j}\left[\sum_{s_m \in S} P_m \phi^H(s_m, \pi_n^{j'}, \pi_n^j)\right]
\end{aligned}
$$

The third equation is because of the independence of private signals. Then we can get the ratio of linking with j under two cases of agent j being high quality and low quality,

$$
R_n^j(G_{n-1}) = \frac{\mathbb{P}(a_n = j \mid G_{n-1}, \omega_j^H)}{\mathbb{P}(a_n = j \mid G_{n-1}, \omega_j^L)} = \Pi_{j' \neq j}\left\{\frac{\sum_{s_m \in S} P_m \phi^H(s_m, \pi_n^{j'}, \pi_n^j)}{\sum_{s_m \in S} P_m \phi^L(s_m, \pi_n^{j'}, \pi_n^j)}\right\} \tag{2.13}
$$

Similarly we have that,

$$
D_n^j(G_{n-1}) = \frac{1 - \mathbb{P}(a_n = j \mid G_{n-1}, \omega_j^H)}{1 - \mathbb{P}(a_n = j \mid G_{n-1}, \omega_j^L)} = \frac{1 - \Pi_{j' \neq j}\left[\sum_{s_m \in S} P_m \phi^H(s_m, \pi_n^{j'}, \pi_n^j)\right]}{1 - \Pi_{j' \neq j}\left[\sum_{s_m \in S} P_m \phi^H(s_m, \pi_n^{j'}, \pi_n^j)\right]} \tag{2.14}
$$

Proof of Theorem 2.1

When the private beliefs of the signal structure are bounded, we show that condensed preferential attachment emerges with probability one.

First we can rank all private signals according to the private likelihood function, $l(s_m)$. Note signal s_M has the highest private likelihood function and s_1 has the lowest private likelihood function. Intuitively, s_M is the most conclusive signal about a high type and s_1 is the lowest conclusive signal.

If $\pi_n^j \cdot l(s_1) > \max_{j'}\left\{\pi_n^{j'}\right\} \cdot l(s_M)$, then agent j will always get the new links. Since the public likelihood of agent j will not decrease and all other agents' public likelihood will not increase, the condition will continue to hold. Therefore it is a sufficient condition that agent j will get all the future links.

We will then provide a lower bound of R_n^j for the agent with the highest public likelihood with equation 2.13. Then we ask if this agent can be linked successively for several periods, whether the condition $\pi_n^j \cdot l(s_1) > \max_{j'} \left\{\pi_n^{j'}\right\} \cdot l(s_M)$ could be achieved.

Let $z_{jj'} = \frac{\pi_n^{j'}}{\pi_n^j}$, define each term in equation 2.13 as

$$\psi_j(z) \triangleq \frac{\sum_{s_m \in S} P_m \phi^H(s_m, \pi_n^{j'}, \pi_n^j)}{\sum_{s_m \in S} P_m \phi^L(s_m, \pi_n^{j'}, \pi_n^j)}.$$

When $z > \frac{l(s_M)}{l(s_1)}$, j cannot be linked when agent j' is available. Then $\psi_j(z)$ is only well defined for all $z \leq \frac{l(s_M)}{l(s_1)}$. Note $\frac{l(s_M)}{l(s_1)}$ as z_0. Also notice that $\psi_j(z)$ is no less than 1 for any $z \leq z_0$.

When $z_{jj'}$ is small, agent j has a good reputation relatively to agent j' then $\psi_j(z)$ is smaller. If $z \leq \frac{1}{z_0}$, then j' cannot be linked when agent j is available. Therefore $\psi_j(z) = 1$.

For any $s_m \in S$, the function $\phi^H(s_m, \pi_n^{j'}, \pi_n^j)$ takes value from the set

$$\Phi = \left\{0, F^H(s_1), F^H(s_2), \ldots, F^H(s_M), 1\right\},$$

depending on the value of $z_{jj'} = \frac{\pi_n^{j'}}{\pi_n^j}$. Not matter how the signals structure looks like, the image of $\psi_j(z)$ is a finite set. Given the MLRP assumption, $\psi_j(z) > 1$ for all $z \in (\frac{1}{z_0}, z_0]$. Then we can get a minimum value of function $\psi_j(z)$. Note as

$$\min_{z \in (\frac{1}{z_0}, z_0]} \{\psi_j(z)\} = 1 + \psi_0, \quad \psi_0 > 0$$

Then if agent j gets a new link, the public likelihood increase at a rate no less than ψ_0. If an agent get a new link for T periods, then the new public likelihood $\pi_{n+T}^j > (1+\psi_0)^T \pi_n^j$. Therefore we can find a finite T such that $(1+\psi_0)^T > z_0$.

Recall that agent j has the highest public likelihood at time n. Then at time $n+T$ we have the following condition:

$$\pi_{n+T}^j > (1+\psi_0)^T \pi_n^j > z_0 \pi_n^j \geq z_0 \max_{j'} \left\{\pi_n^{j'}\right\} \geq z_0 \max_{j'} \left\{\pi_{n+T}^{j'}\right\}$$

The last inequality is because all other agents' public likelihood stop increasing if they don't get any link from period n to period $n+T$. The above condition implies that agent j will become the spotlight agent after time $(n+T)$.

If the private signals about agent j from time n to $(n+T)$ are all s_M, then the aforementioned event will happen, independent of all other agents' private signals. There probability of agent j get T successive links from time n is no less than $P(s_M)^T$.

Now we can conclude that the condensed preferential attachment happens with probability one. Consider the probability of the opposite statement, which requires any T periods an event (agent j doesn't send T signals s_M) with probability less than $\left[1 - P(s_M)^T\right]$must happen.

The case of binary signals are considered in closed form solution in the proof of lemma 2.5. The requirement T in binary case is not very large. For example, when the probability of high type is $p_H = 0.7$, $p(s_G) = 0.6$, $p(s_B) = 0.4$, we have $T = 4$. When $p_H = 0.6$, $p(s_G) = 0.6$, $p(s_B) = 0.4$, then $T = 5$. For most values of p_H, $p(s_G)$, $p(s_B)$, T is less than 10. We also observe the condensed preferential attachment happens quickly in the simulations

Proof of Lemma 2.3 (Fixed Number of Links)

When each agent makes K links, we consider the event that agent j is linked by agent n, $\{a_n^j = 1\}$. The proof of Proposition 1 and the separation in equation (2.10) still hold. Now we need to consider the following two cases.

When $a_n^j = j$, we can write the last term in equation (2.10) as,

$$\frac{E_{S_n}\left[\mathbb{P}(a_n = j \mid G_{n-1}, \omega_j^H)\right]}{E_{S_n}\left[\mathbb{P}(a_n = j \mid G_{n-1}, \omega_j^L)\right]} = \frac{E_{s_n}\left\{E_{S_{n-1}}\mathbb{P}(\left|\{j' \text{ s.t. } \mu_n^j \geq \mu_n^{j'}\}\right| \geq n-K, \text{ given } G_{n-1}, \omega_j^H, s_n^{-j})\right\}}{E_{s_n}\left\{E_{S_{n-1}}\mathbb{P}(\left|\{j' \text{ s.t. } \mu_n^j \geq \mu_n^{j'}\}\right| \geq n-K, \text{ given } G_{n-1}, \omega_j^L, s_n^{-j})\right\}}$$

