1-1-2016

Essays on Reputations and Dynamic Games

Ju Hu
University of Pennsylvania, juhu1@sas.upenn.edu

Essays on Reputations and Dynamic Games

Abstract
This dissertation consists of three essays on reputations and dynamic games. I investigate how incomplete information, Bayesian Learning and strategic behavior interplay in different dynamic settings. In Chapter 1, I study reputation effects between a long-lived seller and different short-lived buyers where buyers enter the market at random times and only observe a coarse public signal about past transactions. The signal measures the difference between the number of good and bad outcomes in a biased way: a good outcome is more likely to increase the signal than a bad outcome to decrease it. The seller has a short-run incentive to shirk, but makes high profits if it were possible to commit to high effort. I show if there is a small but positive chance that the seller is a commitment type who always exerts high effort and if information bias is large, equilibrium behavior of the seller exhibits cyclic reputation building and milking. The seller exerts high effort at some values of the signals in order to increase the chance of reaching a higher signal and build reputation. Once the seller builds up his reputation through reaching a high enough signal, he exploits it by shirking. In chapter 2, I study the reputation effect in which a long-lived player faces a sequence of uninformed short-lived players and the uninformed players receive informative but noisy exogenous signals about the type of the long-lived player. I provide an explicit lower bound on all Nash equilibrium payoffs of the long-lived player. The lower bound shows when the exogenous signals are sufficiently noisy and the long-lived player is patient, he can be assured of a payoff strictly higher than his minmax payoff. In Chapter 3 I study optimal dynamic monopoly pricing when a monopolist sells a product with unknown quality to a sequence of short-lived buyers who have private information about the quality. Because past prices and buyers' purchase behavior convey information about private signals, they jointly determine the public belief about the quality of the monopolist's product. The monopolist is essentially doing experimentation in the market because every price charged generates not only current period profit but also additional information about the quality. I focus on information structures with a continuum of signals. Under a mild regularity condition on information structures, I show that in equilibrium, the optimal price is an increasing function of the public beliefs. In addition, I fully characterize information cascade sets in terms of information structure. I find that the standard characterization in terms of boundedness of information structure in the social learning literature no longer holds in the presence of a monopoly. In fact, whether herding occurs or not depends more on the values of the conditional densities of the signals at the lowest signal.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Economics

First Advisor
George J. Mailath

Keywords
Dynamic games, Learning, Reputations

Subject Categories
Economics | Economic Theory

ESSAYS ON REPUTATIONS AND DYNAMIC GAMES

Ju Hu

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

____

George J. Mailath
Professor of Economics

Graduate Group Chairperson

____

Jesus Fernandez-Villaverde,
Professor of Economics

Dissertation Committee:

George J. Mailath (Chair), Professor of Economics
Rakesh V. Vohra, Professor of Economics
Yuichi Yamamoto, Professor of Economics

ACKNOWLEDGEMENTS

I am most grateful to my advisor, Professor George Mailath, for his guidance and support throughout the time of my dissertation research. During the past years, I have benefited a lot from weekly meetings with Professor Mailath. As an academic advisor, his invaluable guidance, inspiring comments, sharp insights and challeging questions have shaped the way I think and pushed me to pursue higher goals. As a professor, he taught me "what this profession is for." As a life-time mentor, he sets an example of perfectionist. The intellectual journey with Professor Mailath is immensely rewarding and what I have learned from him is the greatest wealth for my whole life.

I also thank the other members of my dissertation commitee. I am deeply indebted to Professor Rakesh Vohra. His great intuitions and constructive criticisms have deepened my understanding about my topics and largely improved the dissertation. Professor Yuichi Yamamoto introduced the frontier of dynamic games to me when I was in my second year. It is his encouragement that led me to this field. I am deeply grateful for his detailed comments, advice and help at every stage of my dissertation.

A slightly shorter version of Chapter 2 appears in Journal of Economic Theory 2014 Volume 153, 64-73. I thank Marciano Siniscalchi (co-editor) and an anonymous referee for helpful comments.

My doctoral research has also benefited from discussions with Aislinn Bohren, Mustafa Dogan, Hanming Fang, Fei Li, Andrew Postlewaite, Cheng Wang and Yanhao Wei. Special thanks go to Jianrong Tian for our countless all-night discussions.

Everything I have achieved would have been impossible without my family. I would like to thank my parents, for their unconditional love. I dedicate this dissertation to my wife, Qing Zheng, for trusting me, supporting me and cheering me up.

ABSTRACT

Ju Hu

George J. Mailath

This dissertation consists of three essays on reputations and dynamic games. I investigate how incomplete information, Bayesian Learning and strategic behavior interplay in different dynamic settings. In Chapter 1, I study reputation effects between a long-lived seller and different short-lived buyers where buyers enter the market at random times and only observe a coarse public signal about past transactions. The signal measures the difference between the number of good and bad outcomes in a biased way: a good outcome is more likely to increase the signal than a bad outcome to decrease it. The seller has a short-run incentive to shirk, but makes high profits if it were possible to commit to high effort. I show if there is a small but positive chance that the seller is a commitment type who always exerts high effort and if information bias is large, equilibrium behavior of the seller exhibits cyclic reputation building and milking. The seller exerts high effort at some values of the signals in order to increase the chance of reaching a higher signal and build reputation. Once the seller builds up his reputation through reaching a high enough signal, he exploits it by shirking. In chapter 2, I study the reputation effect in which a long-lived player faces a sequence of uninformed short-lived players and the uninformed players receive informative but noisy exogenous signals about the type of the long-lived player. I provide an explicit lower bound on all Nash equilibrium payoffs of the long-lived player. The lower bound shows when the exogenous signals are sufficiently noisy and the long-lived player is patient, he can be assured of a payoff strictly higher than his minmax payoff. In Chapter 3 I study optimal dynamic monopoly pricing when a monopolist sells a product with unknown quality to a sequence of short-lived buyers who have private information about the quality. Because past prices and buyers purchase behavior convey information about private signals, they jointly determine the public belief about the quality of the monopolists product. The monopolist is essentially doing experimentation in the market because every price charged generates not only current period profit but also additional information about the quality. I focus on information structures with a continuum of signals. Under a mild regularity condition on information structures, I show that in equilibrium, the optimal price is an increasing function of the public beliefs. In addition, I fully characterize information cascade sets in terms of information structure. I find that the standard characterization in terms of boundedness of information structure in the social learning literature no longer holds in the presence of a monopoly. In fact, whether herding occurs or not depends more on the values of the conditional densities of the signals at the lowest signal.

Contents

1 Biased learning and permanent reputation **1**

1.1 Introduction 1

1.1.1 Related literature 6

1.2 Model 9

1.2.1 Stage game 9

1.2.2 Random entry 10

1.2.3 Biased information 11

1.2.4 Incomplete information and types 13

1.3 Equilibrium 13

1.3.1 Strategies and equilibria 13

1.3.2 Complete information benchmark 15

1.3.3 Reputations 17

1.4 Formal random entry and learning model 21

1.4.1 Random entry model 22

1.4.2 Posterior beliefs, ex post symmetry and stationarity 25

1.4.3 Reputation game with random entry model 27

1.5 Conclusion 29

2 Reputation in the presence of noisy exogenous learning **31**

2.1 Introduction 31

2.2 Model 34

2.2.1 Reputation game with exogenous learning 34

2.2.2 Relative entropy 37

2.3 Main result 39

2.4 An example 41

3 Social learning and market experimentation **45**

3.1 Introduction 45

3.2 Model 50

3.2.1 Preliminary result 52

3.3 Price monotonicity 55

3.4 Characterization of Cascade Sets 58

3.4.1 $h > v > l$ 60

3.4.2 $h > l > v$ 62

3.5 Further discussions 63

Appendix A Proofs for Chapter 1 **65**

A.1 Mathematical preliminaries 65

A.2 Distributions of signals under stationary strategy 66

A.3 Asymptotic behavior of posterior beliefs. 79

A.4 Proof of Lemma 1.3.1 87

A.5 Proof of Theorem 1.3.1 89

A.6 Proof of Theorem 1.3.2 93

A.7 Proof of Theorem 1.3.3 97

A.8 Proofs for Section 1.4 101

Appendix B Proofs for Chapter 2 **107**

B.1 Proof of Proposition 2.3.1 107

Appendix C Proofs for Chapter 3 **113**

C.1 Proof of Lemma 3.2.1 and 3.2.2 113

C.2 Proofs for Section 3.3 115

C.3 Proof of Proposition 3.4.1 117

Bibliography **124**

Chapter 1

Biased learning and permanent reputation

1.1 Introduction

This chapter studies reputation effects between a long-lived seller and different short-lived buyers where buyers enter the market at random times and only observe a coarse and biased signal upon entry. When a potential client walks into a lawyer's office, the client usually does not really have detailed records of this lawyer's career. The client may know in how many cases this lawyer has succeeded in the past, through advertisement or self-introduction, but typically he can rarely know how many cases this lawyer has dealt with in his career. To know how successful this lawyer is, the client has to speculate about how many cases this lawyer has dealt with but failed.

Much of the standard reputation literature assumes that short-lived players observe detailed history about past play. This assumption plays a critical role in the analysis of reputation effects. However, in many markets where a long-lived player repeatedly interacts with different short-lived players, detailed information about past transactions is typically not available to the short-lived players. Even if the short-lived players may have some partial or aggregate information about the past, in many circumstances, this information is more likely to reveal past good outcomes than bad outcomes, as is the case in the lawyer's example. Other examples of this kind of biased information include "reporting bias" in online reputation systems that relies on voluntary reporting (see a discussion below) and manipulated search engine results as a result of so-called "online reputation management," which suppresses negative search contents by promoting positive ones.

Despite its salience in many situations, the implications of biased partial information in reputation effects have been left unexplored. In fact, this information environment has two conflicting effects on reputation building. On one hand, the uninformed short-lived players only have very coarse information about the past and can draw only imprecise inference about the characteristics of the long-lived player. This fact discourages the long-lived player's reputation building because his ability to manipulate the short-lived players' beliefs through past outcomes is limited. On the other hand, because information about past bad outcomes is not as likely as good outcomes to be revealed, the long-lived player is less afraid of producing bad outcomes and hence has a larger incentive to milk his reputation, which in turn encourages him to build reputation first. Hence it is not *a priori* clear whether the long-lived player is willing to build a reputation or not if the short-lived players only have partial biased information.

This chapter addresses this question. We examine repeated interactions between a long-lived seller and different short-lived buyers. The seller faces a moral hazard problem: he has a short-term incentive to exert low effort, but makes higher profits if it were possible to commit to high effort. Buyers are willing to choose the customized product if they are sufficiently confident that the seller exerts high effort; otherwise they would like the standardized product. Buyers are unsure of the characteristics of the seller. In particular, there are two types of the seller. One is a commitment type who always exerts high effort. The other is a normal type who behaves strategically to maximize his long-run payoff.

In every period, a new buyer enters the market. However, departing from standard assumptions in repeated games, we assume that the entering buyers do not know the number of transactions before them and only observe a coarse public signal upon entry. The public signal measures the difference between the number of good and bad outcomes in a biased way: a good outcome is more likely to increase the signal than a bad outcome to decrease it. A prominent example of this information setting is sellers' feedback scores on eBay. When buyers complete a transaction on eBay, they can leave either a positive, negative or neutral feedback, or leave no feedback at all. A seller's feedback score is measured as the difference between the number of positive feedbacks and the number of negative feedbacks received by the seller. The bias contained in the public signals reflects buyers' "reporting bias", a situation where buyers exhibit different propensities to report different outcomes to online feedback system (Dellarocas and Wood (2008)). In fact, a growing empirical literature on eBay's reputation systems has found that satisfied buyers are more likely to post a positive feedback than dissatisfied buyers to leave a negative feedback (see for example Bolton, Greiner, and Ockenfels (2013), Dellarocas and Wood (2008), and Nosko and Tadelis (2015)).

In Theorem 1.3.1, we show that when the bias in the public signals is large, repeated play of the stage Nash equilibrium would be the unique stationary public equilibrium[1] if there were no uncertainty about the characteristics of the seller. In this equilibrium, the long-lived seller always exerts low effort and the buyers always choose the standardized product. This result emphasizes that biased information about the past cannot mitigate the seller's moral hazard problem if there is no incomplete

[1] We focus on symmetric behavior of the buyers. For a detailed discussion of symmetry among buyers, see Section 1.4.

information and hence no room for the seller to build a reputation.[2]

In Theorem 1.3.2, we show that if there is a small but positive chance that the seller is the commitment type who always exerts high effort and if the bias is large, then in every equilibrium, the normal seller must exert high effort at some public signals. In equilibrium, the normal seller is willing to exert high effort because he wants to build up his reputation to induce the choice of the customized product. Hence, even if only very coarse information is revealed to the buyers, the seller still has the ability and incentives to build his reputation, as long as the information is sufficiently biased. By imitating the commitment type and exerting high effort, the normal seller increases his chance of getting a higher signal and thus a higher reputation.

In Theorem 1.3.3 we show that the normal seller builds up his reputation only to milk it. In every equilibrium, once the seller builds up reputation through reaching a high enough signal, the buyers will be convinced that they are facing the commitment type with large probability and hence choose the customized product. The normal seller then exploits by exerting low effort. As a result, for a range of parameter values, reputation is cyclic in equilibrium. When the signal is small, the seller exerts effort to imitate the commitment type and build reputation. Once the signal is high, the seller stops exerting effort and the signal will on average gradually decrease. When the signal becomes small again, the seller then exerts high effort to rebuild reputation. This cyclic feature of reputation building and exploitation is different from the temporary reputation results in Cripps, Mailath, and Samuelson (2004,

[2]This result would be trivial if we focused on Markov perfect equilibrium in which both the seller's and the buyers' strategies only depend on buyers' posterior beliefs about the types of the seller since posterior beliefs are never updated without incomplete information. However, in Theorem 1.3.1, we allow the strategies to depend on public signals as is the case with incomplete information and hence obtain a stronger result.

2007).

Section 1.4 is devoted to resolve a conceptual difficulty in modelling symmetric buyers in a formal random entry model. In the reputation game, we want to focus on symmetric behavior of the buyers for tractability. Because buyers' behavior depends on their beliefs about the types of the seller, symmetric behavior of the buyers requires that all buyers hold the same posterior beliefs about the types of the seller when they enter and observe the same current record. This is what we call *ex post* symmetry. One modelling approach to guarantee ex post symmetry is to assume *ex ante* symmetry—-that all buyers are uniformed distributed across all periods so that they all have the same prior belief about the number of transactions before them which leads to the same posterior belief about the types of the seller. This approach is commonly used in literature in similar contexts,[3] and we take this approach in Sections 1.2 and 1.3 to facilitate the understanding of the reputation game. However, this approach is not consistent with any formal random entry model because there is no uniform distribution over the set of countably many buyers and any formal random entry model must induce nontrivial heterogeneity in buyers' priors. To resolve this difficulty, in Section 1.4 we take a different approach and directly focus on ex post symmetry. The key results are that (a) there are indeed random entry models that induce ex post symmetry and (b) ex ante symmetry and ex post symmetry are equivalent in the reputation game as long as symmetric behavior of the buyers is concerned. This is a technical contribution of this .

[3]For example, in anonymous local games, Kets (2011) assumes that there are countably many potential players and every set of n candidate players with consecutive labels is equally likely to be selected to participate in the local game; in an observational learning context with biased information, Herrera and Hörner (2013) assumes a countable set of agents do not know their indices and believe that they are equally likely to be anywhere in the sequence.

1.1.1 Related literature

The adverse selection approach to reputations was first introduced in the context of finitely repeated prisoners' dilemma and chain-store game in Kreps, Milgrom, Roberts, and Wilson (1982), Kreps and Wilson (1982) and Milgrom and Roberts (1982). In infinitely repeated games, Fudenberg and Levine (1989, 1992) showed that the incomplete information about one player's characteristics imposes lower bounds on ex ante equilibrium payoffs to the informed long-run player. [4]

The long-run reputation effects of these games are explored in Cripps, Mailath, and Samuelson (2004, 2007). They show that in the standard reputation games with imperfect monitoring, reputation effect is only a temporary phenomenon. The uninformed players will eventually learn the true type of the informed player and thus equilibrium play will eventually converge to that of the complete information version of the repeated game. For the results in Fudenberg and Levine (1989, 1992) and Cripps, Mailath, and Samuelson (2004, 2007)to hold, it is important that short-lived players receive detailed information about the play in each of the past period.

Recently, some papers have studied different mechanisms that sustain nondisappearing reputation effects. One strand of these papers consider cases where the types of the long-run player are impermanent, e.g. Mailath and Samuelson (2001) and Phelan (2006). In these models, the type of long-run player is governed by a stochastic process, rather than being determined once and for all at the beginning of the repeated game. As a result, the opponents will never be completely certain of a player's type. Then, permanent reputation arises if the long-lived player has incentives to continually build reputation in order to demonstrate that his type has not changed. Ekmekci, Gossner, and Wilson (2012a) further generalizes this idea and considers general reputation models with replacement of the types. They show

[4] See also Gossner (2011a) for a simpler and unified derivation of these results using relative entropy.

that, under general conditions on the convergence rates of the discount factor to one and of the rate of replacement to zero, in any Nash equilibrium the long-run player can guarantee himself the Stacklberg payoff after any history reached with positive probability.

Another strand of the literature shares with the current model the feature that the observations of short-lived players are restricted, but we focus on different aspects of what kind information is available. Liu (2011) and Liu and Skrzypacz (2014) considers the case where the short-lived players can recall only a finite number of past behavior of the long-lived player.[5] In these two models, while the short-lived players' observations are restricted to truncated histories, the long-lived player's behavior in recent periods is perfectly monitored. Hence, the incentive to build a reputation arises from the long-lived player's desire to "clean up" history. In equilibria of these two models, the short-lived players actually know that they are facing the normal type most of the time. Their posterior beliefs about the commitment type suddenly jump up once the long-lived player cleans up recent history, but drop down to zero once reputation milking occurs and remain so until next time when recent history is cleaned up again. In contrast, because short-lived players in the current model only observe a coarse signal, they can never distinguish the two types of the long-lived player. Posterior beliefs in equilibria gradually change over time as a result of reputation building and milking. Moreover, our analysis on how to model symmetric buyers in Section 1.4 also provides a foundation for the model and analysis in Liu and Skrzypacz (2014).

Ekmekci (2011) studies long-run reputation effects on the long-lived player's payoff when a rating system observes all past play and publicly announces one of a finite

[5]Liu (2011) assumes short-lived players must pay a positive cost $C(n)$ to observe the long-lived player's past behavior in recent n periods. Because $C(n)$ becomes large when n is large, there exists a finite upper bound N such that acquiring information more than N periods is strictly dominated.

number of ratings to the player. Ekmekci (2011) shows that for any value close to long-lived player's Stackelberg payoff, he can construct a particular rating system and an equilibrium under this rating system such that the long-lived player's payoff after every history is as least as high as the targeted value. His analysis is facilitated by the fact that the desired rating system and equilibrium are jointly constructed. For example, the details of the rating system depends on the payoff structure of the long-lived player as well as the targeted value. Instead, the current chapter is interested in the equilibrium behavior under a fixed information disclosure scheme that transmits biased information to the short-lived players.

This chapter is also related to recent development in observational learning literature. Guarino, Harmgart, and Huck (2011) and Monzón and Rapp (2014) consider observational learning models in which each short-lived agent enters at random times and do not know his "position." The common feature of these two models and the current one is the assumption that agents do not know when the relationship started. However, there are two key differences. First, while all agents in the above two models are short-lived, there is one long-lived agent in the current model and how the long-lived agent strategically respond to the randomly entered agents is the main focus of this chapter. Second, these two models both assumed finite population and considered the asymptotic behavior as the population gets large. To model symmetric agents, they simply assume that all agents are equally likely to enter in every period. However, as mentioned above, this chapter considers random entry of infinite population and we need to develop nontrivial random entry model to model symmetric agents. For a more detailed discussion about this, see Section 1.4.

1.2 Model

1.2.1 Stage game

There is a long-lived seller with discount factor $\delta \in (0,1)$ and a countable set of short-lived potential buyers. When a buyer enters into the market, the seller and the entering buyer play the following product-choice game. The seller chooses between high effort (H) or low effort (L). The buyer chooses between a customized product (c) or a standardized product (s). The outcome of each stage is either a product of high quality (h) or low quality (l). We assume that the probability of a high quality product only depends on the effort level of the seller. If the seller exerts high effort, then with probability $\rho(h|H) = \alpha$ the product is of high quality. If the seller exerts low effort, then with probability $\rho(h|L) = \beta$ the product is of high quality. We assume $1 > \alpha > \beta > 0$, i.e., high effort leads to high quality with larger probability. As usual, the distribution $\rho(\cdot|\cdot)$ can be naturally extended to seller's mixed actions.

The ex ante payoffs of this stage game are denoted by $u_s : \{H, L\} \times \{c, s\} \to \mathbb{R}$ and $u_b : \{H, L\} \times \{c, s\} \to \mathbb{R}$. We make the following assumptions about the payoffs of this stage game. [6]

Assumption 1.2.1. $u_s(L, c) - u_s(H, c) \geq u_s(L, s) - u_s(H, s) > 0$.

This assumption states that there is always a positive cost of exerting high effort and the benefit from low effort when the customized product is chosen is weakly higher than that when the standardized product is chosen.

Assumption 1.2.2. $u_b(H, c) > u_b(H, s)$ and $u_b(L, s) > u_b(L, c)$.

[6] In repeated games with imperfect monitoring, it is standard to assume that buyers' ex ante payoff is induced by his ex post payoff $u_b^* : \{c, s\} \times \{h, l\} \to \mathbb{R}$, i.e. $u_b(a_1, a_2) = \rho(h|a_1)u_b^*(a_2, h) + \rho(l|a_1)u_b^*(a_2, l)$ for $a_1 \in \{H, L\}$ and $a_2 \in \{c, s\}$. However, given the information structure of the current model, while such assumption is natural, it is not necessary.

This assumption states that if the seller exerts high (resp. low) effort, then the buyers strictly prefer the customized product (resp. standardized product). When the seller mixes between H and L, this assumption implies that there exists a cutoff $\kappa \in (0,1)$ such that the buyer strictly prefers customized product (resp. standardized product) if the seller makes high effort with probability larger (resp. smaller) than κ. Assumptions 1 and 2 together imply that (L, s) is the unique Nash equilibrium of this stage game.

Assumption 1.2.3. $u_s(H, c) > u_s(L, s)$.

This assumption states that the seller could make higher profits if it were possible to commit to high effort. Hence exerting high effort H is the pure Stackelberg action of this stage game.

1.2.2 Random entry

Time is discrete $t = 0, 1, 2, \cdots$. The above product choice game is repeatedly played in every period. Departing from standard assumptions in repeated games, we assume that the entering buyers do not know the number of transactions before them. Because we assume that each transaction occurs in every period, the number of transactions is equal to the calendar time. In this and next sections, we assume that all buyers ex ante symmetric in the sense that they all have the same prior belief $\{\mu_t\}_{t=0}^{\infty}$ about when they enter where μ_t is the probability of entering in period t. We also assume $\{\mu_t\}_{t=0}^{\infty}$ follows a geometric distribution with parameter δ, i.e. $\mu_t = (1-\delta)\delta^t$ for $t \geq 0$.

Some remarks are in order. This ex ante symmetry assumption allows us to focus on symmetric behavior of the buyers, which makes the model tractable. However, because there are countably many buyers, if buyers' prior beliefs about when they enter are derived from a formal random entry model, there must be nontrivial heterogeneity in these beliefs (see Lemma 1.4.1). Hence, this ex ante symmetry assumption can only be justified if improper prior distribution is invoked. Using improper prior distribution to model symmetry or anonymity in models with countably many agents in which some kind of randomness among the agents is involved is a common approach in literature (See, for example, Kets (2011), Herrera and Hörner (2013), Liu (2011) and Liu and Skrzypacz (2014)). We follow this approach in Sections 1.2 and 1.3 in order to facilitate understanding of the reputation game. Later, in Section 1.4, we will consider formal random entry models without invoking improper prior distributions and take a different approach to model symmetric behavior of the buyers. There, we will show the contradiction between heterogeneous prior beliefs about when they enter and symmetric behavior can be reconciled in formal random entry models if we focus on ex post symmetry (for details, again see Section 1.4). Most importantly, ex ante symmetry and ex post symmetry will lead to exactly the same results in the reputation game, as long as symmetric equilibria are concerned. Hence the development in Section 1.4 can be viewed as a foundation for ex ante symmetry assumption. Moreover, we use the same parameter $\delta \in (0,1)$ to denote both the discount factor of the seller and the parameter in buyers' beliefs about when they enter for two reasons. One is for expositional ease. All the results will carry over qualitatively if these two parameters are different. The other is that if we interpret seller's discount factor as the continuation probability of the repeated game,[7] then the prior belief $\{\mu_t\}_{t=0}^{\infty}$ exactly corresponds to the distribution of the length of play. This is also the interpretation of δ we adopt in Section 1.4.

1.2.3 Biased information

The entering buyers do not observe detailed history about past play. Instead, upon entry, each entering buyer only observes a public signal. This public signal measures

[7] For example, this interpretation is adopted in Jehiel and Samuelson (2012). For a detailed discussion of this interpretation, see Section 4.2 in Mailath and Samuelson (2006a).

the difference between the number of high and low quality products provided in the past in a biased way: high quality product and low quality product are not equally likely to change the public signal.

Formally, let $R \equiv \mathbb{Z}_+$ be the signal space. In period $t = 0$, before the game starts, the initial signal is $R_0 = 0$. As the game evolves, the signal evolves according to a Markovian transition rule $T : R \times \{h, l\} \to \Delta\{R\}$ such that for all $r \in R$

$$\begin{aligned} T(r,h)[r+1] &= 1-\varepsilon, \\ T(r,h)[r] &= \varepsilon, \end{aligned}$$

and

$$\begin{aligned} T(r,l)[r] &= \varepsilon, \\ T(r,l)[\max\{r-1,0\}] &= 1-\varepsilon, \end{aligned}$$

where $\varepsilon \in [0, 1]$. That is, if the current signal is r and an additional high quality product is provided, then the signal increases by 1 with probability $1-\varepsilon$ and remains unchanged with probability ε. If, instead, an additional low quality product is provided, then the signal decreases by 1 with probability ε and stays unchanged with probability $1 - \varepsilon$.

As mentioned in the introduction, here the signal structure models some online rating systems such as sellers' feedback scores on eBay, which measure the difference between the number of positive and negative reviews from the buyers. The parameter ε captures buyers' behavioral bias in leaving reviews in online rating systems. As found in Bolton, Greiner, and Ockenfels (2013), Dellarocas and Wood (2008), and Nosko and Tadelis (2015), buyers with different purchasing experience are not equally likely to leave a review. In particular, satisfied buyers are more likely to leave a positive review than dissatisfied buyers to leave a negative review. This corresponds to the case $\varepsilon < 1/2$. The smaller ε is, the larger the bias is. Notice that in the extreme case, $\varepsilon = 0$, the public signal exactly measures the number of high quality products provided in the past. In this case, low quality products have no effect at all on the evolution of the signals and information contained in the signals is completely biased. In what follows, we will mainly focus on small but positive ε, which is a noisy version of the extreme case.

1.2.4 Incomplete information and types

Following the standard reputation literature, we assume there is incomplete information about the characteristics of the seller. In particular, there are two types of the seller, $\widehat{\xi}$ and $\widetilde{\xi}$. The type $\widehat{\xi}$ seller is a commitment type who always exerts high effort. The type $\widetilde{\xi}$ seller is a normal seller who behaves strategically and chooses between high effort or low effort to maximize his expected long-run profits. The type of seller is the seller's private information. Assume that all buyers have common prior belief about the type of the seller. Let $b_0 \in (0, 1)$ denote the prior probability that the seller is the commitment type. Denote by $\Gamma(b_0, \alpha, \beta, \delta, \varepsilon)$ the reputation game.

1.3 Equilibrium

1.3.1 Strategies and equilibria

Because by assumption all buyers are ex ante symmetric, we consider their symmetric strategies. Denote $\sigma_2 : R \to [0, 1]$ as buyers' symmetric strategy that specifies a probability of choosing the customized product for each signal $r \in R$. Given a symmetric strategy of the buyers, the normal seller always has a stationary best response that only depends on the public signals. Denote $\sigma_1 : R \to [0, 1]$ as the normal seller's stationary public strategy that specifies a probability of exerting high effort for each $r \in R$.

Every stationary public strategy σ_1 of the normal seller, together with the distributions of qualities and transition rule of the signals, induces probability measure $P_{\tilde{\xi}}^{\sigma_1}$ over R^{∞}. This measure $P_{\tilde{\xi}}^{\sigma_1}$ defines a Markov chain $\{R_t\}_{t\geq 0}$ over the state space R. From the buyers' point of view, the commitment type seller's strategy also induces a measure $P_{\hat{\xi}}$ and hence a Markov chain over the signals. After observing a signal r, the entering buyer updates his belief about the commitment type given his prior belief about the types and about when he enters according to Bayes' rule:

$$\nu(r) = \frac{b_0 \sum_{t=0}^{\infty} \delta^t P_{\hat{\xi}}\big(R_t = r\big)}{b_0 \sum_{t=0}^{\infty} \delta^t P_{\hat{\xi}}\big(R_t = r\big) + (1-b_0) \sum_{t=0}^{\infty} \delta^t P_{\tilde{\xi}}^{\sigma_1}\big(R_t = r\big)}. \tag{1.1}$$

The formula (1.1) can be understood as follows. If a buyer enters in period t, then the probability of observing signal r under the commitment type (resp. normal type) is $P_{\hat{\xi}}\big(R_t = r\big)$ (resp. $P_{\tilde{\xi}}^{\sigma_1}\big(R_t = r\big)$). If the buyer knew that he enters in period t, his posterior belief about the commitment type would be $b_0 P_{\hat{\xi}}(R_t = r)/\big(b_0 P_{\hat{\xi}}\big(R_t = r\big) + (1-b_0) P_{\tilde{\xi}}^{\sigma_1}(R_t = r)\big)$. However, the entering buyer does not know when he enters. In fact, he believes that he enters in period t with probability (proportional to) δ^{t-1}. Hence the probability of observing r when he enters under the commitment type and the normal type respectively are $\sum_{t=0}^{\infty} \delta^t P_{\hat{\xi}}\big(R_t = r\big)$ and $\sum_{t=0}^{\infty} \delta^t P_{\tilde{\xi}}^{\sigma_1}\big(R_t = r\big)$. These two probabilities together with the prior belief about the types then yield (1.1) as buyer's posterior belief after observing signal r.

Every stationary public strategy profile (σ_1, σ_2) induces a value function $V : R \to \mathbb{R}$ for the normal seller. It can be written recursively as

$$\begin{aligned} V(r) \quad = \quad & (1-\delta) u_s(\sigma_1(r), \sigma_2(r)) + \delta\Big[(1-\varepsilon)\rho(h|\sigma_1(r))V(r+1) \\ & + \big[\varepsilon\rho(h|\sigma_1(r)) + (1-\varepsilon)\rho(l|\sigma_1(r))\big]V(r) \\ & + (1-\varepsilon)\rho(l|\sigma_1(r))V(\max\{r-1, 0\})\Big]. \end{aligned} \tag{1.2}$$

For every r, $V(r)$ is the normal seller's expected long-run payoff at signal r under the strategy profile (σ_1, σ_2).

A stationary public equilibrium of this game requires that strategies be mutual best responses given beliefs and that the beliefs be consistent with the strategies.

Definition 1.3.1. A stationary public equilibrium of the game $\Gamma(b_0, \alpha, \beta, \delta, \varepsilon)$ is the triple $(\sigma_1^*, \sigma_2^*, \nu^*)$ such that

(a) for all $r \geq 0$,

$$\begin{aligned} \sigma_1(r) \quad \in \quad \underset{\varsigma_1 \in [0,1]}{\arg\max} \; & (1-\delta)u_s(\varsigma_1, \sigma_2^*(r)) + \delta\Big[(1-\varepsilon)\rho(h|\varsigma_1)V^*(r+1) \\ & + \big[\varepsilon\rho(h|\varsigma_1) + (1-\varepsilon)\rho(l|\varsigma_1)\big]V^*(r) \\ & + (1-\varepsilon)\rho(l|\varsigma_1)V^*(\max\{r-1, 0\})\Big], \end{aligned}$$

where V^* is defined in (1.2);

(b) for all $r \geq 0$, $\sigma_2^*(r)$ is a best response to $\nu^*(r) + \big(1 - \nu^*(r)\big)\sigma_1^*(r)$;

(c) $\nu^*(r)$ is formed via Bayes' rule according to (1.1).

The following lemma states that a stationary public equilibrium always exists. It is a standard application of Glicksberg (1952)'s generalization of Kakutani's fixed-point theorem.

Lemma 1.3.1. *A stationary public equilibrium exists.*

Proof. See Appendix A.4. □

1.3.2 Complete information benchmark

In this section, we analyze the complete information version of the above model. The result here serves as a benchmark to see the effects of biased information and random entry on the equilibrium behavior when there is no room for reputation building.

We apply Definition 1.3.1 to the case $b_0 = 0$. It is clear that condition (c) in Definition 1.3.1 becomes $v^*(r) = 0$ for all r as required by (1.1), meaning that the buyers' posterior beliefs about the seller's type are never updated. Moreover, it is straightforward to see that if we focus on equilibrium in which both the seller and buyers' strategies only depend on posterior beliefs (equivalently, strategies that take the same action at all public signals), then repeated play of the stage Nash is the unique equilibrium and the seller never exerts high effort. The following theorem states that this result continues to hold even if we allow the strategies to depend on the public signals, as long as the bias in the signals is large.

Theorem 1.3.1. *Assume* $b_0 = 0$. *For any* $\alpha > \beta$, *there exists* $\overline{\varepsilon}$ *such that for all* $\varepsilon < \overline{\varepsilon}$ *and* $\delta \in (0, 1)$, *there is a unique stationary public equilibrium in* $\Gamma(0, \alpha, \beta, \delta, \varepsilon)$. *In this equilibrium, the seller always exerts low effort and the entering buyers choose the standardized product.*

Proof. See Appendix A.5. □

This theorem states that high effort cannot be supported as equilibrium outcome if there is no incomplete information about the seller's characteristics and if the bias of the signals is large, even if the continuation probability is sufficiently large. In this case, the unique stationary public equilibrium is the repeated play of the stage Nash equilibrium.

The intuition of this result is best understood by considering the extreme case $\varepsilon = 0$. In this case, the public signal measures the number of high quality products provided in the past. Every additional high quality product increases the signal while low quality product does not change the public signal at all. In any equilibrium, if the seller were willing to exert high effort with positive probability at a public signal r, he must be rewarded at signal $r + 1$ since high effort is strictly dominated in the stage game. This leads to a discrete jump between the seller's values at signal $r + 1$ and r. This fact immediately rules out equilibria in which the seller only exerts high effort at finitely many signals. This is because if the seller were willing to exert high effort at

r and low effort at all signals higher than r, then the seller's value at $r+1$ is only his minmax payoff at $r+1$ because the buyers would choose the standardized product at all signals higher than or equal to $r+1$ because of *complete information*, which in turn implies the seller would have no incentives to exert high effort at r. If, instead, the seller were willing to exert high effort at infinitely many signals, then there must be infinitely many discrete jumps in the seller's value function. This intuitively requires that the seller get unboundedly high continuation payoff at high enough signals. But this is impossible because the seller's value function must be bounded by the stage game payoffs.