Given G_{n-1}, by Proposition 2.1 the two probabilities can be expressed by the following, where $q \in q^H, q^L$. **Given $G_{n-1}, \omega_j^H, s_n^{-j}$, the conditional probability of that there is no more than $(n-K)$ agents has posterior beliefs higher than agent j is the**

following:

$$\mathbb{P}(\left|\{j' \text{ s.t. } \mu_n^j \geq \mu_n^{j'}\}\right| \geq n-K) \quad = \quad \mathbb{P}\left[\left|l(s_n^j) \geq z(s_n^{-j}, \pi_n)\right| \geq n-K,\right]$$

where $z(s_n^{-j}, \pi_n) = \frac{\pi_n^{j'}(G_{n-1})}{\pi_n^j(G_{n-1})} l(s_n^{j'})$. Given the distribution of an individual signal s_n^j, we can get a Kth order distribution of the Kth largest value in $\{s_n^0, ..., s_n^{n-1}\}$. We note as $\psi_K^q(z) = \mathbb{P}\left[|l(s_n^j) \geq z| \geq n-K\right]$. Then we can get a similar expression as

$$\frac{E_{S_n}\mathbb{P}(a_n = j \mid G_{n-1}, \omega_j^H)}{E_{S_n}\mathbb{P}(a_n = j \mid G_{n-1}, \omega_j^L)} \quad = \quad \frac{E_{s_n^{-j}}\psi_K^H\left(z(s_n^{-j}, \pi_n)\right)}{E_{s_n^{-j}}\psi_K^L\left(z(s_n^{-j}, \pi_n)\right)} \tag{2.15}$$

When $a_n \neq j$, the probability can be expressed as

$$\frac{E_{S_n}\mathbb{P}(a_n \neq j \mid G_{n-1}, \omega_j^H)}{E_{S_n}\mathbb{P}(a_n \neq j \mid G_{n-1}, \omega_j^L)} \quad = \quad \frac{E_{s_n^{-j}}\left[1 - \psi_K^H\left(z(s_n^{-j}, \pi_n)\right)\right]}{E_{s_n^{-j}}\left[1 - \psi_K^L\left(z(s_n^{-j}, \pi_n)\right)\right]} \tag{2.16}$$

Proof of Proposition 2.3

When the private belief is bounded, if $\pi_n^j(G_{n-1})\pi_n^{(K)}(G_{n-1}) > \beta = \frac{b_h(1-b_l)}{b_l(1-b_h)}$, we have that $\pi_n^j(G_{n-1}) \cdot \frac{b_l}{(1-b_l)} > \pi_n^{(K)}(G_{n-1}) \cdot \frac{b_h}{(1-b_h)}$. Then $l_1 > K$th $\max\{\frac{\pi_n^{j'}(G_{n-1}) \cdot l_2}{\pi_n^j(G_{n-1}) \cdot}\}_{j' \neq j}$. Therefore the probability that agent j get a link from agent n is 1.

On the other hand, if $\pi_n^j(G_{n-1})\pi_n^{(K)}(G_{n-1}) < \frac{1}{\beta} = \frac{b_l(1-b_h)}{b_h(1-b_l)}$, we then have that $\pi_n^j(G_{n-1}) \cdot \frac{b_h}{(1-b_h)} < \pi_n^{(K)}(G_{n-1}) \cdot \frac{b_l}{(1-b_l)}$. Then $l_2 < K$th $\max\{\frac{\pi_n^{j'}(G_{n-1}) \cdot l_1}{\pi_n^j(G_{n-1}) \cdot}\}_{j' \neq j}$. Therefore the probability that agent j get a link from agent n is 0.

With bounded private belief, the probability the first condition is strictly positive even when $q_j = L$. Therefore the asymptotic learning will not occur even when T is very large.

Proof of Lemma 2.4 (Arbitrary Number of Links)

The proof of Lemma 2.3 applies to this case. The only difference is that agent n is making x_n links rather than K links. Therefore we need to substitute $\psi_K^q(z)$ with $\psi_{x_n}^q(z)$ in equation (2.15) and (2.16).

Proof of Lemma 2.5 (Binary Signals)

We first give the algorithm when $x_n = 1$ and then for the general case with arbitrary x_n. Recall that $P(\text{good} \mid H) = a$ and $P(\text{good} \mid L) = a$. We note that $l_G = \frac{a}{b}$, $l_B = \frac{1-a}{1-b}$, and $p_G = pa + (1-p)b$. Then $l_G > 1 > l_B$.

Making one link.

When the signals are binary, in equation (2.11) we have that

$$\phi^H(v) = \begin{cases} 0 & \text{if } v > l_G \\ a, & \text{if } v \in (l_B, l_G] \\ 1, & \text{if} v \leq l_B \end{cases}$$

and

$$\phi^L(v) = \begin{cases} 0 & \text{if } v > l_G \\ b, & \text{if } v \in (l_B, l_G] \\ 1, & \text{if} v \leq l_B \end{cases}$$

where $v = \frac{\pi_n^{j'}(G_{n-1})}{\pi_n^j(G_{n-1})} l(s_n^{j'})$. When $s_n^{j'} = good$, $v = \frac{\pi_n^{j'}}{\pi_n^j} l_G$. Therefore $z \in (l_B, l_G]$ is equivalent with $\frac{\pi_n^{j'}}{\pi_n^j} \in (\frac{l_B}{l_G}, 1]$. When $s_n^{j'} = bad$, $v = \frac{\pi_n^{j'}}{\pi_n^j} l_B$. Therefore $z \in (l_B, l_G]$ is equivalent with $\frac{\pi_n^{j'}}{\pi_n^j} \in (1, \frac{l_G}{l_B}]$.

If we define $z_{jj'} = \frac{\pi_n^{j'}}{\pi_n^j}$, and $z_0 = \frac{l_G}{l_B} = \frac{a(1-b)}{b(1-a)} > 1$, then we have the following result:

$$\psi_j(z) \triangleq \frac{p_G\phi^H(good,z)+p_B\phi^H(good,z)}{p_G\phi^H(bad,z)+p_B\phi^H(bad,z)} = \begin{cases} -(\text{not defined}) & \text{if } z > z_0 \\ \frac{a}{b}, & \text{if } z \in (1, z_0] \\ \frac{p_G a+(1-p_G)}{p_G b+(1-p_G)} & \text{if } z \in (\frac{1}{z_0}, 1] \\ 1, & \text{if} z \le \frac{1}{z_0} \end{cases} \tag{2.17}$$

Note that $R_n^j(G_{n-1}) = \Pi_{j'\neq j} \left\{ \frac{\sum_{s_m \in S} P_m \phi^H(s_m, \pi_n^{j'}, \pi_n^j)}{\sum_{s_m \in S} P_m \phi^L(s_m, \pi_n^{j'}, \pi_n^j)} \right\} = \Pi_{j'\neq j}\psi_j(z_{jj'})$. Then we can

$$K_1(G_{n-1}, \pi_n) = \#\{j' \mid \pi_n^{j'} \in (\frac{1}{z_0}\pi_n^j, \pi_n^j]\} \text{ and } K_2(G_{n-1}, \pi_n) = \#\{j' \mid \pi_n^{j'} \in (\pi_n^j, \pi_n^j z_0]\}$$

From equation (2.5) we can get

$$R_n^j(\pi_n, G_{n-1}) = \left[\frac{p_G a + (1-p_G)}{p_G b + (1-p_G)}\right]^{K_1} \left(\frac{a}{b}\right)^{K_2} \tag{2.18}$$

.

Apply the above expression of $\psi_j(z)$ to equation (2.6) we can get the result for the reputation decay function. If there exists j' such that $\pi_n^{j'} > \pi_n^j z_0$, then $D_n^j(\pi_n, G_{n-1}) = 1$. If for all j' we have $\pi_n^{j'} < \frac{1}{z_0}\pi_n^j$, then agent j has so good public likelihood that she must be linked. In this case, D_n^j is not defined. For all other cases where $K_1 + K_2 \geq 1$, we have

$$D_n^i(\pi_n, G_{n-1}) = \frac{1 - [p_G a + (1-p_G)]^{K_1} [(1-p_G)\, a]^{K_2}}{1 - [p_G b + (1-p_G)]^{K_1} [(1-p_G)\, b]^{K_2}}.$$

Making arbitrary number of links.

Given an existing graph G_{n-1} and agent n's choice A_n, let's look at the updating of posterior. We are interested in the change of reputation $R_n^j(\pi_n, G_{n-1})$ and $D_n^j(\pi_n, G_{n-1})$.