1.3.3 Reputations

Having discussed the inability of biased information to support high effort in stationary public equilibrium under complete information, we now turn to the incomplete information case. When $b_0 > 0$, then ex ante all buyers believe there is a positive chance that the seller is a commitment type who always exerts high effort. This fact gives the normal seller an opportunity to build a reputation by imitating this commitment type and making high effort.

The following theorem states that the normal seller is indeed willing to take this opportunity when the bias and continuation probability are large.

Theorem 1.3.2. *Assume* $0 < b_0 < \kappa$ *and* $\alpha > 2\beta$*. Then there exist* $\underline{\delta} \in (0,1)$ *and* $\overline{\varepsilon} \in (0,1)$ *such that for all* $\delta > \underline{\delta}$ *and* $\varepsilon < \overline{\varepsilon}$*, in every stationary public equilibrium* $(\sigma_1^*, \sigma_2^*, \nu^*)$ *of the game* $\Gamma(b_0, \alpha, \beta, \delta, \varepsilon)$*, there exists* $r \in R$ *such that* $\sigma_1^*(r) > 0$*. If in addition* $\alpha > 3\beta$*, then* $\overline{\varepsilon}$ *can be chosen to be bigger than* β*.*

Proof. See Appendix A.6. □

The theorem states that if there is a small but positive chance of the commitment type who always exerts high effort, and the distributions of the qualities under high effort and low effort are different enough, then always exerting low effort is no longer an equilibrium if the continuation probability is sufficiently high and the bias is sufficiently large. In this case, in every stationary public equilibrium, the normal seller must exert high effort with positive probability at some values of the signals. Notice that the parameter ranges in Theorem 1.3.1 and 1.3.2 overlap. In the common parameter range, the sharp contrast of these two theorems highlight the role of incomplete information in supporting high effort as equilibrium behavior.

From Assumptions 1 and 3, we know that the normal seller would like the buyers to choose the customized product, but the buyers are willing to do so only if they expect the seller to exert high effort with large probability. In the presence of the commitment type, the buyers expect high effort once they believe they are facing the commitment type with large probability. Therefore, the normal seller has an incentive to build a reputation. By imitating the commitment type and making high effort, the normal seller can increase the chance of getting a higher signal and thus higher reputation, which in turn makes the buyers more convinced that they are facing the commitment type and thus induces the choice of the customized product.

Recall that the buyers are willing to choose the customized product if and only if the seller exerts high effort with probability larger than or equal to κ. If $b_0 > \kappa$, then intuitively the normal seller has no incentive to build reputation because the buyers have already been convinced by their prior belief that they are facing the commitment type with large probability. The condition $b_0 < \kappa$ then rules out this uninteresting case. When $b_0 < \kappa$, in order to convince the buyers that they are facing the commitment type, the normal seller has to first build up his reputation by making high effort. But clearly this condition alone does not guarantee that the seller is willing to build up his reputation. For the seller to have incentive to build reputation, the cost of reputation building must be small compared to the value of reputation. Holding the stage game payoffs fixed, this idea is then characterized by the four parameters α, β, ε, and δ. The condition $\alpha > 3\beta$ and small ε simply guarantees that there is a good chance to get a higher signal and thus a higher reputation under high effort while a bad chance under low effort. The large continuation probability δ makes sure that the seller is willing to build a reputation at a current period cost in order to get a higher reputation value from tomorrow on.

The intuition behind the contrast of Theorems 1.3.1 and 1.3.2 is as follows. In the complete information case, the buyers are willing to choose the customized product only if the normal seller exerts high effort with large probability. But for the seller to exert high effort at some signal r, he must induce the choice of the customized product at some higher signal $r' > r$, which in turn requires the seller to exert high effort at signal r' and even at infinitely many signals. Theorem 1.3.1 then states that this is impossible. In contrast, in the presence of the commitment type who always exert high effort, this logic no longer holds. It is still true that high effort at some signal r requires the choice of the customized product at some signal $r' > r$, but for the buyers to choose the customized product at r', the normal seller does not need to exert high effort again if he has built up his reputation and the buyers are convinced they are facing the commitment type with large probability. This is why high effort can be supported as equilibrium behavior, as Theorem 1.3.2 states, compared to the impossibility result in Theorem 1.3.1. Moreover this reasoning also suggests that in any equilibrium, the seller must milk his reputation. The following theorem states that this is indeed true. The normal seller builds up his reputation only to milk it.

Theorem 1.3.3. *In any equilibrium* $(\sigma_1^*, \sigma_2^*, \nu^*)$ *of the game* $\Gamma(b_0, \alpha, \beta, \delta, \varepsilon)$*, there exists a signal* $\underline{r}$ *such that for all* $r \geq \underline{r}$*, the normal type of the seller exerts low effort* $\sigma_1^*(r) = 0$ *while the buyers believe that they are facing the commitment type with large probability and choose customized product* $\sigma_2^*(r) = 1$*.*

Proof. See Appendix A.7. □

This result states that the normal seller can effectively build a reputation and then start milking his reputation when the signal becomes large. In equilibrium, when entering buyers observe high enough signals, they believe that they are facing the commitment type with large probability and expect that the seller exerts high effort with large probability. Consequently, they choose the customized product. However, on the other hand, the normal seller in fact only exerts low effort.

There are two possible consequences of reputation exploiting. If the bias in the signals is extremely large, then the bad outcomes have little impact on the public signals. This corresponds to the case when $\varepsilon < \beta$. In this case, even if the seller exerts low effort, the probability of the public signal going up is larger than that of going down, i.e. $(1-\varepsilon)\beta > \varepsilon(1-\beta)$, because a large fraction of the bad outcomes can not be reflected in the public signal. As a result, the public signal on average will continue to increase and the incentive to build reputation will disappear in the long run. A more interesting case is when the bias is moderate. This corresponds to $\varepsilon > \beta$. In this case, when the seller exerts low effort, a relative large fraction of bad outcomes will be reflected in the public signal. In fact, the probability of the public signal going down is larger than that of going up, i.e. $(1-\varepsilon)\beta < \varepsilon(1-\beta)$. As a result of low effort, the public signal on average gradually decreases and when it becomes low, the seller will exert high effort to build it up. Hence, in this case, the incentive of building reputation never disappear no matter how long the game has been played and cyclic reputation building and exploiting arises in equilibrium. This observation is summarized in the following corollary.

Corollary 1.3.1. *Assume $\varepsilon > \beta$ and $(\sigma_1^*, \sigma_2^*, \nu^*)$ is a stationary public equilibrium in which the seller exerts high effort at some signals. Then the evolution of the signals $P_{\tilde{\xi}}^{\sigma_1^*}$ is recurrent. Hence seller's incentive to build reputation never disappear in the long-run.*

1.4 Formal random entry and learning model

In the above analysis, we focused on symmetric behavior of the buyers. This is justified by the assumption that all buyers are ex ante symmetric and have the same prior belief about the number of transactions before them. In this section, we re-examine this assumption by considering formal random entry models of the buyers. As mentioned previously, this ex ante symmetry assumption is not consistent with any random entry model, as we will formally see in Lemma 1.4.1. Then, to justify symmetric behavior of the buyers, we take a different approach in this section and focus on random entry models that induce the same posterior beliefs about the types of the seller across all buyers. This is ex post symmetry and, as we will see, it can be endogenized by formal random entry models. This ex post symmetry, together with a stationarity condition, will fully rationalize the reputation game we analyzed in Sections 1.2 and 1.3. The analysis also provides a foundation for the model used in Liu and Skrzypacz (2014). Since the results here have their own interests and have potential applications in various contexts other than the reputation game, we states the model and results in a more general form.

Let Ξ be a finite type space and $b = \big(b(\xi_1), \cdots, b(\xi_{|\Xi|})\big) \in \Delta^{|\Xi|-1}$ be a prior distribution over types with full support. We continue to use R to denote a countable signal space. Let $\mathscr{G}$ be the usual product σ-algebra over $\Xi \times R^\infty$. Every probability measure P over $\big(\Xi \times R^\infty, \mathscr{G}\big)$ with marginal distribution b over Ξ defines a stochastic process as follow. At period $t = -1$, nature selects a type $\xi \in \Xi$ according to the prior distribution b. In every period $t \geq 0$, conditional on the realized type ξ, a signal $r_t \in R$ is generated according to $P_\xi \equiv P\big(\,\cdot\,\big|\{\xi\} \times R^\infty\big)$. Let $\mathcal{P}$ be the set of all such probability measures.

Consider the situation where a countable set of agents, denoted by $\mathbb{Z}$, randomly enter into this stochastic process at period $t = 0, 1, 2, \cdots$, one for each period. Upon entry, each agent does not know when he enters and only observes the current signal, and updates his posterior belief about the types given this observation. In what follows, we first model this situation formally and show there exist models in which the posterior beliefs over types are identical across entering agents. Moreover, this common belief property holds not only for a particular process, but for all possible processes. This justifies the common belief assumption used in the reputation game.

1.4.1 Random entry model

A random entry model specifies (a) the distribution over the total number of agents who enter, or in other word the length of the entry process, and (b) conditional on the length of entry, who enters first, second and so on with what probability.

Formally, for each $n = 1, 2, \cdots$, define

$$\Sigma_n \equiv \left\{ (i_0, i_1, \cdots, i_{n-1}) \in \mathbb{Z}^n \middle| i_s \neq i_t \text{ if } s \neq t \right\},$$

and

$$\Sigma \equiv \bigcup_{n=1}^{\infty} \Sigma_n.$$

Let $\mathscr{E}$ be the power set of Σ.

Definition 1.4.1. A random entry model is a probability measure μ over $(\Sigma, \mathscr{E})$.

The set Σ_n contains the set of all possible orders of entry given the length of entry n. For example, a vector $\theta = (i_1, i_2, \cdots, i_n) \in \Sigma_n$ specifies that the length of entry is n, agent i_1 enters in period 0, agent i_2 enters in period 1, and so on until i_n enters in period $n-1$. Notice the requirement that $i_s \neq i_t$ if $s \neq t$ simply means that all agents are short-lived. Each agent enters at most once. If an agent enters in period s, then he cannot enter in period t again. The entry of agents governed by a random entry model μ can be considered as follows. First, the length of entry n is realized according to distribution $\{\mu(\Sigma_n)\}_{n\geq 1}$. Second, conditional the realized length

n, an order of entry $\theta = (i_1, i_2, \cdots, i_n) \in \Sigma_n$ is realized according to the conditional distribution $\mu(\cdot|\Sigma_n)$. Finally, agents $i_1, i_2, \cdots, i_n$ enter successively in each period $t = 0, \cdots, n-1$ and then the entry process ends.

Every random entry model induces the prior belief for each agent about when he enters. For each agent $i \in \mathbb{Z}$ and every realization of orders of entry $\theta \in \Sigma$, define[8]

$$\tau_i(\omega) \equiv \begin{cases} t & \text{if } i_t = i, \\ +\infty & \text{if } i_s \neq i \ \forall s. \end{cases} \tag{1.3}$$

Then, the mapping $\tau_i : \Sigma \to \mathbb{Z} \cup \{+\infty\}$ is the random time at which agent i enters. For example, $(\tau_i = t)$ represents the event that agent i enters in period t and $(\tau_i = +\infty)$ means that agent i never enters. The distribution of random variable τ_i in under μ defines agent i's prior belief about when he enters. Denote by $\mu_t^i \equiv \mu(\tau_i = t)$ the probability that agent i enters in period t.

A special random entry model is the one used in Guarino, Harmgart, and Huck (2011) and Monzón and Rapp (2014). To study observational learning where agents do not know when they enter, these two papers both assume that there are only finitely many agents and they are equally likely to enter in every period. In terms of the formulation in the current chapter, this random entry model is just a uniform distribution over the set of all permutations of $(0, 1, \cdots, n-1)$, which is a subset of Σ_n, for some n. A direct implication of this random entry model is that all agents are ex ante symmetric. This allows these two papers to focus on symmetric equilibria.

However, to study reputation games, this chapter focuses on random entry models where the number of agents who enter with positive probability is unbounded.[9] We do so for two reasons. First, in the reputation game, we have a long-lived agent which

[8]Notice for each $\omega \in \Sigma$, by construction, there exists at most one t such that $i_t = i$.

[9]From Definition 1.4.1, a random entry model does not necessarily induce entry of infinite population, e.g. the one mentioned in previous paragraph. However, in the next subsection, we will restrict attention to a special class of entry models where the number of agents who enter with positive probability is infinity. See Definition 1.4.3.

is absent from Guarino, Harmgart, and Huck (2011) and Monzón and Rapp (2014). We believe that in reality, it is rarely the case that a seller knows exactly the number of potential buyers. Although there is always physical upper bound on the number of buyers, in a market with large number of potential buyers, a seller's behavior is better captured by models in which he believes that there are infinitely many potential buyers.[10] Second, considering infinitely repeated games, together with a stationarity assumption (see Definition 1.4.3 below) makes the seller's problem stationary and more tractable.[11]

Despite its advantages, random entry models that induce entry of unbounded number of agents have their intrinsic difficulty in modeling symmetry of the entering agents.[12]

Lemma 1.4.1. *Let μ be a random entry model. If infinitely many agents enter with positive probability, i.e.* $\#\{i \in \mathbb{Z} | \mu(\tau_i < \infty) > 0\} = \infty$, *then prior beliefs about entering time must be different across agents.*

This lemma states there is an intrinsic conflict between the assumption of unbounded number of entering agents and ex ante symmetry. We hence will focus on ex post symmetry instead and this idea is formalized in the next subsection.

[10]For more detailed discussions about the plausibility of finitely repeated games vs. infinitely repeated games, see Osborne and Rubinstein (1994) Section 8.2 and Mailath and Samuelson (2006a) Section 4.1.

[11]By definition, every entry model considered in this chapter ends for sure in finite time. An alternative way to allow entry of infinite population is to consider random entry models that last forever. However, such entry models do not have properties we need in the analysis of reputation game. See Lemma A.8.1 in the appendix.

[12]Because a random entry model is mathematically equivalent to a random (partial) matching scheme between the set of agents and calendar times, this lemma is essentially a well-known impossibility result in random matching between infinite number of agents, adapted to the current context. See, for example, Section 3 in Boylan (1992).

1.4.2 Posterior beliefs, ex post symmetry and stationarity

For any random entry model μ and a measure $P \in \mathcal{P}$, the two spaces $\big(\Xi \times R^\infty, \mathscr{G}, P\big)$ and $\big(\Sigma, \mathscr{E}, \mu\big)$ form a product probability space $\big(\Omega, \mathscr{F}, P \otimes \mu\big)$ where $\Omega \equiv \Xi \times \big(R \times \mathbb{Z} \times \{0,1\}\big)^\infty$ and $\mathscr{F}$ is the corresponding σ-algebra. For each agent $i \in \mathbb{Z}$, signal $r \in R$ and type $\xi \in \Xi$, let $\nu_i^{P\otimes\mu}\big(\xi\big|r\big)$ be agent i's posterior belief about type ξ when he enters and observes signal r, i.e.

$$
\begin{aligned}
\nu_i^{P\otimes\mu}(\xi|r) &\equiv P\otimes\mu\big(\{\xi\}\big|R_{\tau_i}=r, \tau_i<\infty\big) \\
&= \frac{b(\xi)P\otimes\mu\Big(R_{\tau_i}=r,\tau_i<+\infty\Big|\{\xi\}\Big)}{\sum_{\xi'} b(\xi')P\otimes\mu\Big(R_{\tau_i}=r,\tau_i<+\infty\Big|\{\xi'\}\Big)} \\
&= \frac{b(\xi)\sum_{t=0}^{\infty}\mu_t^i P_\xi\Big(R_t=r\Big)}{\sum_{\xi'} b(\xi')\sum_{t=0}^{\infty}\mu_t^i P_{\xi'}\Big(R_t=r\Big)}.
\end{aligned}
\tag{1.4}
$$

where $\{\xi\}$ denotes, for notational simplicity, the event $\{\xi\} \times \big(R \times \mathbb{Z} \times \{0,1\}\big)^\infty$, i.e. type ξ.

The following definition formalize the idea that all entering agents are ex post symmetric.

Definition 1.4.2. A random entry model μ satisfies ex post symmetry (EPS) if for every probability measure $P \in \mathcal{P}$, we have

$$
\nu_i^{P\otimes\mu}\big(\xi\big|r\big) = \nu_j^{P\otimes\mu}\big(\xi\big|r\big) \quad \forall i,j \in \mathbb{Z}, \xi \in \Xi, r \in R. \tag{1.5}
$$

Hence, if a random entry model μ satisfies EPS, then for every stochastic processes $P \in \mathcal{P}$, if agents enter according to μ, then they will have the same posterior belief about the types given the same signal. In other words, they are ex post symmetric.

The following lemma provides a characterization of EPS. It reduces EPS into a condition on the ratios of entering probabilities between every pair of agents. A random entry model μ satisfies EPS if and only if these ratios are constant over time.

Lemma 1.4.2. *Let μ be a random entry model. Then μ satisfies EPS if and only if for all i, j, there exists c^{ij} such that $\mu_t^i = c^{ij}\mu_t^j$ for all $t \geq 0$.*

As a simple application of this lemma, consider again the entry model used in Guarino, Harmgart, and Huck (2011) and Monzón and Rapp (2014). We have already known the uniform entry model satisfies EPS because all agents are even ex ante symmetric. In this case, $c^{ij} = 1$. Moreover this lemma implies that the uniform entry model is the unique entry model for finite population that satisfies EPS. To see this, suppose μ is such a model for population size n. Then we know $\mu_t^i = c^{i1}\mu_t^1$ for $0 \leq i, t \leq n-1$. This implies that $\sum_t \mu_t^i = c^{i1} \sum_t \mu_t^1$ for all $0 \leq i \leq n-1$. Because each agent must enter in some period, we have $\sum_t \mu_t^i = 1$ for all i. This implies that $c^{i1} = 1$ and hence uniform random entry.

As mentioned above, this chapter focuses on random entry models where infinitely many agents enter with positive probability. In particular, we consider models in which the arrival probability of an additional agent is stationary.

Definition 1.4.3. A random entry model μ satisfies stationarity (S) if there exists $\delta \in (0, 1)$ such that for all $n \geq 1$,

$$\mu\Big(\bigcup_{k=n+1}^{\infty} \Sigma_k \Big| \bigcup_{k=n}^{\infty} \Sigma_k\Big) = \delta.$$

In this case, we call δ the continuation probability.

Stationarity simply states that the arrival probability of a new agent is constant over time, independent of the number of agents who have entered. Thus, if there were an outside observer, he would always believe that with probability δ one new agent would enter in next period, regardless of the number of agents who entered in the past.

It is easy to see that stationarity is equivalent to geometric distribution over the length of entry, i.e. $\mu(\Sigma_n) = (1-\delta)\delta^{n-1}$ for all $n \geq 1$. Hence, in any random entry model that satisfies stationarity, there are infinitely many agents who enter with positive probability. According to Lemma 1.4.1, this implies that any random entry model that satisfies S must induce ex ante heterogeneity across agents. However, the following proposition shows that there exist random entry models which satisfy S and render all agents ex post identical.

Proposition 1.4.1. *For every $\delta \in (0,1)$, there exists a random entry model that satisfies EPS and S with continuation probability δ.*

The random entry model that satisfies EPS and S is not unique. For example, if μ is such a model, then so is $\mu \circ \zeta^{-1}$, where $\zeta : \mathbb{Z} \to \mathbb{Z}$ is any permutation of the agents. Despite this multiplicity, the following proposition shows different random entry models that satisfy EPS and S with the same continuation probability are in fact "equivalent": the common posterior beliefs induced by these two models must be the same. This is because EPS and S jointly pin down the form of common posterior beliefs.

Proposition 1.4.2. *Let μ be a random entry model that satisfies EPS and S with continuation probability $\delta \in (0,1)$. Then for any $P \in \mathcal{P}$, the common posterior belief can be written as*

$$\nu^P(\xi|r) = \frac{b(\xi)\sum_{t=1}^{\infty}\delta^t P_\xi\big(R_t = r\big)}{\sum_{\xi'} b(\xi')\sum_{t=1}^{\infty}\delta^t P_{\xi'}\big(R_t = r\big)},$$

for all $\xi \in \Xi$ and $r \in R$.

1.4.3 Reputation game with random entry model

We now apply the results developed in this section to the reputation game studied in Section 1.2. In the reputation game, $\Xi = \{\widehat{\xi}, \widetilde{\xi}\}$, $b = (b_0, 1 - b_0)$ and $R = \mathbb{Z}_+$ as before.

Given any random entry model μ that satisfies EPS and S with continuation probability δ, we incorporate it into the game and modify the original game as follows.

Before the game starts in period $t = -1$, nature selects (a) the type of the seller $\xi \in \{\widehat{\xi}, \widetilde{\xi}\}$ according to prior probability b_0 and (b) the total demand n and the order of entry $\omega = (i_0, \cdots, i_{n-1}) \in \Sigma$ according to the random entry model μ. As before, the realized type of the seller is observed only by the seller. However the realized ω is neither observed by the seller nor by the buyers. Then, the game starts. Buyer i_0 enters in period 0, i_1 enters in period 1 and so on until buyer i_{n-1} enters in period $n-1$, then the game ends. Assume everything else is the same as before, including the stage game, the evolution of the public signals and that entering buyers only observe the current signal and do not know the number of transactions before them. Denote this game as $\Gamma^{\mu}(b_0, \alpha, \beta, \delta, \varepsilon)$.

Because μ satisfies EPS, for any stationary public strategy of the normal seller, all buyers will have the same posterior beliefs about the types and thus the same expectation about the seller's behavior. This allows us to restrict attention to symmetric strategies of the buyers. Because of Proposition 1.4.2, the posterior beliefs of the buyers in the current model have exactly the same form as in (1.1). Because of S, the seller always expects the arrival of next buyer with probability δ. Thus, the definition of stationary public equilibrium in Definition 1.3.1 applies to $\Gamma^{\mu}(b_0, \alpha, \beta, \delta, \varepsilon)$. Therefore, we have

Proposition 1.4.3. *Let μ be a random entry model that satisfies EPS and S with continuation probability δ. The set of stationary public equilibria of the game $\Gamma^{\mu}(b_0, \alpha, \beta, \delta, \varepsilon)$ coincide with that of $\Gamma(b_0, \alpha, \beta, \delta, \varepsilon)$.*

This proposition states that the symmetry assumption of the buyers can indeed be rationalized by formal random entry model if we replace ex ante symmetry by ex post symmetry. As long as symmetric equilibria are, these two notions are in fact equivalent in terms of the set of equilibria. Moreover, this proposition implies that the details of the entry model are immaterial because all entry models that satisfy

EPS and S with the same continuation probability lead to the same set of stationary public equilibria. All these facts suggest that if our focus is symmetric behavior of the buyers, then we should be comfortable with any choice of these models.[13]

1.5 Conclusion

This chapter studies reputation effects when the short-lived players do not know how long the game has been played and there is only coarse and biased information about the past available.

One key new feature of this model is that short-lived players enter the game at random times and upon entry only observe biased signals about past outcomes. This setting departs from the standard assumptions in repeated games where short-lived players observe detailed history about the past and enter in fixed order, and it indeed results in a different set of equilibria even in the complete information case. Without the commitment type, we show that within the class of stage games we study, when bias is large, repeated play of the stage Nash equilibrium is the unique stationary public equilibrium.

In the presence of a Stackelberg type, we show that even if only coarse information is revealed to the short-lived players, the normal type still has incentives to build reputation, as long as the coarse information is sufficiently biased. In every stationary public equilibrium, the normal type player must play the Stackelberg action at some

[13]In the definition of EPS, a random entry model must induce common posterior beliefs for *all* stochastic processes over the signal space. When applied to the reputation game, this requirement seems conceptually too strong because for symmetric behavior of the buyers, it is sufficient that the buyers have identical posterior beliefs for all those processes *that can be generated by some strategies of the seller.* One may worry that these two notions have different implications in the reputation game. However, in the appendix, we show these two notions are in fact equivalent in the current reputation model.

values of the public signals. By doing so, the normal type increases his chance of reaching a higher signal and hence a higher reputation. We also show that the normal type is not willing to always imitate the Stackelberg type. In fact, he builds up his reputation only to milk it. As a result of reputation building and milking, cyclic reputation arises in equilibria.

Chapter 2

Reputation in the presence of noisy exogenous learning [1]

2.1 Introduction

This chapter studies the reputation effect in the long-run interactions in which a long-lived player faces a sequence of uninformed short-lived players and the uninformed players receive informative but noisy exogenous signals about the type of the long-lived player. In the canonical reputation models without exogenous learning (Fudenberg and Levine (1989), Fudenberg and Levine (1992)), the long-lived player can effectively build a reputation by mimicking the behavior of a commitment type because the short-lived player will play a best response to the commitment action in all but a finite number of periods after always seeing the commitment action. The underlying reason is the fact that the short-lived player cannot be surprised too many times: every time the short-lived player expects the commitment action with small probability and yet this action is actually chosen, the posterior belief on this

[1] A slightly shorter version of this chapter appears in Journal of Economic Theory 2014 Volume 153, 64-73.

commitment type jumps up, but at the same time the beliefs can not exceed unity. However this "finite number of surprises" intuition does not carry over to the case with exogenous learning. It is still true that each surprise leads to a discrete jump of the posterior beliefs. But after a surprise during the periods of no surprises, the exogenous learning can drive down the posterior beliefs. After a long history without surprises, the posterior beliefs may return to the original level, resulting in another surprise. Typically, this can happen infinitely many times. Hence in the presence of exogenous learning, there is no guarantee that we have a finite number of surprises.

Wiseman (2009) first presented an infinitely repeated chain store game example with perfect monitoring and exogenous signals taking two possible values. He shows that when the long-lived player is sufficiently patient and there is sufficient noise in the signals, the long-lived player can effectively build a reputation and assure himself of a payoff strictly higher than his minmax payoff.

This chapter extends Wiseman (2009) to more general reputation models with exogenous learning. We provide an *explicit* lower bound on all Nash equilibrium payoffs to the long-lived player. The lower bound is characterized by the commitment action, discount factor, prior belief and how noisy the learning process is. For fixed commitment action and discount factor, the lower bound increases in both prior probability and noise in the exogenous signals. This is intuitive as a higher prior probability on the commitment type and a noisier and slower exogenous learning process both correspond to easier reputation building. When the long-lived player become sufficiently patient, the effect of the prior probability vanishes while that of the exogenous learning remains. This is again intuitive because the prior probability represents the cost of reputation building in the initial periods. When the long-lived player places arbitrarily high weight on future periods, the cost in the initial periods becomes negligible. In contrast, learning has a long run effect. The longer the history, the more the uninformed player can learn about the type of his opponent. Hence the effect of learning remains even if the long-lived player become sufficiently patient.

Not surprisingly, the lower bound we derive is generally lower than that if there is no exogenous learning, reflecting the negative effect of learning on reputation building. In the case that signals are completely uninformative, these two bounds coincide. Nonetheless, when the signals are sufficiently noisy, the lower bound shows that in any Nash equilibrium, the long-lived player is assured of a payoff strictly higher than his minmax value.

To derive the lower bound, we apply the relative entropy approach first introduced by Gossner (2011b) to the study of reputations. Gossner (2011b) uses this approach to the standard reputation game in Fudenberg and Levine (1992) and obtains an explicit lower bound on all equilibrium payoffs. He also shows when the commitment types are sufficiently rich and the long-lived player is arbitrarily patient, the lower bound is exactly the Stackelberg payoff which confirms the result in Fudenberg and Levine (1992). Ekmekci, Gossner, and Wilson (2012b) applied this method to the reputation game in which the type of the long-lived player is governed by an underlying stochastic process. They calculate explicit lower bounds for all equilibrium payoffs at the beginning of the game and all continuation payoffs. In these two papers, relative entropy only serves as a measure of prediction errors. However, in this chapter, in addition to a measure of prediction errors, the concept of relative entropy is also naturally adapted to the learning situation as a measure of noise in the exogenous signals. This again makes relative entropy as a more suitable tool.

The rest of the chapter is organized as follows. In section 2, we describe the reputation model with exogenous learning and introduce relative entropy. Section 3 presents and discusses the main result, which is proved in Section 4.

2.2 Model

2.2.1 Reputation game with exogenous learning

We consider the canonical reputation model (Mailath and Samuelson (2006b), Chapter 15) in which a fixed stage game is infinitely repeated. The stage game is a two-player simultaneous-move finite game of private monitoring. Denote by A_i the finite set of actions for player i in the stage game. Actions in the stage game are imperfectly observed. At the end of each period, player i only observes a private signal z_i drawn from a finite set Z_i. If an action profile $a \in A_1 \times A_2 \equiv A$ is chosen, the signal vector $z \equiv (z_1, z_2) \in Z_1 \times Z_2 \equiv Z$ is realized according to the distribution $\pi(\,\cdot\,|a) \in \Delta(Z)$.[2] The marginal distribution of player i's private signals over Z_i is denoted by $\pi_i(\,\cdot\,|a)$. Both $\pi(\,\cdot\,|a)$ and $\pi_i(\,\cdot\,|a)$ have obvious extensions $\pi(\,\cdot\,|\alpha)$ and $\pi_i(\,\cdot\,|\alpha)$ respectively to mixed action profiles. Player i's ex-post stage game payoff from his action a_i and private signal z_i is given by $u_i^*(a_i, z_i)$. Player i's ex ante stage game payoff from action profile $(a_i, a_{-i}) \in A$ is $u_i(a_i, a_{-i}) = \sum_{z_i} \pi_i(z_i|a_i, a_{-i}) u_i^*(a_i, z_i)$. Notice this setting includes as special cases the perfect monitoring environment (Fudenberg and Levine (1989)) in which $Z_1 = Z_2 = A$ and $\pi(z_1, z_2|a) = 1$ if and only if $z_1 = z_2 = a$, and the public monitoring environment (Fudenberg and Levine (1992)) in which $Z_1 = Z_2$ and $\pi(z_1, z_2|a) > 0$ implies $z_1 = z_2$. Player 1 is a long-lived player with discount factor $\delta \in (0, 1)$ while player 2 is a sequence of short-lived players each of whom only lives for one period. In any period t, the long-lived player 1 observes both his own previous actions and private signals, but the current generation of the short-lived player 2 only observes previous private signals of his predecessors.

There is uncertainty about the type of player 1. Let $\Xi \equiv \{\xi_0\} \cup \hat{\Xi}$ be the set of all possible types of player 1. ξ_0 is the *normal type* of player 1. His payoff in the repeated game is the average discounted sum of stage game payoffs $(1 - \delta) \sum_{t \geq 0} \delta^t u_1(a^t)$. Each

[2]For a finite set X, $\Delta(X)$ denotes the set of all probability distributions over X.

$\xi(\hat{\alpha}_1) \in \hat{\Xi}$ denotes a *simple commitment type* who plays the stage game (mixed) action $\hat{\alpha}_1 \in \Delta(A_1)$ in every period independent of histories. Assume $\hat{\Xi}$ is either finite or countable. The type of player 1 is unknown to player 2. Let $\mu \in \Delta(\Xi)$ be player 2's prior belief about player 1's type, with full support.

At period $t = -1$, nature selects a type $\xi \in \Xi$ of player 1 according to the initial distribution μ. Player 2 does not observe the type of player 1. However, we assume that the uninformed player 2 has access to an exogenous channel which gradually reveals the true type of player 1. More specifically, conditional on the type ξ, a stochastic process $\{\eta_t(\xi)\}_{t\geq 0}$ generates a signal $y^t \in Y$ after every period's play, where Y is a finite set of all possible signals. To distinguish the signals $z \in Z$ generated from each period's play and the signals $y \in Y$ generated by $\{\eta_t(\xi)\}_{t\geq 0}$, we call the former *endogenous signals* and the latter *exogenous signals*. In addition to observing previous endogenous signals, each generation of player 2 also observes all the exogenous signals from earlier periods. We assume that for each type $\xi \in \Xi$, the stochastic process $\{\eta_t(\xi)\}_{t\geq 0}$ is independent and identically distributed across t. Conditional on ξ, the distribution of the exogenous signals in every period is denoted by $\rho(\,\cdot\,|\xi) \in \Delta(Y)$. Notice this assumes that the realization of the exogenous signals are independent of the play, hence it models the exogenous learning of the uninformed player 2.

For expositional convenience, we assume player 1 does not observe the exogenous signals. This assumption is not crucial for our result. The same lower bound will apply if we assume player 1 also observes the exogenous signals.

A private history of player 1 in period t consists of his previous actions and endogenous signals, denoted by $h_1^t \equiv (a_1^0, z_1^0, a_1^1, z_1^1, \ldots, a_1^{t-1}, z_1^{t-1}) \in H_{1t} \equiv (A_1 \times Z_1)^t$, with the usual notation $H_{10} = \{\emptyset\}$. A behavior strategy for player 1 is a map

$$\sigma_1 : \Xi \times \bigcup_{t=0}^{\infty} H_{1t} \to \Delta(A_1),$$

with the restriction that for all $\xi(\hat{\alpha}_1) \in \hat{\Xi}$,

$$\sigma_1(\xi(\hat{\alpha}_1), h_1^t) = \hat{\alpha}_1 \quad \text{for all } h_1^t \in \bigcup_{t=0}^{\infty} H_{1t}.$$

A private history of player 2 in period t contains both previous endogenous and exogenous signals, denoted by $h_2^t \equiv (z_2^0, y_2^0, z_2^1, y_2^1, \ldots, z_2^{t-1}, y_2^{t-1}) \in H_{2t} \equiv (Z_2 \times Y)^t$, with $H_{20} = \{\emptyset\}$. A behavior strategy for player 2 is a map

$$\sigma_2 : \bigcup_{t=0}^{\infty} H_{2t} \to \Delta(A_2).$$

Denote by Σ_i the strategy space of player i.

Any strategy profile $\sigma \equiv (\sigma_1, \sigma_2) \in \Sigma_1 \times \Sigma_2$, together with the prior μ and the signal distributions $\{\pi(\cdot|a)\}_{a \in A}$ and $\{\rho(\cdot|\xi)\}_{\xi \in \Xi}$, induces a probability measure P^σ over the set of states $\Omega \equiv \Xi \times (A_1 \times A_2 \times Z_1 \times Z_2 \times Y)^\infty$. The measure P^σ describes how the uninformed player 2 expects play to evolve. Let $\widetilde{P}^\sigma$ be the conditional probability of P^σ given the event that player 1 is the normal type. The measure $\widetilde{P}^\sigma$ describes how play evolves if player 1 is the normal type. We use $E^\sigma[\cdot]$ (resp., $\widetilde{E}^\sigma[\cdot]$) to denote the expectation with respect to the probability measure P^σ (resp., $\widetilde{P}^\sigma$).