When agent j is high type,$\phi_{x_n}^H(z(s_n^{-j}, \pi_n))$ is the probability of the signals s_n^j such that

$(\pi_n^j l(s_n^j))$ larger than the x_n-th largest $(\pi_n^{j'} l'(s_n^{j'}))$ of all other l', note as $\left(\pi_n^{j'} l'(s_n^{j'})\right)^{(x_n)}$. Then $\phi_{x_n}^H(z) = 1$ if both good signal and bad signal can make $\pi_n^j l(s_n^j) \geq \left(\pi_n^{j'} l'(s_n^{j'})\right)^{(x_n)}$. If only good signal can make $\pi_n^j l(s_n^j) \geq \left(\pi_n^{j'} l'(s_n^{j'})\right)^{(x_n)}$, $\phi_{x_n}^H(z) = a$. Otherwise if no signal can make $\pi_n^j l(s_n^j)$ be one of the x_n largest value, $\phi_{x_n}^H(z) = 0$. Therefore $\psi_{x_n}^H(z)$ has only three values $1, a$ and 0.

When agent j is high type, $\psi_{x_n}^L(z(s_n^{-j}, \pi_n))$ has three values $1, b$ and 0, similarly. We summarize these three cases as following:

$$\begin{cases} \left(\pi_n^{j'} l'(s_n^{j'})\right)^{(x_n)} > \pi_n^j \cdot l_G & \text{no signal s.t. agent } n \text{ links with } j \\ \pi_n^j \cdot l_G \geq \left(\pi_n^{j'} l'(s_n^{j'})\right)^{(x_n)} > \pi_n^j \cdot l_B & \text{only good signal s.t. agent } n \text{ links with } j \\ \pi_n^j \cdot l_B \geq \left(\pi_n^{j'} l'(s_n^{j'})\right)^{(x_n)} & \text{both signals s.t. agent } n \text{ links with } j \end{cases}$$

We note the *ex-ante* probabilities of signals s_n^{-j} such that these three cases happen as p_0, p_1 and p_2. Then we can have

$$R_n^j = \frac{p_1 a + p_0}{p_1 b + p_0} \tag{2.19}$$

$$D_n^j = \frac{1 - [p_1 (1-a)] - p_0}{1 - [p_1 (1-b)] - p_0} \tag{2.20}$$

Since x_n is arbitrary, to calculate p_0 and p_1 we need to consider different cases of x_n. We first define

$$K_1 = \#\{j' \mid \pi_n^{j'} \in (\frac{1}{z_0}\pi_n^j, \pi_n^j]\},\ K_2 = \#\{j' \mid \pi_n^{j'} \in (\pi_n^j, \pi_n^j z_0]\} \text{ and } K_3 = \#\{j' \mid \pi_n^{j'} > \pi_n^j z_0\}.$$

- If $x_n \leq K_3$, no matter what others private signals $(s_n^{j'})_{j'}$ are, no signal s_n^j can make n link with j. We have $p_0 = 1$ and $p_1 = 0$.

- If $K_3 < x_n \leq K_3 + K_2$, $p_2 = 0$. And $p_0 = 1$ if all the agents j' in K_2 can send no less than $(x_n - K_3)$ number of good signals. This event happen with probability $\sum_{i=x_n-K_3}^{K_2} C_{K_2}^i p_G^i (1 - P_G)^{K_2 - i}$.

- If $K_3 + K_2 < x_n \leq K_3 + K_2 + K_1$, $p_0 = 0$. And $p_2 = 1$ if agents j' in K_1 can send no more than $(x_n - K_3 - K_2 - 1)$ number of good signals. This event happen with probability $\sum_{i=0}^{x_n-K_3-K_2-1} C_{K_1}^i p_G^i (1 - P_G)^{K_1 - i}$.

- If $K_3 + K_2 + K_1 < x_n$, no matter what others private signals $(s_n^{j'})_{j'}$ are, both signals s_n^j can make n link with j. We have $p_0 = 0$ and $p_2 = 1$.

Given any G_{n-1}, $(\pi_n^j)_j$ and x_n, it must be one of above four cases. Then we can use p_0 and p_1 in equations (2.19) and (2.20) to track the change of public likelihood. Therefore solve the agent $n+1$'s decision problem.

CHAPTER 3

Information Design in Contests[1]

3.1 Introduction

Contests are widely used to motivate economic agents to exert effort. Firms adopt contests to promote risky innovations,[2] the modern patent system is a contest by rewarding the property right to the first innovator. Many papers in economics focus on the design of reward structure in contests. However, in some cases, we have little leeway to choose the reward (monetary) and it is rather given.

Consider the case in school teachers cannot choose reward in terms of exam, students' ranking might be the only reward. In this case, what the teachers could do is to motivate through giving information to the students about their own and relative ability to the other students. The teachers usually may want to motivate the students to make efforts regardless of their ability. Needless to say, a similar situation happens in workplaces.

In this paper, we consider the setting where motivating takes place only through giving information. The reward structure of the contest is given. The designer of a contest chooses an information disclosure policy to maximize the total effort of participants.

We assume an ex-ante asymmericity of agents. There is a *strong* agent who has a larger probability to become a high ability agent, therefore can get a better outcome than the low ability, given the same effort level. The optimal information disclosure policy we characterize in Proposition 3.2 implies that the principal will discriminate two agents. She will disclose

[1]This chapter is a joint work with Daehyun Kim. (Kim: daehyunkim@ucla.edu, Department of Economics, UCLA.)

[2]Che and Gale (2003) and Halac et al. (2017) study the use of contests in promoting research and innovations. They also document many examples from both historical events and current business practice.

no information to the strong agent. When the weak agent has a disadvantage in abilities, the principal will partially disclose the state to him privately.

Comparing the no disclosure policy in benchmark case (Proposition 3.1), the optimal disclosure increase the total effort. We also consider the case when principal can only send public messages, the result (Proposition 3.4) shows no improvement to the benchmark case.

We also provide a linear programming approach (Proposition 3.3) to solve the optimal information disclosure policy, which is the straightforward method used in the information design literature. We expect to extend the approach to a more general setting with a richer set of actions (or states).

3.1.1 Literature Review

Most works in the literature about the design of contests focus on the design of reward structure.[3] Here we emphasis a few papers which focus on the information revelation in contests and the asymmericity between contestants.

Aoyagi (2010) studies the information revelation problem in a dynamic contest. The principal can observe a signal of agents past effort, and then chose a feedback policy to reveal this private signals to other agents. It discussed when the no-feedback policy and full-feedback policy are optimal for principal who wants to maximize the total effort. While our work provide a optimal revelation policy in-between the two extreme policies in a static environment. Zhang and Zhou (2016) also studies a information design problem in contest. In their model one agent's valuation is common knowledge while the other agent's valuation is private information. The designer chooses a information disclosure policy about the private valuation. It is essentially a persuasion problem between the designer and the uninformed agent. In our model the information disclosure is for all contestants.

[3]See Che and Gale (2003), Moldovanu and Sela (2006), Siegel (2009), Siegel (2014) and Halac et al. (2017) for the design of reward structure, i.e. "contest architecture", and the equilibrium behavior of contestants in different contests. See Moscarini and Smith (2007) for design of dynamic contests. Judd et al. (2012) studies the optimal rules of requirement to accomplishing a patent and the allocation of the benefits. Akcigit and Liu (2015) studies firms' strategy of whether to exert effort to the dead-end or to abandon risky projects in early stage in a contest of innovation. Szymanski and Valletti (2005) considers the effect of having a second prizes in contests.

Bimpikis et al. (2015) studies the design of the reward structure and the information disclosure policy in a dynamic contest. It provides a justification of the existence of an intermediate reward which can disclosure information about the status of competition, besides providing a reward. Gill (2008) considers the strategic information disclosure of a leading innovator in a patent contests. It finds when the development costs are high, the leading innovator has an incentive to signal its commitment to the ongoing project to deter potential rivals. In his model the principal (patent office) may not always want to maximize the total effort considering the wasteful effort of duplication R&D, which may reduce the social welfare.

Gershkov and Perry (2009) studies the relative importance between a midterm review and a final review in the design of the evaluation in a contest. Ederer (2010) studies the effect of a feedback policy in dynamic tournaments. It shows that a interim feedback also creates signal-jamming incentive prior to the evaluation. Smolin (2015) studies the optimal feedback policy in a principal-agent setting.

Kubitz (2015) considers a repeated contests with asymmetric players. It concludes that in general the aggregate output (hence effort) per contest in repeated setting is lower than that of a single contests. The private information of being a weak or strong players can lead to productive inefficiencies. Therefore it implies a potential social improvement of reveling private information in repeated settings. Other papers studies the asymmericity in contests includes Cornes and Hartley (2005), Parreiras and Rubinchik (2010), Siegel (2010), and Olszewski and Siegel (2016).

3.2 Model

We study a stylized model of contests: there are a designer and two agents $I = \{1, 2\}$. Each of the agents is endowed with ability θ_i, where $\theta_i \in \{\theta_H, \theta_L\}$, θ_H represents high ability and θ_L represents low ability. Assume $\theta_H > \theta_L$.

Agents can choose effort level $e_i \in \{0, 1\}$, then yields a outcome $y_i = e_i \cdot \theta_i$. The agent with higher outcome will win a reward $r > 0$, and we normalize $r = 1$. If both agents make effort higher ability agent wins; if their types are the same or no agent exert effort, a winner is randomly chosen. Exerting effort incurs a cost $c > 0$, agents want to make effort only when their is sufficient expected reward which would compensate the cost.

As we have the same payoffs when they are the same type; from now on we denote the state space by $\Omega = \{\omega_0, \omega_1, \omega_2\}$, where ω_0 represents the case that $\theta_1 = \theta_2$, ω_1 and ω_2 represent $\theta_1 > \theta_2$ and $\theta_1 < \theta_2$, respectively. Note that the payoff structure has the following properties: at ω_0, the marginal benefit of playing $e = 1$ is the same regardless of the opponent's play; at ω_1, the marginal benefit is larger when the opponent plays $e = 1$ (strategically complement): at ω_1, the marginal benefit is smaller (strategic substitute).

We assume two agents are ex-ante heterogeneous: they are endowed with high ability with different probabilities. Assume the prior of each agent being high type is given by μ_i. Without loss of generality we assume $\mu_1 > \mu_2$, where $\mu_1 = Prob(\theta_1 = \theta_H)$, $\mu_2 = Prob(\theta_2 = \theta_H)$. In this case, we call agent 1 a *strong* agent and agent 2 as *weak*. From them, we can induce $p_0 \equiv \Pr(\omega = \omega_0) = \mu_1\mu_2 + (1-\mu_1)(1-\mu_2)$; $p_1 \equiv \Pr(\omega = \omega_1) = \mu_1(1-\mu_2)$ and $p_2 \equiv \Pr(\omega = \omega_2) = (1-\mu_1)\mu_2$. Notice that $p_1 > p_2$.

The principal has an informational advantage in the sense that she can observe the true state, i.e., the relative value of two agents' ability.

The principal wants to maximize the total effort of two agents. For example, the employer want all employees to work hard, the central government prefers every local official to exert high effort. In this model, we focus on the information disclosure policy of the designer in contests. The principal first choose a information disclosure policy, which is common knowledge. Then she commits to this message structure. After receiving the messages, agents choose the effort level.

3.3 Optimal information disclosure

We explore the case where agents do not know their own type; this might be the case in the workplace when employees are just hired, they may have not realized their type, e.g., productivity is determined by how they are well matched to a task. On the other hand, it would be natural that the employer has better information as they saw many workers before.

3.3.1 Benchmark: No Information Disclosure

It is easy to see that our characterization crucially depends on the cost of effort as well as the prior. When the state is ω_1, it is always better for agent 1 to make an effort given the opponent does so. It is the opposite when the state is ω_2. When the state is ω_0, it depends on the cost. With sufficiently low cost, it is better to make an effort when the other agent does, and vice versa. Thus, whether $(1,1)$ is a Nash equilibrium depends on the prior and also the cost. We summarize the outcome of the contest under no information disclosure case in the following proposition.

Proposition 3.1 *When the cost of effort $c < \frac{1}{2}$. If $\frac{1}{2}p_0 + p_1 - c > 0$, then $e = 1$ is strictly dominant for both agents. If $\frac{1}{2}p_0 + p_i - c < 0$ for each $i = 1, 2$. Then $(0,0)$ is the unique Nash equilibrium with no information. Conversely, suppose $(1,1)$ is a Nash equilibrium with no information. Then $c < 1/2$ regardless of the priors.*

When the cost of effort $c > \frac{1}{2}$, there is no Nash equilibrium in which an agent plays $e = 1$. Therefore, the unique Nash equilibrium is $(0,0)$.

Proof. Given $c > \frac{1}{2}$, each agent is better not to exert effort given the opponent does not. Note that either $\frac{1}{2}p_0 + p_1 - c < 0$ or $\frac{1}{2}p_0 + p_2 - c < 0$, otherwise it contradicts to $c > \frac{1}{2}$. Assume the above is the case. Then, playing $e = 0$ is strictly dominant for agent 1, so agent 1 plays $e = 0$ in any NE. Given this, agent 2's best response is also to play $e = 0$. A symmetric argument applies to the second case. □

Therefore, only when $c < \frac{1}{2}$, there is room for manipulating information which incurs more effort in equilibrium. Also by Proposition 3.1, it should be the case that there is one and only one of agents satisfies $\frac{1}{2}p_0 + p_i - c < 0$, otherwise, $(1,1)$ is a NE and no further informational manipulation would be necessary. This implies that $\frac{1}{2}p_0 + p_1 - c > 0$ and

$\frac{1}{2}p_0 + p_2 - c < 0$ as agent 1 is stronger than agent 2 (i.e., $\mu_1 > \mu_2$). Note that $(0, 1)$ is a unique Nash equilibrium in this case.

When the two agents are ex-ante symmetric (i.e., $\mu_i = \frac{1}{2}$ for each i), a low cost $c < \frac{1}{2}$ ensures both agents playing $e = 1$ constitutes unique Nash equilibrium. In this case, the principal does not need to give any information to the agents.

3.3.2 Optimal Information Disclosure

Based on the previous discussion, from now on we focus on the case where $\frac{1}{2}p_0 + p_1 - c > 0$ and $\frac{1}{2}p_0 + p_2 - c < 0$.

First we argue that it is without loss to assume that agent 1 is given no information from the principal when it comes to maximizing the expected total effort. Recall that we have already found a Nash equilibrium where agent 1 makes an effort and agent 2 makes no effort at each state. Thus, we can see that there is no point of recommending $(1, 0)$ at any state. In addition, the Nash equilibrium is weakly better to recommend $(0, 1)$ at some state as both recommends total 1 of effort to the agents. This means we only need to consider either $(1, 0)$ or $(1, 1)$ at each state and this exactly means that the principal always recommends $e = 1$ to agent 1 or equivalently gives no information to agent 1.