A Nash equilibrium in this reputation game is a pair of mutual best responses.

Definition 2.2.1. A strategy profile $\sigma^* = (\sigma_1^*, \sigma_2^*) \in \Sigma_1 \times \Sigma_2$ is a Nash equilibrium if it satisfies:

(a) for all $\sigma_1 \in \Sigma_1$,

$$\widetilde{E}^{\sigma^*}\Big[(1-\delta)\sum_{t=0}^{\infty}\delta^t u_1(a^t)\Big] \geq \widetilde{E}^{(\sigma_1, \sigma_2^*)}\Big[(1-\delta)\sum_{t=0}^{\infty}\delta^t u_1(a^t)\Big],$$

(b) for all $h_2^t \in \bigcup_{\tau \geq 0} H_{2\tau}$ with positive probability under P^{σ^*},

$$\sigma_2^*(h_2^t) \in \operatorname*{arg\,max}_{\alpha_2 \in \Delta(A_2)} E^{\sigma^*}\Big[u_2\big(\sigma_1^*(h_1^t, \xi), \alpha_2\big)\Big| h_2^t\Big].$$

Condition (a) states that given σ_2^*, the normal type of player 1 maximizes his expected lifetime utility. Condition (b) requires that given σ_1^*, player 2 updates his belief via Bayes' rule along the path of play and plays a myopic best response since he is short lived.

2.2.2 Relative entropy

The *relative entropy* between two probability distributions P and Q over a finite set X is the expected log likelihood ratio

$$d(P\|Q) \equiv E_P \log \frac{P(x)}{Q(x)} = \sum_{x \in X} P(x) \log \frac{P(x)}{Q(x)},$$

with the usual convention that $0 \log \frac{0}{q} = 0$ if $q \geq 0$ and $p \log \frac{p}{0} = \infty$ if $p > 0$. Relative entropy is always nonnegative and it is zero if and only if the two distributions are identical (See Cover and Thomas (2006), Gossner (2011b) and Ekmekci, Gossner, and Wilson (2012b) for more details on relative entropy).

Relative entropy measures the speed of the learning process of the uninformed player 2. For each commitment type $\xi(\hat{\alpha}_1) \in \hat{\Xi}$, let $\lambda_{\xi(\hat{\alpha}_1)}$ be the relative entropy of the exogenous signal distributions when player 1 is the normal type and when he is the commitment type $\xi(\hat{\alpha}_1)$, i.e.

$$\lambda_{\xi(\hat{\alpha}_1)} \equiv d\big(\rho(\,\cdot\,|\xi_0)\big\|\rho(\,\cdot\,|\xi(\hat{\alpha}_1))\big).$$

Relative entropy measures how different the two distributions $\rho(\,\cdot\,|\xi_0)$ and $\rho(\,\cdot\,|\xi(\hat{\alpha}_1))$ are. In terms of learning, $\lambda_{\xi(\hat{\alpha}_1)}$ measures how *fast* player 2 can learn from exogenous signals that player 1 is *not* the commitment type $\xi(\hat{\alpha}_1)$ when player 1 is indeed the normal type. The larger $\lambda_{\xi(\hat{\alpha}_1)}$ is, the faster the learning process is. This is illustrated by the two polar cases. If $\lambda_{\xi(\hat{\alpha}_1)} = 0$, then the distributions of the exogenous signals when player 1 is the normal type and when he is of type $\xi(\hat{\alpha}_1)$ are identical. In this case, from the exogenous signals, player 2 can never distinguish the normal type from the commitment type $\xi(\hat{\alpha}_1)$ when player 1 is the normal type. If $\lambda_{\xi(\hat{\alpha}_1)} = \infty$, there must be some signal $y \in Y$ which will occur when player 1 is the normal type but will not occur when player 1 is the commitment type $\xi(\hat{\alpha}_1)$. Hence in this case, player 2 will learn that player 1 is not the commitment type $\xi(\hat{\alpha}_1)$ for sure in finite time when player 1 is the normal type. For other intermediate values $0 < \lambda_{\xi(\hat{\alpha}_1)} < \infty$, conditional on the normal type, player 2 will eventually know that player 1 is not the commitment type $\xi(\hat{\alpha}_1)$.

The following assumption rules out extremely fast learning. Technically, it requires that the support of $\rho(\,\cdot\,|\xi)$ be contained in the support of $\rho(\,\cdot\,|\xi(\hat{\alpha}_1))$ for every commitment type $\xi(\hat{\alpha}_1)$.

Assumption 2.2.1. $\lambda_{\xi(\hat{\alpha}_1)} < \infty$ for all $\xi(\hat{\alpha}_1) \in \hat{\Xi}$.

Relative entropy measures the error in player 2's one step ahead prediction on the endogenous signals. Gossner (2011b) first introduced the following notion of ε-entropy-confirming best response (see also Ekmekci, Gossner, and Wilson (2012b)):

Definition 2.2.2. The mixed action $\alpha_2 \in \Delta(A_2)$ is an *ε-entropy-confirming best response* to $\alpha_1 \in \Delta(A_1)$ if there exists $\alpha_1' \in \Delta(A_1)$ such that

(a) α_2 is a best response to α_1',

(b) $d\big(\pi_2(\,\cdot\,|\alpha_1, \alpha_2)\big\|\pi_2(\,\cdot\,|\alpha_1', \alpha_2)\big) \leq \varepsilon$.

The set of all ε-entropy confirming best responses to α_1 is denoted by $B_\varepsilon(\alpha_1)$.

The idea of ε-entropy-confirming best response is similar to ε-confirming best response defined in Fudenberg and Levine (1992). If player 2 plays a myopic best response α_2 to his belief that player 1 plays the action α_1', then player 2 believes that his endogenous signals realize according to the distribution $\pi_2(\,\cdot\,|\alpha_1', \alpha_2)$. If the true action taken by player 1 is α_1 instead of α_1', then the true distribution of player 2's endogenous signals is indeed $\pi_2(\,\cdot\,|\alpha_1, \alpha_2)$. Hence player 2's one step ahead prediction error on his endogenous signals is, measured by relative entropy, $d\big(\pi_2(\,\cdot\,|\alpha_1, \alpha_2)\big\|\pi_2(\,\cdot,\,|\alpha_1', \alpha_2)\big)$. The mixed action α_2 is an ϵ-entropy-confirming best response of α_1 if the prediction error is no greater than ϵ.

For any commitment type $\xi(\hat{\alpha}_1) \in \hat{\Xi}$, let

$$\underline{V}_{\xi(\hat{\alpha}_1)}(\varepsilon) \equiv \inf_{\alpha_2 \in B_\varepsilon(\hat{\alpha}_1)} u_1(\hat{\alpha}_1, \alpha_2)$$

be the lowest possible payoff to player 1 if he plays $\hat{\alpha}_1$ while player 2 plays an ε-entropy-confirming best response to $\hat{\alpha}_1$. Let $V_{\xi(\hat{\alpha}_1)}(\,\cdot\,)$ be the pointwise supremum of all convex functions below $\underline{V}_{\xi(\hat{\alpha}_1)}$. Clearly $V_{\xi(\hat{\alpha}_1)}$ is convex and nonincreasing.

2.3 Main result

For any $\delta \in (0, 1)$, let $\underline{U}_1(\delta)$ denote the infimum of all Nash equilibrium payoffs to the normal type of player 1 if the discount factor is δ. Our main result is the following:

Proposition 2.3.1. *Under Assumption 2.2.1, for all* $\delta \in (0, 1)$,

$$\underline{U}_1(\delta) \geq \sup_{\xi(\hat{\alpha}_1) \in \hat{\Xi}} V_{\xi(\hat{\alpha})}\Big(-(1-\delta) \log \mu\big(\xi(\hat{\alpha}_1)\big) + \lambda_{\xi(\hat{\alpha}_1)} \Big).$$

To understand the equilibrium lower bound in Proposition 2.3.1, it suffices to consider the reputation building on each $\xi(\hat{\alpha}_1) \in \hat{\Xi}$ since the overall lower bound is obtained by considering all possible commitment types. Fix a commitment type $\xi(\hat{\alpha}_1) \in \hat{\Xi}$. Proposition 2.3.1 states that in any Nash equilibrium, the normal type of player 1 is assured of a payoff no less than $V_{\xi(\hat{\alpha}_1)}\big(-(1-\delta) \log \mu(\xi(\hat{\alpha}_1)) + \lambda_{\xi(\hat{\alpha}_1)} \big)$. Recall that $V_{\xi(\hat{\alpha}_1)}$ is a nonincreasing function. For fixed δ, this lower bound increases with $\mu(\xi(\hat{\alpha}_1))$ while decreases with $\lambda_{\xi(\hat{\alpha}_1)}$. The intuition is straightforward. A larger prior probability on the commitment type $\xi(\hat{\alpha}_1)$ makes it easier for the normal type of player 1 to build a reputation on this commitment type. In another word, the cost of reputation building in the initial periods is smaller in this case which leads to a higher lower bound. However the learning process goes against reputation building because player 2 eventually learns that player 1 is not the commitment type $\xi(\hat{\alpha}_1)$. It is then intuitive that the speed of learning matters. If the exogenous signals are sufficiently noisy, then $\lambda_{\xi(\hat{\alpha}_1)}$ is small and it is hard for player 2 to distinguish the normal type and the commitment type. This results in a rather slow learning process and hence a high lower bound. If the learning process is completely uninformative, $\lambda_{\xi(\hat{\alpha}_1)} = 0$, then the lower bound is given by $V_{\xi(\hat{\alpha}_1)}\big(-(1-\delta)\log\mu(\xi(\hat{\alpha}_1))\big)$ which is exactly the same lower bound derived in Gossner (2011b) without exogenous learning. In general, when $\lambda_{\xi(\hat{\alpha}_1)} > 0$, the lower bound is lower than that in Gossner (2011b) due to the learning effect.

Another parameter in the lower bound is player 1's discount factor δ. An interesting feature in the lower bound is that δ only appears as a coefficient for the term $\log\mu(\xi(\hat{\alpha}_1))$, not for $\lambda_{\xi(\hat{\alpha}_1)}$. This is because $-\log\mu(\xi(\hat{\alpha}_1))$ captures the cost of reputation building in the initial periods while $\lambda_{\xi(\hat{\alpha}_1)}$ is the learning effect which remains active as the game evolves. As a result, when player 1 becomes arbitrarily patient, $\delta \to 1$, the cost of reputation building in the initial periods becomes negligible since player 1 places higher and higher weight on the payoff obtained in later periods, whereas the learning effect remains unchanged. In this case, the lower bound becomes $V_{\xi(\hat{\alpha}_1)}\big(\lambda_{\xi(\hat{\alpha}_1)}\big)$.[3] Moreover, in the presence of multiple commitment types, which commitment type is the most favorable is now ambiguous. Intuitively, this is because the effectiveness of reputation building does not only depend on the stage game payoff from the commitment type but also on the learning process. Even if player 2 assigns positive probability on the Stackelberg action, committing to the Stackelberg action may not help player 1 effectively build a reputation because the exogenous signals may reveal quickly to player 2 that player 1 is not the Stackelberg commitment type. This is in a sharp contrast with the result in standard models without exogenous learning.

We use the following example which is first considered in Wiseman (2009) to illustrate the lower bound obtained in Proposition 2.3.1.

[3]Since $V_{\xi(\hat{\alpha}_1)}(\varepsilon)$ is convex, it is continuous at every $\varepsilon > 0$.

2.4 An example

There is a long-lived incumbent, player 1, facing a sequence of short-lived entrants, player 2. In every period, the entrant chooses between entering (E) and staying out (S) while the incumbent decides whether to fight (F) or accommodate (A). The stage game payoff is given in Figure 2.1, where $a > 1$ and $b > 0$.

	E	S
F	$-1,\ -1$	$a,\ 0$
A	$0,\ b$	$a,\ 0$

Figure 2.1: Chain store stage game.

The stage game is infinitely repeated with perfect monitoring. There are two types of player 1, the normal type, denoted by ξ_0, and a simple commitment type, denoted by $\xi(F)$ who plays the stage game Stackelberg action F in every period independent of histories. The prior probability of $\xi(F)$ is $\mu(\xi(F))$. The exogenous signals observed by player 2 only take two values: $\overline{y}$ and $\underline{y}$. Assume $\rho(\overline{y}|\xi_0) = \beta$, $\rho(\overline{y}|\xi(F)) = \alpha$ and $\beta > \alpha$. Thus

$$\lambda_{\xi(F)} = \beta \log \frac{\beta}{\alpha} + (1-\beta) \log \frac{1-\beta}{1-\alpha}.$$

Now we apply Proposition 2.3.1 in this setting. Because monitoring is perfect, it is easy to see $B_\varepsilon(F) = \{S\}$ when $\varepsilon < \log \frac{b+1}{b}$. Therefore, we have

$$V_{\xi(F)}(\varepsilon) = \begin{cases} a - \frac{a+1}{\log \frac{b+1}{b}} \varepsilon, & \text{if } \varepsilon < \log \frac{b+1}{b}, \\ -1 & \text{if } \varepsilon \geq \log \frac{b+1}{b}. \end{cases}$$

Proposition 2.3.1 then implies for all $\delta \in (0,1)$

$$\underline{U}_1(\delta) \geq a - \frac{a+1}{\log \frac{b+1}{b}} \Big(-(1-\delta) \log \mu(\xi(F)) + \lambda_{\xi(F)} \Big),$$

and in the limit

$$\liminf_{\delta \to 1} \underline{U}_1(\delta) \geq a - (a+1) \frac{\lambda_{\xi(F)}}{\log \frac{b+1}{b}}. \tag{2.1}$$

Wiseman (2009) considers symmetrically distributed signals, i.e., $\beta = 1-\alpha > 1/2$, and derives a lower bound of $a - (a+1)\frac{\log \frac{\beta}{1-\beta}}{\log \frac{b+1}{b}}$. Because in this symmetric case $\lambda_{\xi(F)} = (2\beta - 1)\log \frac{\beta}{1-\beta}$, this bound is lower than that in (2.1). As signals become less informative, i.e. $\beta \to \frac{1}{2}$, both lower bounds become arbitrarily close to player 1's Stackelberg payoff.

Although it is not surprising that exogenous learning affects reputation building, why does it take this particular form, i.e. the relative entropy of the exogenous signals? As mentioned previously, the "finite number of surprises" argument in Fudenberg and Levine (1989) does not apply because of the downward pressure on posterior beliefs due to exogenous learning. In this particular example, the uninformed entrants may enter infinitely many times even if he is always fought after any entry. Moreover, receiving the signal $\overline{y}$ always decreases the posterior beliefs (recall $\rho(\overline{y}|\xi_0) > \rho(\overline{y}|\xi(F))$) which is the source of the downward pressure. Thus the strength of this downward pressure depends exactly on how frequently the entrants can receive the signal $\overline{y}$ which, together with the size of surprise, in turn determines how long it takes for the posterior beliefs to return after a surprise. In other words, the size of surprise and the relative frequency of exogenous signals together determine the frequency of entries. If it takes a long time for the posterior beliefs to return, then the entrants can not enter too frequently and the incumbent can effectively build a reputation.