Thus, we assume without loss of generality that agent 1 is not given any further information and we focus on agent 2's information. As agent 1 plays $e = 1$, agent 2's incentive compatibility with no further information is

$$\frac{1}{2}p_0 + p_2 - c < 0, \tag{3.1}$$

So agent 2 will make no effort. One way to improve this outcome would be to recommend agent 2 to play $e = 1$ only when the state is ω_2. Conditional on being recommended to play $e = 1$, agent 2's incentive compatibility constraint $1 - c > 0$ is clearly satisfied. Also, suppose that the principal recommends agent 2 only when the state is ω_0. Then, the IC constraint becomes $\frac{1}{2} - c > 0$, which is also satisfied. Together this means that if we recommend when the state is either ω_0 or ω_2, the IC constraint vis satisfied with slack:

$$\frac{1}{2}\frac{p_0}{p_0+p_2}+\frac{p_2}{p_0+p_2}-c>0. \tag{3.2}$$

Thus, we can make agent 2 make more effort by recommending agent 2 to play $e = 1$ even at ω_1 with some probability. For the optimality, we should increase the probability of recommendation at ω_1 until the IC is satisfied with equality:

$$\frac{1}{2}\frac{p_0}{p_0+p_1\alpha+p_2}+0\cdot\frac{p_1\alpha}{p_0+p_1\alpha+p_2}+\frac{p_2}{p_0+p_1\alpha+p_2}-c=0 \tag{3.3}$$

where $\alpha \in (0, 1]$ represents the conditional probability that the principal recommends $e = 1$ given ω_1. Then, because $\frac{1}{2}p_0 + p_2 - c < 0$, we have

$$\alpha=\frac{1}{p_1}\left(\frac{\frac{1}{2}p_0+p_2}{c}-p_0-p_2\right). \tag{3.4}$$

We also need to check after this chance agent 1's IC is still satisfied; comparing to the initial Nash equilibrium we started, as now agent 2 plays $e = 1$ with some probability at some states. However, remember that $e = 1$ is strictly dominant; therefore, still agent 1's IC is satisfied.

With the previous discussion, we have our main result:

Proposition 3.2 [Optimal Disclosure] *The optimal disclosure (including private disclosure) by the principal is characterized as follows:*

(i) The principal always recommends $e = 1$, and gives no further information to agent 1 (stronger agent);

(ii) For agent 2 (weak agent), the principal sends message $e = 1$ with probability 1 at state ω_0 and ω_2. When the state is ω_1, she sends message $e = 1$ with probability α as in equation 3.4, and sends $e = 0$ with probability $(1 - \alpha)$.

When $\mu_1 = 0.8, \mu_2 = 0.2$, the solid line in Figure 3.1 shows the total effort of two agents under different cost c. When the cost is between 0.2 and 0.5,the optimal disclosure result in

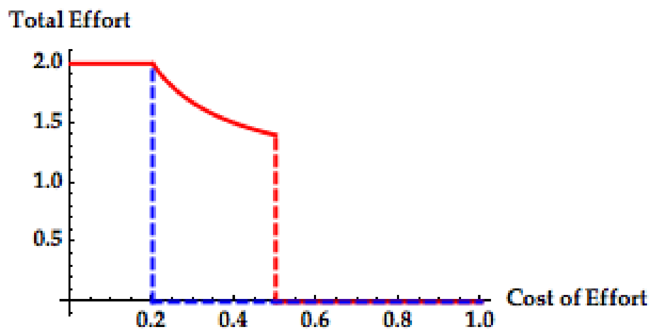


Figure 3.1: Total Effort under the Optimal Information Policy
($\mu_1 = 0.8$, $\mu_2 = 0.2$. The solid red line shows the total effort of two agents under the optimal disclosure policy. The blue dashed line is the total effort under no information disclosure.)

	$e_2 = 1$	$e_2 = 0$	$e_2 = 1$	$e_2 = 0$	$e_2 = 1$	$e_2 = 0$
$e_1 = 1$	a_0	b_0	a_1	b_1	a_2	b_2
$e_1 = 0$	c_0	d_0	c_1	d_1	c_2	d_2
	$\omega_0 : \ \theta_1 = \theta_2$		$\omega_1 : \ \theta_1 > \theta_2$		$\omega_2 : \ \theta_1 < \theta_2$	

Table 3.1: Probability of General Messages

a much higher total effort level than the benchmark case.

3.4 Extensions

Next we take a linear programing approach to solve the optimal information policy directly. This approach has the potential to work with a larger state space and a richer action set.

3.4.1 A Linear Programing Approach

We allow the designer to send general messages. According to the revelation principle, any information policy can be considered as a recommendation policy. The following table specifies the probability of each pair of actions in every state, e.g., a_1 represents the probability of sending message $(1, 1)$ that recommends both agents to make effort 1 when the state is $\theta_1 > \theta_2$. We then have $d_i = 1 - a_i - b_i - c_i$ for $i = 0, 1, 2$.

Denote p_i as the posterior of state ω_i, $i = 0, 1, 2$. Then $p_0 = \mu_1\mu_2 + (1-\mu_1)(1-\mu_2)$, $p_1 = \mu_1(1-\mu_2)$, $p_2 = (1-\mu_1)\mu_2$.

A recommendation policy should satisfy the incentive compatibility condition of two players. After receiving the recommendation from designer, each player will update his belief about the state, and then choose an effort.

When the designer recommend $e = 1$, the conditions such that both players will make an effort are:

Player 1:

$$p_0\left(\frac{1}{2}a_0 + b_0\right) + p_1\left(a_1 + b_1\right) + p_2 b_2 - P_1 c \geq p_0\left(\frac{1}{2}b_0\right) + p_1\left(\frac{1}{2}b_1\right) + p_2\left(\frac{1}{2}b_2\right) \qquad (1E)$$

Player 2:

$$p_0\left(\frac{1}{2}a_0 + c_0\right) + p_1 c_1 + p_2\left(a_2 + c_2\right) - P_2 c \geq p_0\left(\frac{1}{2}c_0\right) + p_1\left(\frac{1}{2}c_1\right) + p_2\left(\frac{1}{2}c_2\right) \qquad (2E)$$

Where $P_1 = p_0(a_0 + b_0) + p_1(a_1 + b_1) + p_2(a_2 + b_2)$ is the probability that effort 1 is recommended to player 1, and $P_2 = p_0(a_0 + c_0) + p_1(a_1 + c_1) + p_2(a_2 + c_2)$ is the probability that effort 1 is recommended to player 2.

When the designer recommends $e = 0$, the conditions such that both players will make no effort are:

Player 1:

$$p_0\left(\frac{1}{2}d_0\right) + p_1\left(\frac{1}{2}d_1\right) + p_2\left(\frac{1}{2}d_2\right) \geq p_0\left(\frac{1}{2}c_0 + d_0\right) + p_1\left(c_1 + d_1\right) + p_2\left(d_2\right) - \left(1 - P_1\right)c \qquad (1N)$$

Player 2:

$$p_0\left(\frac{1}{2}d_0\right)+p_1\left(\frac{1}{2}d_1\right)+p_2\left(\frac{1}{2}d_2\right)+\geq p_0\left(\frac{1}{2}b_0+d_0\right)+p_1\left(d_1\right)+p_2\left(b_2+d_2\right)-\left(1-P_2\right)c \quad (2N)$$

Note that we use total probabilities in these IC constraints. When the designer never recommend an effort to an agent, the respective IC constraint is not necessary.

The designer want to maximize

$$\max E(e_1+e_2)=p_0\left(2a_0+b_0+c_0\right)+p_1\left(2a_1+b_1+c_1\right)+p_2\left(2a_2+b_2+c_2\right) \tag{3.5}$$

This optimization problem is a linear programming problem. We can obtain the optimal solution by considering the value of each probabilities in the optimal policy, as the claims of the four steps in the proof of the following proposition. The result is the same as Proposition 3.2.

Proposition 3.3 *(1). If $c \leq \frac{1}{2}-\frac{1}{2}\Delta\mu$, the policy of no information disclosure can make $e=(1,1)$ the unique equilibrium.*

(2). If $c \in \left(\frac{1}{2}-\frac{1}{2}\Delta\mu,\ \frac{1}{2}\right)$, there is an unique optimal information disclosure policy, in which $a_1=1-\frac{(\mu_1-\mu_2)-(1-2c)}{2\mu_1(1-\mu_2)c}$, $b_1=\frac{(\mu_1-\mu_2)-(1-2c)}{2\mu_1(1-\mu_2)c}$, $a_2=a_0=1$. The two agents will make effort according to the designer's recommendation. The expected total effort is $E(e_1+e_2)=2-\frac{(\mu_1-\mu_2)-(1-2c)}{2c}$.

(3). If $c \geq \frac{1}{2}$, no message policy could satisfy all the IC conditions. The only equilibrium is $e=(0,0)$.

Proof. Part (1). When $c \leq \frac{1}{2}-\frac{1}{2}\Delta\mu$, with no information disclosure, the players use the priors (p_0,p_1,p_2) to evaluate the payoffs.

Given that player 2 will make effort with probability q_2, agent 1 will make effort if and only if

$$p_1 q_2 + (1 - q_2) + \frac{1}{2} p_0 q_2 - c > \frac{1}{2}(1 - q_2). \tag{3.6}$$

Then we have $\left(p_1 + \frac{1}{2}p_0 - \frac{1}{2}\right) q_2 > c - \frac{1}{2}$. To make $e = (1, 1)$ an pure strategy equilibrium, we need $p_1 + \frac{1}{2}p_0 > c$ and $p_2 + \frac{1}{2}p_0 > c$. When $c \leq \frac{1}{2} - \frac{1}{2}\Delta\mu$, both two conditions are satisfied.

Part (3). Given an arbitrary information policy, denote $A = p_0 a_0 + p_1 a_1 + p_2 a_2$, $B = p_0 b_0 + p_1 b_1 + p_2 b_2$, and $C = p_0 c_0 + p_1 c_1 + p_2 c_2$. Adding the LHS of $(1E)$ and $(2E)$ we have

$$(A + B + C) - (2A + B + C)\, c \geq \frac{1}{2}(B + C) \tag{3.7}$$

It implies $c \leq \frac{1}{2}$. When $c \geq \frac{1}{2}$, there does not exist any private message to induce effort 1 with positive probability. $\square$

Corollary 3.1 *If $c \in \left(\frac{1}{2} - \frac{1}{2}\Delta\mu,\ \frac{1}{2}\right)$, the probability of agent 2 not making effort under the optimal recommendation policy, $p_1 b_1$ is strictly increasing in c. The expected total effort is strictly decreasing in c. When fixing μ_1 or μ_2, $p_1 b_1$ is strictly increasing in $\Delta\mu = \mu_1 - \mu_2$, and the expected total effort is strictly decreasing in $\Delta\mu$.*

3.4.2 Public Message

Now we restrict to the case that the designer can only send public messages to the agents. Signal spaces can be arbitrary. Then a public message will result in a posterior p_n at state ω_n, $n \in \{0, 1, 2\}$.

So the agent i problem is to maximize the payoff given the other agent's choice of effort,

$$\text{Max}_{e_i} \qquad P(\theta_i e_i > \theta_j e_j \mid m) + \frac{1}{2} P(\theta_i e_i = \theta_j e_j \mid m) - c \cdot e_i \tag{3.8}$$

Agent i will choose $e_i = 1$ if and only if

$$P(\theta_i > \theta_j \mid m)P(e_j = 1) + P(e_j = 0) + \frac{1}{2}P(\theta_i = \theta_j \mid m)P(e_j = 1) - c > \frac{1}{2}P(e_j = 0). \quad (3.9)$$

For agent 1, the condition can be reduced to

$$\left(p_1 + \frac{1}{2}p_0 - \frac{1}{2}\right) P(e_2 = 1) > c - \frac{1}{2} \quad (3.10)$$

To make $e = (1, 1)$ an equilibrium, we need $\left(p_1 + \frac{1}{2}p_0 - \frac{1}{2}\right) > c - \frac{1}{2}$. Therefore $\left(p_1 + \frac{1}{2}p_0\right) > c$.

Compare the above condition with the benchmark case in Proposition 3.1, we can see the similarities between the conditions of these two problems. Furthermore, we can show that with public message the designer cannot do better than the no disclosure benchmark.

Proposition 3.4 *When the designer can only send public messages, she cannot get any equilibrium better than the no disclosure policy.*

Proof. When $c \leq \frac{1}{2}$, if $p_1 + \frac{1}{2}p_0 > c$, the policy of no information disclosure can make $e = (1, 1)$ the unique equilibrium. There are also other policies can achieve this equilibrium. Otherwise if $p_i + \frac{1}{2}p_0 < c$ for each i, then $(0, 0)$ is the unique equilibrium. The condition of exerting effort is the same as that of the benchmark case.

When $c > \frac{1}{2}$, $e = (0, 0)$ is always an equilibrium. No policy can make $e = (1, 1)$ or $(0, 1)$ pure strategy equilibria.

Suppose there exists a mix strategy equilibrium, then we need $\left(p_1 + \frac{1}{2}p_0 - \frac{1}{2}\right) P(e_2 = 1) = c - \frac{1}{2}$ and $\left(p_2 + \frac{1}{2}p_0 - \frac{1}{2}\right) P(e_1 = 1) = c - \frac{1}{2}$. Therefore $\left(p_1 + \frac{1}{2}p_0 - \frac{1}{2}\right) > 0$ and $\left(p_2 + \frac{1}{2}p_0 - \frac{1}{2}\right) > 0$. Since these two conditions will not hold at the same time. No policy can achieve mix strategy equilibrium. $\square$

Therefore, with only public messages, the principal can only yield the same payoff with the no disclosure case. It create an incentive for the designer of a contest to be able to communicate with individual participants privately.

3.5 Conclusion

We provide a model studying the optimal information design in contests. When there is asymmericity between contestants, the designer can increase the total effort by partially revealing information to the weak player.

We completely solve the simple model with two agents, two types, and two actions. Comparing with the no information disclosure benchmark, the optimal policy significantly increase the total effort level. Private information disclosure is important, since the public message can only yield the effort level as the benchmark case.

3.6 Appendix

Proof to Proposition 3.3

The optimization problem is a linear programming problem. First we can replace the probability d_i with $(1 - a_i - b_i - d_i)$. Then the problem becomes the following.

$$\max E(e_1 + e_2) = p_0\,(2a_0 + b_0 + c_0) + p_1\,(2a_1 + b_1 + c_1) + p_2\,(2a_2 + b_2 + c_2) \quad \text{s.t.}$$

$$\begin{pmatrix} \kappa p_0 & \kappa p_0 & 0 & p_1(1-c) & \kappa p_1 & 0 & \underline{-p_2 c} & \kappa p_2 & 0 \\ \kappa p_0 & 0 & \kappa p_0 & \underline{-p_1 c} & 0 & \kappa p_1 & p_2(1-c) & 0 & \kappa p_2 \\ p_0 & p_0 & 0 & p_1 & p_1 & \underline{-\frac{p_1}{2\kappa}} & p_2 & p_2 & \frac{p_2}{2\kappa} \\ p_0 & 0 & p_0 & p_1 & \frac{p_1}{2\kappa} & p_1 & p_2 & \underline{-\frac{p_2}{2\kappa}} & p_2 \end{pmatrix} X_1 \geq \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

where $X_1 = (a_0, b_0, c_0, a_1, b_1, c_1, a_2, b_2, c_2)^T$, $\kappa = \frac{1}{2} - c$. Note that there are extra constraints that $a_i + b_i + c_i \leq 1$ and $a_i, b_i, c_i \geq 0$ for $i = 0, 1, 2$.

Claim 1 *In the optimal solution, $d_0 = d_1 = d_2 = 0$.*

First observe the coefficients of a_0, b_1, c_2 in all constraints are nonnegative. We underline an element in the above matrix if it's negative. Then if $d_0 > 0$, we can decrease d_0 and increase a_0 by the same amount; if $d_1 > 0$, we can decrease d_1 and increase b_1; if $d_2 > 0$, we can decrease d_2 and increase c_2. Each of above adjustment will leave the inequality still hold and results in a weakly higher object function.