To see this, fix any Nash equilibrium σ. For any history h^∞ in which F is always played, let $\{\mu_t\}_{t\geq 0}$ be player 2's posterior belief on the commitment type along this history. Player 2 is willing to enter in period t only if

$$\text{Prob}(F) \equiv \mu_t + (1-\mu_t)\sigma_1(\xi_0, h^t)(F) \leq \frac{b}{b+1}.$$

So, if player 2 enters in period t, we must have

$$\mu_t \leq \frac{b}{b+1} \tag{2.2}$$

and

$$\sigma_1(\xi_0, h^t)(F) \leq \frac{b}{b+1}. \tag{2.3}$$

We examine the odds ratio $\{\mu_t/(1-\mu_t)\}_{t\geq 0}$ along this history. Since the entrant is always fought along this history, the odds ratio evolves as

$$\frac{\mu_{t+1}}{1-\mu_{t+1}} = \big(\frac{\alpha}{\beta}\big)^{\mathbb{1}_{\overline{y}}(y^t)}\big(\frac{1-\alpha}{1-\beta}\big)^{\mathbb{1}_{\underline{y}}(y^t)}\frac{\mu_t}{(1-\mu_t)\sigma_1(\xi_0, h^t)(F)} \quad \forall t \geq 0,$$

where for $y \in \{\overline{y}, \underline{y}\}$, $\mathbb{1}_y$ is the indicator function, $\mathbb{1}_y(y^t) = 1$ if $y^t = y$ and 0 otherwise. Because $\sigma_1(\xi, h^t)(F)$ is always less than or equal to 1, we have

$$\frac{\mu_{t+1}}{1-\mu_{t+1}} \geq \big(\frac{\alpha}{\beta}\big)^{\mathbb{1}_{\overline{y}}(y^t)}\big(\frac{1-\alpha}{1-\beta}\big)^{\mathbb{1}_{\underline{y}}(y^t)}\frac{\mu_t}{1-\mu_t} \tag{2.4}$$

if player 2 stays out in period t. Because inequality (2.3) holds if player 2 enters in period t, we have

$$\frac{\mu_{t+1}}{1-\mu_{t+1}} \geq \big(\frac{\alpha}{\beta}\big)^{\mathbb{1}_{\overline{y}}(y^t)}\big(\frac{1-\alpha}{1-\beta}\big)^{\mathbb{1}_{\underline{y}}(y^t)}\frac{b+1}{b}\frac{\mu_t}{1-\mu_t} \tag{2.5}$$

if he enters in period t. For any $t \geq 1$, let $n_E(t)$, $n_{\overline{y}}(t)$ be the number of entries and the number of signal $\overline{y}$'s respectively in history h^t. Inequalities (2.4), (2.5) and simple induction imply

$$\frac{\mu_t}{1-\mu_t} \geq \big(\frac{b+1}{b}\big)^{n_E(t)}\big(\frac{\alpha}{\beta}\big)^{n_{\overline{y}}(t)}\big(\frac{1-\alpha}{1-\beta}\big)^{t-n_{\overline{y}}(t)}\frac{\mu(\xi(F))}{1-\mu(\xi(F))} \quad \forall t \geq 1. \tag{2.6}$$

Moreover, if player 2 enters in period t, inequality (2.2) implies

$$b \geq \frac{\mu_t}{1-\mu_t}. \tag{2.7}$$

Hence inequalities (2.6) and (2.7) together yield

$$b \geq \big(\frac{b+1}{b}\big)^{n_E(t)}\big(\frac{\alpha}{\beta}\big)^{n_{\overline{y}}(t)}\big(\frac{1-\alpha}{1-\beta}\big)^{t-n_{\overline{y}}(t)}\frac{\mu(\xi(F))}{1-\mu(\xi(F))} \tag{2.8}$$

for all t at which player 2 enters. Let $\{t_k\}_{k\geq 0}$ be the sequence of periods in which entry occurs. By taking log and dividing both sides by t_k, inequality (2.8) implies

$$\begin{aligned}\limsup_{k\to\infty}\frac{n_E(t_k)}{t_k} &\leq \frac{1}{\log\frac{b+1}{b}}\lim_{k\to\infty}\Big[\frac{n_{\overline{y}}(t_k)}{t_k}\log\frac{\beta}{\alpha} + \big(1 - \frac{n_{\overline{y}}(t_k)}{t_k}\big)\log\frac{1-\beta}{1-\alpha}\Big] \\ &= \frac{\lambda_{\xi(F)}}{\log\frac{b+1}{b}},\end{aligned}$$

because $\lim_t n_{\overline{y}}(t)/t = \beta$ by law of large numbers. Because for every $t \geq 1$, there exists $k \geq 0$ such that $t_k \leq t < t_{k+1}$ and $n_E(t)/t = n_E(t_k)/t \leq n_E(t_k)/t_k$, the above inequality also holds for the whole sequence

$$\limsup_{t\to\infty} \frac{n_E(t)}{t} \leq \frac{\lambda_{\xi(F)}}{\log \frac{b+1}{b}}.$$

This inequality states exactly what we have mentioned above: the fraction of entries along a typical history is determined by the size of surprise $\frac{b+1}{b}$ and the relative frequency of the exogenous signals $\lambda_{\xi(F)}$. Lastly, because this inequality holds for all Nash equilibria, we have

$$\liminf_{\delta\to 1} \underline{U}_1(\delta) \geq \big(1 - \frac{\lambda_{\xi(F)}}{\log \frac{b+1}{b}}\big)a + \frac{\lambda_{\xi(F)}}{\log \frac{b+1}{b}}(-1) = a - (a+1)\frac{\lambda_{\xi(F)}}{\log \frac{b+1}{b}}.$$

This is exactly the lower bound in (2.1).

Chapter 3

Social learning and market experimentation

3.1 Introduction

This chapter studies optimal dynamic monopoly pricing when a monopolist sells a product with unknown quality to a sequence of short-lived buyers who have private information about the quality. Because buyers purchase behavior conveys information about their private signals, the market, including both the monopolist and subsequent buyers, can gradually learn the quality of the product. Examples include that readers buy books that are best sellers, that smart phone users download apps that are heavily downloaded, that diners order food that are popular among other diners.[1]

Standard social learning literature (e.g. Bikhchandani, Hirshleifer, and Welch (1992) and Banerjee (1992)) only focused on the buyers' behavior of the above environment and analyzed the implications of information externality and learning,

[1]For example, Cai, Chen, and Fang (2009) found that in a randomized natural experiment in a restaurant dinning setting, when customers are given ranking information of the five most popular dishes, the demand for those dishes increases by 13 to 20 percent.

ignoring the fact that the monopolist can and will strategically adjust the price of the product as the learning proceeds. In fact, the different prices charged by the monopolist have important effects on the buyers' learning because the market belief is jointly determined by the price and the buyer's purchase behavior. For example, when we see that an expensive restaurant is crowded, we may think that the food there is very good. But when we see that a less expensive restaurant is crowded, we attribute it to low price. Hence from the monopolist's point of view, each price not only extracts rents in current period, but also creates an experiment that determines the information available to the market. As a result, when choosing prices, the monopolist is essentially doing experimentation in the market.

What is the optimal experimentation strategy for the monopolist? What are the implications on social learning when the monopolist optimally experiments? Will the monopolist have incentives to stop experimentation by either leaving the market or charging low price so that buyer's purchase behavior no longer provides information about their private signals? This chapter addresses these questions in a simple model. We assume the quality of the monopolist's product can be either high or low. Neither the monopolist nor the buyers know initially the true quality. The buyers enter the market sequentially and each of them is endowed with a private signal about the quality of the product. At the begining of each period, the monopolist can post a price and the entering buyer, after observing previous prices and purchase behavior his predecessors, decides whether to buy or not. Bose, Orosel, Ottaviani, and Vesterlund (2006) and Bose, Orosel, Ottaviani, and Vesterlund (2008) first studied a similar model and they focused on information structures that have only finitely many signals. Unlike their settings, this chapter focuses on information structures that have a continuum of signals and we show that the characterization of informational casacades is qualitatively different from their results.

We first observe that the monopolist's pricing problem is equivalent to setting cut-offs in posterior beliefs because in any Bayesian perfect equilibrium, after any history the entering buyer's willingness to pay increases with his poterior belief that the quality is high. This simple observation reduces the monopolist's problem into a dynamic programming problem. Based on this, we further show that, under a belief monotonicity condition on the information structure, the optimal cut-offs as a correspondence of market belief increases in the strong set order. This then implies that there always exists a Bayesian perfect equilibrium in which the monopolist posted prices increase with market belief. Moreover, if in any Bayesian perfect equilibrium the price charged after a history with high market belief is lower than that after a history with low market belief, then the monopolist must be indifferent between these two prices at these two histories.

We then characterize informational cascades and answer the question that whether and when the monopolist has incentive to stop experimentation. Here, we distiguish two cases about value of the product. One case is that the value of the low quality product is lower than the buyers' outside option, and the other is that the buyers' outside option is lower than the value of the product even if its quality is low. Propositions 3.4.1 and 3.4.2 fully characterize whether the monopolist will stop experimentation in terms of the information structure in these two cases respectively. In the first case, if the private signals are of unbounded informativeness, then the monopolist never stops experimentation. This is the same result as in Smith and Sørensen (2000). However, if the private signals are of bounded informativeness, although the monopolist always leaves the market when the market belief is low, whehther the monopolist induces herding on purchasing when the market belief is high depends on whether private signals are heavily distributed around the lowest possible signal. If the density of the private signals at the lowest signal is strictly positive, then the monopolist indeed will induce herding on purchasing when the market belief is high. But if the density of the private signals at the lowest signal is zero, then the monopolist will never stop experimentation. This is intuitive because in this case, the monopolist would never want to charge a low price even if he were myopic. Hence he never incentive to stop experimentation because experimentation will also bring in future value. This finding is very different from Smith and Sørensen (2000). The chracterization for the second case is similar to the first one, except the fact the monopolist never wants to leave the market. Whether the monopolist will induce herding on purchasing will again depends on whether the density at the lowest signal is zero, for *both* bounded and unbounded informativeness. Again, this is also different from Smith and Sørensen (2000). A large body of social learning papers have found Smith and Sørensen (2000)'s characterization robust to various modifications of the standard model. To our knowledge, this is the first time in this literature to find the subtle relationship between informational cascades and the density of private signals, besides its support.

Related Literature. The social learning framework was first introduced independently by Bikhchandani, Hirshleifer, and Welch (1992) and Banerjee (1992). These two papers have shown that in a sequential decision problem where each agent must make a decision from a finite set of actions after receiving his private signal about the unknow states and observing previous agents' actions, a herd arises. That is eventually all agents will choose the same (possibly wrong) action regardless of their own private signals. Smith and Sørensen (2000) provides a formal framework to analyze herding behavior systematically. They found that herding on a wrong action can occur if and only if the private signals are of bounded informativeness. That is the public information contained in previous decisions will eventually swamp the agents' private information if agents can only receive signals of uniformly bounded precision. In all these papers, all the agents face the same fixed, exogenously given decision problem. In contrast, the buyers in this chapter face different decision problems as the monopolist endogenously chooses prices.

Smith, Sørensen, and Tian (2015) conducted a welfare analysis in the above social learning framework. They assume that a social planner, who does not know the underlying state and can not directly observe private signals, can dictate each agent's decision rule and wants to maximize the discounded sum of expected utilities. They show that the social planner's optimal solution is cut-off rules and exhibits "contrarianism": agent should lean against taking the myopically better actions. Their paper and the current one share the similarity that a forward-looking agent maximizes long-run expected payoff by changing and learning from short-lived agents' behavior. In fact, the techniques they developed in showing contrarianism can be adapted to the current setting to show that the monopolist's optimal pricing rule is monotonic with respect to the market belief. However, we find very different characterization of informational cascades from theirs which is basically the same as Smith and Sørensen (2000). This is because in the current setting whehter a herd occurs depends on whether the monopolist has incentive to stop experimentation, and the monopolist's incentive to stop experimentation is determined by the distribution of private signals around the lowest possible signal.

Some papers have studied firm's pricing behavior in a market where the buyers have private information about the quality of the product. Moscarini and Ottaviani (2001) studies static price competition between two firms in a setting where each firm offers a variaty of a good to a buyer who receives a private binary signal on their relative quality. Because this is a static setting, there is no learning from the buyer's behavior. More closely related papers to the current one are Bose, Orosel, Ottaviani, and Vesterlund (2006) and Bose, Orosel, Ottaviani, and Vesterlund (2008). Similarly as this chapter, both papers study the monopolist's dynamic optimal pricing problem when buyers have private signals and can learn from other buyers' purchase behavior. The major difference between these two papers and the current one is that we focus on different kinds of information structures. While Bose, Orosel, Ottaviani, and Vesterlund (2006) considers information structure that contains finitely many possible signals and Bose, Orosel, Ottaviani, and Vesterlund (2008) considers binary information structure, this current chapter extend their analysis to information structures that has a continuum of signals. Both Bose, Orosel, Ottaviani, and Vesterlund (2006) and Bose, Orosel, Ottaviani, and Vesterlund (2008) show that information cascade must arise in these two models, as is the case in Bikhchandani, Hirshleifer, and Welch (1992) and Banerjee (1992), even in the prensence of a monopolist who endogenously and optimally charges prices. This result can be considered as robustness of Smith and Sørensen (2000)'s characterization of informational cascade, because discrete signals are always of bounded informativeness. In contrast, the characterization of cascade sets in Section 3.4 in this chapter shows different results. We show whether informational cascade occurs depends not only on the boundedness of the information structure, but also on how private signals are distributed at the lowest possible signal. Our result coincide with Bose, Orosel, Ottaviani, and Vesterlund (2006) and Bose, Orosel, Ottaviani, and Vesterlund (2008) when the information structure is bounded *and* private signals are heavily distributed around the lowest possible signal.[2] However if private signals are not heavily distributed around the lowest possible signal, then herding will not occur *even if* information structure is bounded.

3.2 Model

Time is discrete $t = 0, 1, 2, \cdots$. There is a long-lived monopolist who sells a product with unknown quality to an infinite sequence of short-lived buyers. The quality of the product can be either high (h) or low (l). Initially, the market, including both the monopolist and the buyers, has common prior belief about the quality of the product.

[2]Notice, when private signals are discretely distributed, there is always positive mass on the lowest possible signal.

Let $\pi_0 \in (0,1)$ be the common prior belief that the quality is high. Let $\delta \in (0,1)$ be the monopolist's discount factor.

Each buyer receives a private signal that conveys information about the quality of the product. The private signal, via Bayes' rule, results in a private belief $r \in (0,1)$. We assume private beliefs are conditionally i.i.d among the buyers with distribution F^h and F^l given quality h and l respectively. Moreover, F^h and F^l are mutually absolutely continuous. Therefore, they have the same support and no private signal perfectly reveals the quality of the product.[3] In the following analysis, we will concentrate on F^h and F^l that have continuous and strictly positive densities f^h and f^l respectively over the common support. Such a pair of distributions (F^h, F^l) is referred to as an information structure. Following Smith and Sørensen (2000), we say an information structure (F^h, F^l) is *bounded* if there exists $0 < \underline{r} < \overline{r} < 1$ such that $\text{supp}F^h = [\underline{r}, \overline{r}]$, and is *unbounded* if $\text{supp}F^h = [0,1]$.

The timing is as follows. At the beginning of each period, the monopolist announces a price $p_t \in \mathbb{R}$. Then a new buyer comes into the market. The buyer then decides whether to buy the product or not, based his information. If the quality of the product is high (resp., low) and the buyer buys at price p_t, his payoff is $h - p_t$ (resp. $l - p_t$). In both cases, the monopolist gets p_t (we normalize the cost of production to 0). If the buyer decides not to buy the product, then he gets his outside option v and the monopolist gets 0.

We assume that past prices and purchase behavior are all publicly observed by the monopolist and the buyers. However buyers' private signals are only observed by

[3] Smith and Sørensen (2000) Appendix A shows that given any pair of mutually absolutely continuous private signal distributions $(\hat{F}^h, \hat{F}^l)$, there exists an equivalent pair of mutually absolutely continuous private belief distributions (F^h, F^l) in the sense that $(\hat{F}^h, \hat{F}^l)$ and (F^h, F^l) always give the same distribution of posterior beliefs given any prior belief. Hence focusing on distributions of private beliefs is without loss of generality.

themselves. Hence a public history h^t of length t includes prices and the corresponding purchase behaviors in periods $s = 0, \cdots, t-1$, i.e. $h^t = (p_0, a_0, \cdots, p_{t-1}, a_{t-1}) \in \mathscr{H}^t \equiv (\mathbb{R} \times \{0,1\})^t$ where $a_s = 1$ means the sth buyer buys the product while $a_s = 0$ means he does not buy. A strategy of the monopolist is a mapping $\sigma_M : \cup_{t=0} \mathscr{H}^t \to \mathbb{R}$ and a strategy of the buyers is a mapping $\sigma_B : \cup_{t=0} \mathscr{H}^t \times \mathbb{R} \times \text{supp} F^H \to \{0,1\}$. Given a strategy σ_B of the buyers and a public history h^t, both the monopolist and the buyers can update their beliefs about the quality of the product via Bayes' rule. We call this belief as *market belief* since it is shared by all market participants. A Bayesian perfect equilibrium of this game is a pair of strategies (σ_M^*, σ_B^*) such that both the monopolist and the buyers are maximizing after any history. Formally,

Definition 3.2.1. A Bayesian perfect equilibrium of this game is a pair of strategies (σ_M^*, σ_B^*) such that for any history h^t, price p_t, and private signal r_t

1. given σ_B^*, $\sigma_M^*|_{h^t}$ maximizes the monopolist's expected continuation payoff $E\Big[(1-\delta)\sum_{s=t}^{\infty} \delta^{s-t} \sigma_M(h^s) \sigma_B^*(h^s, \sigma_M(h^s), r_s) \Big| h^t \Big]$,
2. $\sigma_B^*(h^t, p_t, r_t) \in \arg\max_{\varsigma \in \{0,1\}} E\Big[\big(Q - p_t\big)\varsigma + (1-\varsigma)v \Big| h^t, r_t\Big]$, where Q is the random variable for quality.

3.2.1 Preliminary result

For any $\pi \in (0,1)$ and $r \in \text{supp} F^h(r)$, let $q(\pi, r)$ be the buyer's posterior belief if the market belief is π and he receives private signal r. That is

$$q(\pi, r) = \frac{\pi r}{\pi r + (1-\pi)(1-r)}.$$

Then condition 2 in Definition 3.2.1 simply requires that after any history (h^t, p_t, r_t),

$$\sigma_B^*(h^t, p_t, r_t) = 1 \quad \text{if and only if} \quad q\big(\pi(h^t), r_t\big)h + \big[1 - q\big(\pi(h^t), r_t\big)\big]l - p_t \geq v, \tag{3.1}$$

where $\pi(h^t)$ denotes the market belief that the quality is high after public history h^t.[4] Thus in any Bayesian perfect equilibrium, after any history, the monopolist's price simply induces a cut-off in the buyer's potential posterior beliefs such that the buyer will buy if and only if his posterior belief is about this cut-off. As a result, we can equivalently think about the monopolist's problem as choosing a cut-off in the postential posterior beliefs given the current market belief.

Formally, let $G(\cdot|\pi)$ be the distribution of the potential posterior beliefs given market belief $\pi \in (0,1)$, i.e.,

$$G(q|\pi) = \int_{q(\pi,\cdot)\leq q} \big(\pi f^h(r) + (1-\pi) f^L(r)\big)\mathrm{d}r \quad \forall q \in [0,1].$$

Notice, because both f^h and f^l are continuous, $G(\cdot|\pi)$ also has a continuous density $g(\cdot|\pi)$. We have the following lemma.