Claim 2 *The IC constraint for player 2 to make effort (2E), and that for player 1 to not make effort (1N), are binding under the optimal solution.*

Since $d_i = 0$, we can replace c_i with $(1 - a_i - b_i)$. The problem then becomes

$$\max E(e_1 + e_2) = 1 + p_0 a_0 + p_1 a_1 + p_2 a_2 \quad \text{s.t.}$$

$$\begin{pmatrix} \kappa p_0 & \kappa p_0 & p_1(1-c) & \kappa p_1 & \underline{-p_2 c} & \kappa p_2 \\ 0 & \underline{-\kappa p_0} & \underline{-\frac{1}{2}p_1} & \underline{-\kappa p_1} & \frac{1}{2}p_2 & -\kappa p_2 \\ p_0 & p_0 & p_1(1+\frac{1}{1-2c}) & p_1(1+\frac{1}{1-2c}) & \underline{p_2(1-\frac{1}{1-2c})} & \underline{p_2(1-\frac{1}{1-2c})} \\ 0 & \underline{-p_0} & 0 & p_1(\frac{1}{1-2c}-1) & 0 & \underline{p_2(-\frac{1}{1-2c}-1)} \end{pmatrix} X_2 \geq \begin{pmatrix} 0 \\ -(\frac{1}{2}-c) \\ 1+\frac{p_1-p_2}{1-2c} \\ 0 \end{pmatrix}$$

where $X_2 = (a_0, b_0, a_1, b_1, a_2, b_2)^T$.

Suppose the second condition (IC of 2E) is not binding. If $a_1 \neq 1$, we can increase a_1 a little bit and decrease c_1. Otherwise if $a_1 = 1$ and $a_0 \neq 1$, we can increase a_0 and decrease b_0. Otherwise if $a_2 \neq 1$ and $a_0 = a_1 = 1$, we can increase a_2 and decrease b_2 by the same amount and it will not affect the first condition. All these adjustments will make the IC constraints still hold and get a higher effort. We argue that $a_0 = a_1 = a_2 = 1$ cannot be the optimal solution. Otherwise the second condition becomes $\frac{1}{2}(p_2 - p_1) > -(\frac{1}{2} - c)$. Therefore $c < \frac{1}{2} - \frac{1}{2}(p_1 - p_2) = \frac{1}{2} - \frac{1}{2}\Delta\mu$, which violates the assumption of cost.

Suppose the third condition (IC of 1N) is not binding. If $a_0 \neq 1$, then we can increase a_0 and decrease b_1 or c_1. Otherwise if $a_0 = 1$ and $a_1 \neq 1$, we can increase a_1 a little bit and decrease b_1 and c_1, while leaving the second condition unchanged. Otherwise if $a_2 \neq 1$ and $a_0 = a_1 = 1$, we can increase a_2 and decrease b_2 by the same amount, and this will not affect the first condition. From above argument we know that $a_0 = a_1 = a_2 = 1$ cannot be the optimal solution.

Claim 3 *In the optimal solution, $b_0 = c_0 = 0$, $a_0 = 1$.*

If $b_0 \neq 0$, we can increase a_0 and decrease b_0 by the same amount and all the constraints still hold. Note the coefficient of a_0 are all nonnegative, therefore if $c_0 \neq 0$, we can always increase a_0 and leave the constraint unchanged. Therefore in the optimal solution $a_0 = 1$.

Divide both sides of the first two constraints by $(\frac{1}{2} - c)$, and note $\tau = \frac{1}{1-2c} = \frac{1}{2\kappa}$. The problem becomes

$$\max\ E(e_1 + e_2) = 1 + p_0 + p_1 a_1 + p_2 a_2 \ \text{ s.t.}$$

$$\begin{pmatrix} p_1(1+\tau) & p_1 & \underline{-p_2 2c\tau} & p_2 \\ \underline{-p_1\tau} & \underline{-p_1} & p_2\tau & -p_2 \\ p_1(1+\tau) & p_1(1+\tau) & \underline{p_2(1-\tau)} & \underline{p_2(1-\tau)} \\ 0 & p_1(\tau-1) & 0 & \underline{p_2(-\tau-1)} \end{pmatrix} X_3 \geq \begin{pmatrix} -p_0 \\ -1 \\ p_1(1+\tau) + p_2(1-\tau) \\ 0 \end{pmatrix}$$

where $X_3 = (a_1, b_1, a_2, b_2)^T$.

Claim 4 *In the optimal solution, $c_1 = c_2 = 0$.*

If $c_1 \neq 0$, then we can decrease c_1 by 2ϵ, and increase both a_1 and b_1 by ϵ. Given that the third constraint is binding, we have $p_1(1+\tau)c_1 = p_2(\tau - 1)c_2$. Therefore we should also decrease c_2 by $\frac{p_1(1+\tau)}{p_2(\tau-1)} \cdot 2\epsilon$. Then we increase both a_2 and b_2 by $\frac{p_1(1+\tau)}{p_2(\tau-1)} \cdot \epsilon$. Now the change in the LHS of (2E) is $\Delta = -p_1(1+\tau)\epsilon + p_2(\tau-1) \cdot \frac{p_1(1+\tau)}{p_2(\tau-1)} \cdot \epsilon = 0$. The other two constraints still hold. Thus $c_1 = 0$ and from the third constraint, $c_2 = 0$.

Now we can ignore the third constraint and the problem can be reduced into

$$\max\ E(e_1 + e_2) = 1 + p_0 + p_1 a_1 + p_2 a_2 \ \text{ s.t.}$$

$$\begin{pmatrix} p_1\tau & \underline{-p_2\tau} \\ \underline{-p_1(\tau-1)} & p_2(\tau+1) \\ \underline{-p_1(\tau-1)} & p_2(\tau+1) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \geq \begin{pmatrix} -1 \\ -1+p_1+p_2 \\ p_2(\tau+1)-p_1(\tau-1) \end{pmatrix}$$

The third constraint is redundant because $-1+p_1+p_2 \geq p_2(\tau+1)-p_1(\tau-1)$ implies $c \geq \frac{1}{2} - \frac{1}{2}\Delta\mu$. Since the second constraint is binding, then $p_1(\tau-1)a_1 = p_2(\tau+1)a_2 + p_0$. Therefore the optimal solution must be a corner solution.

If $a_2 = 1$, then $a_1 = \frac{p_2(\tau+1)+p_0}{p_1(\tau-1)}$. This solution satisfies the other two constraint. (The first condition implies $a_1 \geq \frac{p_2+2c-1}{p_1}$, then we have $p_2 + p_0 \geq -p_2 - 1 + \frac{1}{\tau}$ which is true.

If $a_1 = 1$, then $a_2 = \frac{p_1(\tau-1)-p_0}{p_2(\tau+1)}$. But $\frac{p_1(\tau-1)-p_0}{p_2(\tau+1)} = \frac{p_1\tau+p_2-1}{p_2\tau+p_2} > 1$, since $\tau > \frac{1}{p_1-p_2}$.

The optimal solution is $a_2 = 1$, $b_2 = 0$, $a_1 = \frac{p_2(\tau+1)+p_0}{p_1(\tau-1)}$, and $b_1 = 1 - a_1. This is the optimal disclosure po$

Bibliography

Acemoglu, Daron, Kostas Bimpikis, and Asuman Ozdaglar, "Dynamics of information exchange in endogenous social networks," Theoretical Economics, 2014, 9 (1), 41–97.

_ , Munther A Dahleh, Ilan Lobel, and Asuman Ozdaglar, "Bayesian learning in social networks," The Review of Economic Studies, 2011, 78 (4), 1201–1236.

Admati, Anat R and Paul Pfleiderer, "A monopolistic market for information," Journal of Economic Theory, 1986, 39 (2), 400–438.

_ and _ , "Direct and indirect sale of information," Econometrica: Journal of the Econometric Society, 1990, pp. 901–928.

Akcigit, Ufuk and Qingmin Liu, "The role of information in innovation and competition," Journal of the European Economic Association, 2015, 14 (4), 828–870.

Aoyagi, Masaki, "Information feedback in a dynamic tournament," Games and Economic Behavior, 2010, 70 (2), 242–260.