Lemma 3.2.1. *Let (σ_M, σ_B) be a strategy profile. For each history h^t, define*

$$q_M(h^t) \equiv \begin{cases} q\big(\pi(h^t), \overline{r}\big) & \text{if } q\big(\pi(h^t), \overline{r}\big)h + \big(1 - q\big(\pi(h^t), \overline{r}\big)\big)l - \sigma_M(h^t) \leq v, \\ q\big(\pi(h^t), \underline{r}\big) & \text{if } q\big(\pi(h^t), \underline{r}\big)h + \big(1 - q\big(\pi(h^t), \underline{r}\big)\big)l - \sigma_M(h^t) \geq v, \\ \frac{\sigma_M(h^t)+v-l}{h-l} & \text{if otherwise} \end{cases} \tag{3.2}$$

Then (σ_M, σ_B) is a Bayesian perfect equilibrium if and only if for all h^t, p_t and r_t

(i) (3.1) holds,

(ii) $q_M(h^t)$ solves

$$\max_{q\in G(\cdot|\pi(h^t))} (1-\delta)\big[1 - G(q|\pi(h^t))\big]\big[qh + (1-q)l - v\big] \\ + \delta\big[1 - G(q|\pi(h^t))\big]V\big(\pi^P(\pi(h^t), q)\big) + \delta G(q|\pi(h^t))V\big(\pi^N(\pi(h^t), q)\big),$$

[4] We assume for simplicity when the buyer is indifferent between buying and not buying, he always buys the product. Because we focus on information structures that are absolutely continuous, this assumption only simplifies exposition and is not essential.

where $\pi^P(\pi, q) = \frac{\int_{\tilde{q}\geq q} \tilde{q} g(\tilde{q}|\pi)\mathrm{d}\tilde{q}}{1-G(q|\pi)}$ *and* $\pi^N(\pi, q) = \frac{\int_{\tilde{q}<q} \tilde{q} g(\tilde{q}|\pi)\mathrm{d}\tilde{q}}{G(q|\pi)}$, *and the function* $V(\cdot)$ *solves the following Bellman equation*

$$V(\pi) = \max_{q\in G(\,\cdot\,|\pi)} (1-\delta)\big[1-G(q|\pi)\big]\big[qh + (1-q)l - v\big] + \delta\big[1-G(q|\pi)\big]V\big(\pi^P(\pi, q)\big) + \delta G(q|\pi)V\big(\pi^N(\pi, q)\big). \tag{3.3}$$

(iii) $q_M(h^t) = q\big(\pi(h^t), \underline{r}\big)$ *implies* $\sigma_M(h^t) = q_M(h^t)h + (1-q_M(h^t))l - v$.

Lemma 3.2.1 reduces the equilibrium problem into a dynamic porgramming problem. Specifically, condition (ii) and (iii) states that a monopolist's strategy σ_M is part of a Bayesian perfect equilibrium if and only if its induced cut-offs in terms of posterior beliefs satsify Bellman equation (3.3), and when the monopolist decides to charge low price so that the incoming buyer will buy regardless of his private signal, the monopolist's must charge the highest possible price that induces this kind of behavior.

Equation (3.2) explains how each price is transformed into the posterior belief cut-off given the buyers' equilibrium behavior (3.1). Given the current market belief π, If the charged price p is so high (resp. low) that even the buyer with the most optimistic (resp. pessimistic) signal will not buy (resp. will buy), then the effective cut-off is just the highest (resp. lowest) possible posterior belief given market belief π. If the price is in the intermediate range so that the incoming buyer will buy if he receives optimistic private signal and will not buy if receives a pessimistic signal, then the effective cut-off is determined by $qh + (1-q)l - v = p$. From the buyers' equilibrium behavior (3.1), we know in this case the incoming buyer will buy if and only if his posterior belief is above q.

Then the Bellman equation (3.3) can be easily understood as follows. If the monopolist's price leads to cut-off q, then the probability that the incoming buys this product, or the demand, is $1 - G(q|\pi)$. Hence the expected myopic payoff to the monopolist is $\big[1 - G(q|\pi)\big]\big[qh + (1-q)l - v\big]$. Moreover, if the current period buyer purchases the product, then the next period market belief, after the market observes this purchase behavior, is going to be updated to $\pi^P(\pi, q) = \frac{\int_{\tilde{q} \geq q} \tilde{q} g(\tilde{q}|\pi) \mathrm{d}\tilde{q}}{1 - G(q|\pi)}$. On the other hand, if the buyer does not buy, then the next period market belief is $\pi^N(\pi, q) = \frac{\int_{\tilde{q} < q} \tilde{q} g(\tilde{q}|\pi) \mathrm{d}\tilde{q}}{G(q|\pi)}$. In this case, the expected continuation value is $\big[1 - G(q|\pi)\big] V\big(\pi^P(\pi, q)\big) + G(q|\pi) V\big(\pi^N(\pi, q)\big)$. The monopolist's optimal behavior is just to choose the cut-off that maximizes his total payoff. We call $V(\cdot)$ the monopolist's value function. We say q is an optimal cut-off at π if q solves the maximization problem on the right hand side of the Bellman equation (3.3) when the market belief is π. The following lemma summarizes some basic properties of V.

Lemma 3.2.2. *The monopolist's value function V is convex, increasing and Lipschitz continuous.*

The most insteresting property of V is its convexity. This means that the monopolist benefits from the information generated from buyers' purchase behavior. Because whether an incoming buyer buys or not depends on the current period price, the monopolist, by choosing different prices and thus different cut-offs in the potential posteriors, can determine in every period the nature of the information that will be generated from the buyer's response. In other words, different prices lead to different experiments. Because the monopolist can benifit from the information, when choosing a price, the monopolist takes into account his current period payoff and what kind of information to be generated. This is in the same spririt of experimentation.

3.3 Price monotonicity

This section studies an important feature of the monopolist's equilibrium price. We show under the following regularity conditions on the information structure, the monopolist's equilibrium price must satisfy centain monotonic patern. In particular, there must exists a Bayesian perfect equilibrium in which the monopolist's price increases with the market belief.

Definition 3.3.1. An information structure (F^h, F^l) satisfies belief monotonicity if for all $0 < \underline{\pi} < \overline{\pi} < 1$, the ratio of the densities of the correponding posterior beliefs

$$\frac{g(q|\overline{\pi})}{g(q|\underline{\pi})}$$

increases in $q \in \mathrm{supp} G(\,\cdot\,|\overline{\pi}) \cap \mathrm{supp} G(\,\cdot\,|\underline{\pi})$.

Belief monotonicity states that the distributions of the posterior beliefs updated from different market beliefs (prior beliefs) satisfy monotonicity likelihood ratio property. Notice, for any information structure (F^h, F^l), it is always true that higher prior belief leads to larger probability of getting a high posterior belief. That is $G(\,\cdot\,|\overline{\pi})$ first order stochastic dominates $G(\,\cdot\,|\underline{\pi})$ for all $0 < \underline{\pi} < \overline{\pi} < 1$. But belief monotonicity is stronger than this since it requires that the likelihood ratio be monotone. Hence, there are indeed information structures that violate belief monotonicity. Lemma 4 in Smith, Sørensen, and Tian (2015) provides a sufficient condition for the information structure (F^h, F^l) to satisfy belief monotonicity.

Lemma 3.3.1. *Assume the information structure (F^h, F^l) satisfies belief monotonicity. For $\underline{\pi} < \overline{\pi}$, assume $q^*(\underline{\pi})$ and $q^*(\overline{\pi})$ are optimal cut-offs at $\underline{\pi}$ and $\overline{\pi}$ respectively. If $q^*(\underline{\pi}) > q^*(\overline{\pi})$, then $q^*(\overline{\pi})$ is also an optimal cut-off at $\underline{\pi}$ and $q^*(\underline{\pi})$ is also an optimal cut-off at $\overline{\pi}$.*

Lemma 3.3.1 states that the set of optimal cut-offs as a correspondence of market belief increases with repect to the strong set order (see, for example, Milgrom and Shannon (1994) and Topkis (1998)). It is worthnoting that Lemma 3.3.1 implies that if the monopolist's optimal cut-off is unique for every market belief, then the optimal cut-offs as a function of the market belief must be increasing.[5] Moreover, Lemma 3.3.1 also implies that there always exists a Bayesian perfect equilibrium in which the monopolist's price increases with the market belief. This is summarized in the following proposition.

Proposition 3.3.1. *If the information structure* (F^h, F^l) *satisfies belief monotonicity, then there exists a Bayesian perfect equilibrium in which the monopolist's strategy* σ_M^* *increasese with respect to the market belief: for any* h^t *and* h^s*, if* $\pi(h^t) > \pi(h^s)$*, then* $\sigma_M^*(h^t) \geq \sigma_M^*(h^s)$*.*

As mentioned above, each price charged by the monopolist determines current period demand, which in turn leads to the current period payoff and additional information about the quality. The relationship between the price and the current period payoff is relatively straightforward. In fact, under belief monotonicity, if the monopolist were myopic, then it is easy to show that the monopolist' optimal prices satisfy the same monotonic property stated in Lemma 3.3.1.[6]

But how the current price changes the value of the information to the monopolist is less obvious. When the monopolist charges different intermediate prices, different kinds of experiments are induced. But how one experiment compares to another in terms of their values to the monopolist is not simple. To see this, notice that both $\pi^N(\pi, q)$ and $\pi^P(\pi, q)$ increases with q for a given π. This simply means that the experiments induced by different prices are not ranked in the usual Blackwell order. As a result, even though the monopolist always benefits from the additional

[5] If this is the case, then there exists an essentially unique Bayesian perfect equilibrium. In this equilibrium, the monopolist's strategy is Markovian. That is the monopolist's equilibrium strategy only depends on the current period market belief.

[6] To guarantee this, we only need that $G(q|\pi)$ satisfies monotone hazard rate property: $\frac{1-G(q|\underline{\pi})}{g(q|\underline{\pi})} \leq \frac{1-G(q|\overline{\pi})}{g(q|\overline{\pi})}$ for all $\underline{\pi} < \overline{\pi}$ and $q \in \text{supp}G(\cdot|\underline{\pi}) \cap \text{supp}G(\cdot, |\overline{\pi})$. It is well known that monotone hazard rate property is an implication of monotone likelihood ration property.

information, the relationship between the price and the value of information is not straightforward.

Nonetheless, Lemma 3.3.1 and Proposition 3.3.1 shows that if the information structure satisfies belief monotonicity, then roughly speaking, in equilibrium the monopolist still has incentive to charge high price when the market belief is high and low price when the market belief is low. One main step in the proof of Lemma 3.3.1 is to show that belief monotonicity guarantees that even if the monopolist only cares about his future payoff, then he has incentive to charge higher price when market belief is higher. This, cominbing with the fact that the monopolist is willing to charge higher price when market belief is higher if he is myopic, lead to the conclusion of Lemma 3.3.1.

3.4 Characterization of Cascade Sets

This section characterizes the cascade sets in any Bayesian perfect equilibrium of our model. In the standard social learning model without a monopolist (Banerjee (1992), Bikhchandani, Hirshleifer, and Welch (1992) and Smith and Sørensen (2000)), the cascade sets are defined to be the set of public beliefs at which the new coming agent's behavior only depend on the public belief and thus social learning stops. In the presence of a monopolist, these sets correspond to the set of market beliefs at which the monopolist stops market experimentation. That is the monopolist charges a price so that the incoming buyer's purchase behavior does not depend on his own private signal and thus later buyers can not draw inference based this buyer's purchase behavior. Therefore, the characterization of cascade sets will also characterize whether the learning process is complete or not.

For ease of exposition in this section, when we say a Bayesian perfect equilibrium q^*, we mean any Bayesian perfect equilibrium that is equivalent to q^* by Lemma 3.2.1.

Given any Bayesian perfect equilibrium q^*, define C^N to be the set of all market beliefs at which trade occurs with probability 0. In other words, the monopolist charges so high a price at these market beliefs so that the incoming buyer will not buy regardless of his private belief. Formally, we define

$$C^N \equiv \Big\{\pi \in (0,1) \Big| q^*(\pi) = \frac{\pi\overline{r}}{\pi\overline{r} + (1-\pi)(1-\overline{r})}\Big\} \bigcup \Big\{\pi \in \{0,1\} \Big| q^*(\pi)h + (1-q^*(\pi))l - v < 0\Big\}.$$

Similarly, define C^P to be the set of all market beliefs at which trade occurs with probability 1. That is, the monopolist charges so low a price at these market beliefs so that the incoming buyer will buy regardless of his private belief. Formally, we define

$$C^P = \Big\{\pi \in (0,1) \Big| q^*(\pi) = \frac{\pi\underline{r}}{\pi\underline{r} + (1-\pi)(1-\underline{r})}\Big\} \bigcup \Big\{\pi \in \{0,1\} \Big| q^*(\pi)h + (1-q^*(\pi))l - v \geq 0\Big\}.$$

The market beliefs contained in $C^N \cup C^P$ are absorbing states in the learning process. In any equilibrium, if the market belief π_t in some period t is in the set $C^N \cup C^P$, then in equilibrium the monopolist simply stops experimentation. He either charges a high enough or low enough price so that the incoming buyer's purchase decision will not depend on his own private signal and hence his purchase behavior provides no further information at all about the quality of the product. As a result, the market belief will not be updated and remain the same at π_t in all later periods and hence the learning process completely stops. In contrast, if the market period π_t in period t is outside $C^N \cup C^p$, then the monopolist does experimentation in the market by charing an intermediate price. As a result, the incoming buyer will buy the product only if he has received a high enough signal. The charged price and the purchase behavior of the incoming buyer together provide additional information about the quality of the product to both the monopolist and future buyers.

In the following, We distinguish two cases and characterize C^N and C^P in Bayesian perfect equilibria respectively.

3.4.1 $h > v > l$

Because $v > l$, when the market believes that the quality of the product is low for sure, the buyers will not buy the product since $q * (0) = 0$ for all Bayesian perfect equilibrium q^*. Hence $0 \in C^N$. Similarly, because $h > v$, when the market believes that the quality of the product is high for sure, the buyers will buy the product since $q^*(1) = 1$. Hence $1 \in C^P$. Therefore we have the following simple lemma.

Lemma 3.4.1. *When $h > v > l$, in any Bayesian perfect equilibrium, both C^N and C^P are nonempty. In particular, we have $0 \in C^N$ and $1 \in C^P$.*

Based on this lemma, we have the following definition

Definition 3.4.1. Assume $h > v > l$. The set C^k for $k \in \{N, P\}$ is *degenerate* if it is a singleton and *non-degenerate* if it contains at least one interval.

The following Proposition provides a full characterization of C^N and C^P in terms of various private information structure in any Bayesian perfect equilibrium.

Proposition 3.4.1. *Assume $h > v > l$. Then in any Bayesian perfect equilibrium,*

1. *if the information structure is unbounded, then both C^N and C^P are degenerate, i.e. $C^N = \{0\}$ and $C^P = \{1\}$,*

2. *if the information structure is bounded, then C^N is always non-degenerate, and C^P is degenerate if and only if $f^l(\underline{r}) = 0$.*

The first part of Proposition 3.4.1 states that both C^N and C^P are degenerate if $h > v > l$ and if the information structure is unbounded and $h > v > l$, which is the same as the result in Smith and Sørensen (2000). In any equilibrium, the monopolist never charges a price that induces purchase or stops sale regardless of what the buyer's private signal is. Instead, the monopolist always charges an intermediate price so that the incoming buyer buys the product only if his private signal is high enough. The reason is straightforward. Because the information structure is unbounded, there is always positive probability that the incoming buyer receives an extremely optimistic or pessimistic private signals. If the monopolist were to stop sale or sell to buyer with all possible signals, he can only get zero or negative profit respectively. However, by charging an intermediate price, the monopolist can always guarantee himself a strictly positive profit. As a result of this pricing behavior, learning in this market must be complete. In the long run, the market will eventually find out the true quality of the product.

The second part of Proposition 3.4.1 differs from that in Smith and Sørensen (2000). It states that although C^N is always non-degenerate, whether C^P is degenerate or not depends on the value of the density function at the lowest possible signal. The reason that C^N is always non-degenerate is straightforward. Specifically, in any equilibrium, when the market belief becomes very low, the monopolist stops sale because any price that leads to sale with positive probability must result in negative profit in the current period. The only reason that the monopolist would be willing to receive a negative profit in the current period is that current sale can largely boost the market belief and improve future profitability. But this is impossible when the market belief is already very low. The characterization of C^P is more complicated. When $f^L(\underline{r}) = 0$, the probability that a buyer receives a very low private signal is very small. Therefore, it is never the monopolist's myopic incentive to charge the lowest price in order to induce purchase since by charging a slightly higher price, the demand will not decline much. Moreover, if it is not myopically optimal to charge the lowest possible price, then it is not optimal to charge such a price because other prices can both increase myopic payoff and provide valuable information for the future. In contrast, when $f^L(\underline{r}) > 0$, then the probability that a buyer receives very low signal is larger than the case where $f^L(\underline{r}) = 0$. Moreover, this probability becomes nonnegligible especially when the market becomes very optimistic about the quality. In this case, it is indeed the monopolist's myopic incentive to charge a price so that the incoming buyer will buy regardless of his private information, because increasing price will lead to a large decline in demand. Furthermore, as the market belief becomes higher, the potential value of additional information to the monopolist becomes very small. As a result, when the market belief is high, the monopolist simply stops experimentation and charges a price that all buyers will buy regardless of their private information.

Hence, we have

Corollary 3.4.1. *Assume $h > v > l$. In any Bayesian perfect equilibrium, if the information structure is unbounded, then learning is complete, and if the information structure is bounded, then learning is incomplete.*

3.4.2 $h > l > v$

Because $l > v$, even if the market belief is $\pi = 0$, the monopolist is still willing to sell the product. Hence it is clear that $C^N = \emptyset$ and $\{0, 1\} \subset C^P$ in any equilibrium. Similarly as Definition 3.4.1, we now have

Definition 3.4.2. Assume $h > l > v$. The set C^P is *degenerate* if $C^P = \{0, 1\}$, and *non-degenerate* if it contains at least one interval.