Banerjee, Abhijit V, "A simple model of herd behavior," The quarterly journal of economics, 1992, 107 (3), 797–817.

Barabási, Albert-László, "Network science: Luck or reason," Nature, 2012, 489 (7417), 507.

_ and Réka Albert, "Emergence of scaling in random networks," science, 1999, 286 (5439), 509–512.

Bardone, Emanuele, Seeking chances: From biased rationality to distributed cognition, Vol. 13, Springer Science & Business Media, 2011.

Bergemann, Dirk, Alessandro Bonatti, and Alex Smolin, "The design and price of information," American Economic Review, 2018, 108 (1), 1–48.

_ and _ , "Markets for Information: An Introduction," Cowles Foundation for Research in Economics Discussion paper, 2018.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," Journal of political Economy, 1992, 100 (5), 992–1026.

Bimpikis, Kostas, Davide Crapis, and Alireza Tahbaz-Salehi, "Information Sale and Competition," Management Science, 2019.

_ , Shayan Ehsani, and Mohamed Mostagir, "Designing Dynamic Contests.," 2015, pp. 281–282.

Board, Simon and Jay Lu, "Competitive information disclosure in search markets," Journal of Political Economy, 2018, 126 (5), 1965–2010.

Boleslavsky, Raphael, Ilwoo Hwang, and Kyungmin Kim, "Competitive Advertising and Pricing," Unpublished, 2018.

Che, Yeon-Koo and Ian Gale, "Optimal design of research contests," The American Economic Review, 2003, 93 (3), 646–671.

Cornes, Richard and Roger Hartley, "Asymmetric contests with general technologies," Economic theory, 2005, 26 (4), 923–946.

D'souza, Raissa M, Christian Borgs, Jennifer T Chayes, Noam Berger, and Robert D Kleinberg, "Emergence of tempered preferential attachment from optimization," Proceedings of the National Academy of Sciences, 2007, 104 (15), 6112–6117.

Ederer, Florian, "Feedback and motivation in dynamic tournaments," Journal of Economics & Management Strategy, 2010, 19 (3), 733–769.

Even-Dar, Eyal and Michael Kearns, "A small world threshold for economic network formation," in "Advances in Neural Information Processing Systems" 2007, pp. 385–392.

Frankel, Alexander and Emir Kamenica, "Quantifying information and uncertainty," Unpublished, 2018.

Freno, Antonino, Mikaela Keller, and Marc Tommasi, “Fiedler random fields: A large-scale spectral approach to statistical network modeling,” in “Advances in Neural Information Processing Systems” 2012, pp. 1862–1870.

Gabel, Alan and Sidney Redner, “Sublinear but never superlinear preferential attachment by local network growth,” Journal of Statistical Mechanics: Theory and Experiment, 2013, 2013 (02), 02–043.

Gentzkow, Matthew and Emir Kamenica, “Costly persuasion,” American Economic Review, 2014, 104 (5), 457–62.

_ and _ , “Competition in persuasion,” The Review of Economic Studies, 2016, 84 (1), 300–322.

Gershkov, Alex and Motty Perry, “Tournaments with midterm reviews,” Games and Economic Behavior, 2009, 66 (1), 162–190.

Gill, David, “Strategic disclosure of intermediate research results,” Journal of Economics & Management Strategy, 2008, 17 (3), 733–758.

Golosovsky, Michael and Sorin Solomon, “Stochastic dynamical model of a growing citation network based on a self-exciting point process,” Physical Review Letters, 2012, 109 (9), 098701.

Gopalan, Prem K, Chong Wang, and David Blei, “Modeling overlapping communities with node popularities,” in “Advances in neural information processing systems” 2013, pp. 2850–2858.

Halac, Marina, Navin Kartik, and Qingmin Liu, “Contests for experimentation,” Journal of Political Economy, 2017, 125 (5), 1523–1569.

Jackson, Matthew O and Brian W Rogers, “Meeting strangers and friends of friends: How random are social networks?,” American Economic Review, 2007, 97 (3), 890–915.

Judd, By Kenneth L, Karl Schmedders, and Şevin Yeltekin, "Optimal rules for patent races," International Economic Review, 2012, 53 (1), 23–52.

Kakade, Sham M, Michael Kearns, Luis E Ortiz, Robin Pemantle, and Siddharth Suri, "Economic properties of social networks," in "Advances in Neural Information Processing Systems" 2005, pp. 633–640.

Kim, Kyungmin and Philipp Kircher, "Efficient competition through cheap talk: the case of competing auctions," Econometrica, 2015, 83 (5), 1849–1875.

Kown, Suehyun, "Selling Complementary Goods: Information and Products," Working paper, 2018.

Krapivsky, Paul and Dmitri Krioukov, "Scale-free networks as preasymptotic regimes of superlinear preferential attachment," Physical Review E, 2008, 78 (2), 026114.

Kubitz, Greg, "Repeated Contests with Private Information," Technical Report, Mimeo, November 2015.

Leskovec, Jure, Lars Backstrom, Ravi Kumar, and Andrew Tomkins, "Microscopic evolution of social networks," 2008, pp. 462–470.

Li, Fei and Peter Norman, "Sequential Persuasion," Unpublished, 2018.

Mirkin, Kenneth and Marek Pycia, "Directed Search and the Futility of Cheap Talk," Working paper, 2016.

Moldovanu, Benny and Aner Sela, "Contest architecture," Journal of Economic Theory, 2006, 126 (1), 70–96.

Moscarini, Giuseppe and Lones Smith, "Optimal dynamic contests," Science (June), 2007, pp. 1–18.

Olszewski, Wojciech and Ron Siegel, "Large contests," Econometrica, 2016, 84 (2), 835–854.

Papadopoulos, Fragkiskos, Maksim Kitsak, M Ángeles Serrano, Marián Boguná, and Dmitri Krioukov, "Popularity versus similarity in growing networks," Nature, 2012, 489 (7417), 537.

Parreiras, Sérgio O and Anna Rubinchik, "Contests with three or more heterogeneous agents," Games and Economic Behavior, 2010, 68 (2), 703–715.

Perc, Matjaž, "The Matthew effect in empirical data," Journal of The Royal Society Interface, 2014, 11 (98), 20140378.

Ronnen, Uri, "Minimum quality standards, fixed costs, and competition," The RAND Journal of economics, 1991, pp. 490–504.

Siegel, Ron, "All-Pay Contests," Econometrica, 2009, 77 (1), 71–92.

_ , "Asymmetric contests with conditional investments," The American Economic Review, 2010, 100 (5), 2230–2260.

_ , "Asymmetric all-pay auctions with interdependent valuations," Journal of Economic Theory, 2014, 153, 684–702.

Smith, Lones and Peter Sørensen, "Pathological outcomes of observational learning," Econometrica, 2000, 68 (2), 371–398.

Smolin, Alex, "Optimal feedback design," November, 2015, 4, 11–17.

Spence, Michael, "Product selection, fixed costs, and monopolistic competition," The Review of economic studies, 1976, 43 (2), 217–235.

Szymanski, Stefan and Tommaso M Valletti, "Incentive effects of second prizes," European Journal of Political Economy, 2005, 21 (2), 467–481.

Toikka, Juuso, "Ironing without control," Journal of Economic Theory, 2011, 146 (6), 2510–2526.

Vázquez, Alexei, "Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations," Physical Review E, 2003, 67 (5), 056104.

Zhang, Jun and Junjie Zhou, "Information disclosure in contests: A Bayesian persuasion approach," The Economic Journal, 2016, 126 (597), 2197–2217.

Zhou, Tao, Bing-Hong Wang, Ying-Di Jin, Da-Ren He, Pei-Pei Zhang, Yue He, Bei-Bei Su, Kan Chen, Zhong-Zhi Zhang, and Jian-Guo Liu, "Modelling collaboration networks based on nonlinear preferential attachment," International Journal of Modern Physics C, 2007, 18 (02), 297–314.