We have the following full characterization of C^P when $h > l > v$ which is an analogue of Proposition 3.4.1. The main idea of the proofs is similar to that of Proposition 3.4.1 and thus is omitted.

Proposition 3.4.2. *Assume $h > l > v$. Then for both unbounded and bounded information structure, in any equilibrium C^P is degenerate if and only if $f^l(\underline{r}) = 0$.*

Proposition 3.4.2 states that whether C^P is degenerate when $h > l > v$ now does not depend on whether the information structure is bounded or not. Rather, it solely depends on the value of the density of the private signals at the lowest possible signal. When information structure is bounded, the result is essentially identical to Proposition 3.4.1. When $f^L(\underline{r}) = 0$, then in any equilibrium, it is never the monopolist's myopic incentive to charge a low price so that the buyer will buy regardless of his private information. As a result, the monopolist is never willing to do so because there is another price that give him higher current period profit as well as valuable information. However, when $f^L(\underline{r}) > 0$, as in the case with $h > v > l$, the monopolist has strict myopic incentive to charge a low price when the market is either extremely high or extremely low. Then it will also be the monopolist's long-run incentive to do so because the value of information from experimentation would be small with extreme market beliefs.

When information structure is unbounded, the result is different from that with $h > v > l$ because now even if the market knows that the quality is low, the monopolist is still willing to sell the product at price l and make strictly positive profit. Hence the same logic as above then applies. When $f^L(\underline{r}) > 0$ and the market belief is low, the monopolist just has incentive to stop experimentation by charging low price and sell to buyers with all values of private signals.

In terms of learning, we have

Corollary 3.4.2. *When $h > l > v$, for both bounded and unbounded information structures, in any Bayesian perfect equilibrium, learning is complete if and only if $f^l(\underline{r}) = 0$.*

3.5 Further discussions

This chapter studies monopolist's optimal dynamic pricing problem in a market where buyers have private information regarding the quality of the product and they can also infer information from other buyers' purchasing behavior. Unlike standard social learning papers where the buyers all face a fixed price throughout the learning process, this chapter focuses on the monopolist's incentives to maximize profits and control information for future buyers by charging different prices at different market beliefs.

Moreover, we fully characterize the informational cascades in any equilibrium. Our results show that whether the monopolist has incentives to stop experimentation and induce herding on purchasing depends not only on whether the buyers' private signals are of bounded informativeness, but also on whether private signals are heavily distributed around the lowest possible signals. This characterization differs from the standard results in social learning literature and largely extends the results obtained in similar problems. In particular, this is the first time in social learning literature to find the relationship between herding and the detailed property of the densities of private signals besides its support.

One interesting avenue for future research is to consider the case that subsequent buyers only get sales information about the past. In this case, the buyers' inference problem become much more complicated since they have to speculate on how many buyers have endtered the market but decided not to buy. How to find a analytically tractable model to capture this idea remains a challege.

Appendix A

Proofs for Chapter 1

A.1 Mathematical preliminaries

This section lists three basic identities that will be used repeatedly in later analysis.

Lemma A.1.1. *Let $m \geq 0$ and $n \geq 1$ be integers. Then the total number of ways to put m indistinguishable balls into n distinguishable bins is $\begin{pmatrix} m+n-1 \\ n-1 \end{pmatrix}$. Formally*

$$\sum_{\substack{m_1+\cdots+m_n=m \\ m_i \geq 0}} 1 = \begin{pmatrix} m+n-1 \\ n-1 \end{pmatrix} \equiv \frac{(m+n-1)!}{m!(n-1)!}.$$

Proof. See Feller (1968), Section II.5. □

Lemma A.1.2. *Let $\sum_{m=0}^{\infty} c_{1m} x_1^m, \cdots, \sum_{m=0}^{\infty} c_{nm} x_n^m$ be n power series. Then in their convergence ranges, we have*

$$\sum_{m=0}^{\infty} \sum_{\substack{m_1+\cdots+m_n=m \\ m_i \geq 0}} \prod_{k=1}^{m} c_{km_k} x_k^{m_k} = \prod_{k=1}^{n} \Big(\sum_{m=0}^{\infty} c_{km} x_k^m \Big).$$

In particular,

$$\Big(\sum_{m=0}^{\infty} x^m \Big)^{n+1} = \sum_{m=0}^{\infty} \begin{pmatrix} m+n \\ n \end{pmatrix} x^m \quad \forall x \in [0,1).$$

Proof. The first part is just standard multiplication of convergent series. The second part comes from Lemma A.1.1. □

Lemma A.1.3 (Euler's continued fraction formula)**.** *Let* $a_1, \cdots, a_n$ *be real numbers. Then*

$$\sum_{k=1}^{n} \prod_{i=1}^{k} a_i = \cfrac{a_1}{1 - \cfrac{a_2}{1 + a_2 - \cfrac{a_3}{1 + a_3 - \cfrac{a_4}{\cfrac{\ddots}{1 + a_{n-1} - \cfrac{a_{n-1}}{1 + a_{n-1} - \cfrac{a_n}{1 + a_n}}}}}}}.$$

Proof. See Wall (1967), Theorem 2.1. □

A.2 Distributions of signals under stationary strategy

Fix a stationary public strategy σ_1 of the seller. Recall σ_1 induces a probability measure $P_{\tilde{\xi}}^{\sigma_1}$ over R^∞.

This measure $P_{\tilde{\xi}}^{\sigma_1}$ defines a Markov chain $\{R_t\}_{t \geq 0}$ with initial state 0 and the following transition rules:

$$\tilde{x}_0 \equiv P_{\tilde{\xi}}^{\sigma_1}(R_{t+1} = 1 | R_t = 0) = (1 - \varepsilon) \times \rho(h|\sigma_1(r)), \tag{A.1}$$

$$\tilde{z}_0 \equiv P_{\tilde{\xi}}^{\sigma_1}(R_{t+1} = 0 | R_t = 0) = 1 - (1 - \varepsilon) \times \rho(h|\sigma_1(r)), \tag{A.2}$$

and for $r \geq 1$

$$\begin{aligned}
\tilde{x}_r &\equiv P^{\sigma_1}_{\tilde{\xi}}(R_{t+1} = r+1 | R_t = r) = (1-\varepsilon) \times \rho(h|\sigma_1(r)), && \text{(A.3)} \\
\tilde{z}_r &\equiv P^{\sigma_1}_{\tilde{\xi}}(R_{t+1} = r | R_t = r) = \varepsilon \times \rho(h|\sigma_1(r)) + (1-\varepsilon) \times \rho(l|\sigma_1(r)), && \text{(A.4)} \\
\tilde{y}_r &\equiv P^{\sigma_1}_{\tilde{\xi}}(R_{t+1} = r-1 | R_t = r) = \varepsilon \times \rho(l|\sigma_1(r)). && \text{(A.5)}
\end{aligned}$$

The values $\tilde{x}_r$, $\tilde{y}_r$ and $\tilde{z}_r$ represent respectively the probabilities of *upward*, *downward* and *horizontal* transitions of the signals at r given seller's strategy σ_1.[1]

In the remainder of this section, we derive the formula for $P^{\sigma_1}_{\tilde{\xi}}(R_t = r)$, the probability of signal r in period t, in terms of $\{\tilde{x}_{\tilde{r}}, \tilde{y}_{\tilde{r}}, \tilde{z}_{\tilde{r}}\}_{\tilde{r}}$ for arbitrary $t \geq r$.[2] For any $r \in R$ and $t \geq r$, let $\Phi^{r,t} \subset R^{t+1}$ be the set of all possible paths from signal 0 to signal r in t periods that have positive probability under $P^{\sigma_1}_{\tilde{\xi}}$, i.e.

$$\Phi^{r,t} \equiv \left\{ (r_0, r_1, \cdots, r_t) \in R^{t+1} \;\middle|\; \begin{array}{l} (1) r_0 = 0 \text{ and } r_t = r, \\ (2) r_1 - r_0 \in \{0, 1\}, \\ (3) r_s - r_{s-1} \in \{-1, 0, 1\} \text{ for } 2 \leq s \leq t. \end{array} \right\}.$$

Clearly, $P^{\sigma_1}_{\tilde{\xi}}(R_t = r) = P^{\sigma_1}_{\tilde{\xi}}(\Phi^{r,t}) = \sum_{h^{t+1} \in \Phi^{r,t}} P^{\sigma_1}_{\tilde{\xi}}(h^{t+1})$. The main idea of the following calculation is to partition $\Phi^{r,t}$ into several cells of paths within which all paths have the same probability. We then count the number of paths in each cell and sum up all the probabilities. The partition of paths is based on the (1) highest signals reached (2) the number of upward and downward moves at each signal, and (3) the number of horizontal moves at each signal. The whole analysis is divided into 5 small steps. In Step 1, we partition $\Phi^{r,t}$ according to (1). In Steps 2-3, we refine partitions into smaller cells according to (2) and (3) respectively. In Step 4, we calculate the number of paths in each final cell. In Step 5, we give the expression for $P^{\sigma_1}_{\tilde{\xi}}(\Phi^{r,t})$.

[1]For notational simplicity, we drop the dependence of $\{\tilde{x}_r, \tilde{y}_r, \tilde{z}_r\}_r$ on σ_1.

[2]It is easy to see $P^{\sigma_1}_{\tilde{\xi}}(R_t = r) = 0$ if $t < r$.

Step 1: Partitioning $\Phi^{r,t}$ according to the highest signal reached along each path.

For each $k = 0, 1, \cdots, \lfloor \frac{t-r}{2} \rfloor$, define

$$\Phi^{r,t}(k) \equiv \Big\{(r_0, \cdots, r_t) \in \Phi^{r,t} \Big| \max_{0 \le s \le t} r_s = r + k\Big\},$$

where $\lfloor \frac{t-r}{2} \rfloor$ is the largest integer less than or equal to $(t-r)/2$. The set $\Phi^{r,t}(k)$ contains the set of all possible paths in $\Phi^{r,t}$ such that the highest signal reached along these paths is $r+k$. Because the signals can go up or down by at most 1, if a path $h^{t+1} \in \Phi^{r,t}$ reached signal $r+k$, then it needs at least $r+2k$ periods because it takes at least $r+k$ periods to reach $r+k$ and another k periods to decreases to r. Hence there is an upper bound of k: $r + 2k \le t$, or equivalently $k \le \lfloor \frac{t-r}{2} \rfloor$. Then it is easy to see $\{\Phi^{r,t}(k)\}_{k=0}^{\lfloor \frac{t-r}{2} \rfloor}$ forms a partition of $\Phi^{r,t}$.

Step 2: Partitioning $\Phi^{r,t}(k)$ according to number of upward and downward moves at each signal.

First consider a path $(r_0, r_1, \cdots, r_t) \in \Phi^{r,t}(k)$. We know the highest signal reached along this path is $r+k$ for some $0 \le k \le \lfloor \frac{t-r}{2} \rfloor$. Let $n_{\tilde{r}}$ and $l_{\tilde{r}}$ denote the number of upward moves and downward moves at $\tilde{r}$ respectively, for $0 \le \tilde{r} \le r+k$. That is, for $0 \le \tilde{r} \le r+k$,

$$n_{\tilde{r}} \equiv \#\Big\{0 \le s \le t-1 \Big| r_s = \tilde{r} \text{ and } r_{s+1} = \tilde{r}+1\Big\}$$

and

$$l_{\tilde{r}} \equiv \#\Big\{0 \le s \le t-1 \Big| r_s = \tilde{r} \text{ and } r_{s+1} = \tilde{r}-1\Big\}.$$

For $\{n_{\tilde{r}}, l_{\tilde{r}}\}$, we make the following simple observations.

Lemma A.2.1. *1.* $n_{\tilde{r}} \ge 1$ *for* $0 \le \tilde{r} \le r+k-1$ *and* $n_{r+k} = l_0 = 0$.

2. $l_{\tilde{r}+1} = n_{\tilde{r}} - 1, \ \forall\, 0 \le \tilde{r} \le r-1$.

3. $l_{\tilde{r}+1} = n_{\tilde{r}}, \ \forall\, r \le \tilde{r} \le r+k-1$.

4. $\sum_{\tilde{r}=0}^{r+k}\left(n_{\tilde{r}}+l_{\tilde{r}}\right)=\sum_{\tilde{r}=0}^{r+k-1}n_{\tilde{r}}+\sum_{\tilde{r}=1}^{r+k}l_{\tilde{r}}\leq t.$

Proof. 1. Because the signals can increase by only 1 in every period, for a path to reach $r+k$ from 0, there must be at least one upward move at every signal $\tilde{r}\in\{0,1,\cdots,r+k-1\}$. Hence $n_{\tilde{r}}\geq 1$ for all $0\leq\tilde{r}\leq r+k-1$. It is clear $n_{r+k}=0$, otherwise the highest reached signal would be $r+k+1$. Because there is no downward move at 0, $l_0=0$.

2 and 3. Because the path eventually reaches r from 0, the number of upward moves at $0\leq\tilde{r}\leq r-1$ must exceed that of downward moves at $\tilde{r}+1$ by exact 1. Thus 2 follows. Every upward move at signal $\tilde{r}\geq r$ must be coupled with a downward move at $\tilde{r}+1$. Thus 3 follows.

4. The equality comes from property 1. The inequality comes form the fact that the total number of upward and downward moves must be bounded above by t, i.e. $\sum_{\tilde{r}=0}^{r+k}\left(n_{\tilde{r}}+l_{\tilde{r}}\right)\leq t$. □

We are now ready to partition $\Phi^{r,t}(k)$ according to the number of upward and downward moves at each signal. For $0\leq k\leq\lfloor\frac{t-r}{2}\rfloor$, define

$$\mathcal{I}^{r,t}(k)\equiv \tag{A.6}$$

$$\left\{(n_0,\cdots,n_{r+k-1},l_1,\cdots,l_{r+k})\in\mathbb{Z}_+^{2r+2k}\;\middle|\;\begin{array}{l}(1)\; n_{\tilde{r}}\geq 1\;\forall\,0\leq\tilde{r}\leq r+k-1,\\(2)\; l_{\tilde{r}+1}=n_{\tilde{r}}-1,\;\forall\,0\leq\tilde{r}\leq r-1,\\(3)\; l_{\tilde{r}+1}=n_{\tilde{r}},\;\forall\,r\leq\tilde{r}\leq r+k-1,\\(4)\;\sum_{\tilde{r}=0}^{r+k-1}n_{\tilde{r}}+\sum_{\tilde{r}=1}^{r+k}l_{\tilde{r}}\leq t.\end{array}\right\}. \tag{A.7}$$

Notice conditions (1)-(4) in this definition correspond to properties 1-4 in Lemma A.2.1. For notational simplicity, we write $(\vec{n},\vec{l})$ to denote a generic element in $\mathcal{I}^{r,t}(k)$. Also as a convention, for any $(\vec{n},\vec{l})\in\mathcal{I}^{r,t}(k)$, when n_{r+k} and l_0 are involved in the following context, they should be understood as 0 because of property 1 in Lemma A.2.1.

For every $(\vec{n},\vec{l}) \in \mathcal{I}^{r,t}(k)$, let $\Phi^{r,t}(k)(\vec{n},\vec{l})$ be the set of paths in $\Phi^{r,t}(k)$ whose number of upward and downward moves at each signal corresponds to $(\vec{n},\vec{l})$. Lemma A.2.1 shows for each history $h^{t+1} \in \Phi^{r,t}(k)$, its number of upward and downward moves at each signal $(\vec{n},\vec{l})$ must be in $\mathcal{I}^{r,t}(k)$. Hence $\big\{\Phi^{r,t}(k)(\vec{n},\vec{l})\big\}_{(\vec{n},\vec{l})\in\mathcal{I}^{r,t}(k)}$ forms a partition of $\Phi^{r,t}(k)$.

Step 3: Partitioning $\Phi^{r,t}(k)(\vec{n},\vec{l})$ according to the number of horizontal moves at each signal.

Given $(\vec{n},\vec{l}) \in \mathcal{I}^{r,t}(k)$, consider a history $(r_0,\cdots,r_t) \in \Phi^{r,t}(k)(\vec{n},\vec{l})$. Let $m_{\tilde{r}}$ denote the number of horizontal moves at $0 \le \tilde{r} \le r+k$ along this path. That is for $0 \le \tilde{r} \le r+k$,

$$m_{\tilde{r}} = \#\big\{0 \le s \le t-1 \big| r_s = r_{s+1} = \tilde{r}\big\}.$$

It is clear that $m_{\tilde{r}} \ge 0$ for all $\tilde{r}$. Moreover, the sum of upward, downward and horizontal moves over all signals corresponds to the length of this path, i.e. t. Hence

$$\sum_{\tilde{r}=0}^{r+k}\big(n_{\tilde{r}} + l_{\tilde{r}} + m_{\tilde{r}}\big) = t,$$

or equivalently

$$\sum_{\tilde{r}=0}^{r+k} m_{\tilde{r}} = t - \sum_{\tilde{r}=0}^{r+k}\big(n_{\tilde{r}} + l_{\tilde{r}}\big).$$

Define

$$J^{r,t}(k)(\vec{n},\vec{l}) \equiv \Big\{(m_0,\cdots,m_{r+k}) \in \mathbb{Z}_+^{r+k+1} \Big| \sum_{\tilde{r}=0}^{r+k} m_{\tilde{r}} = t - \sum_{\tilde{r}=0}^{r+k}\big(n_{\tilde{r}} + l_{\tilde{r}}\big).\Big\}. \tag{A.8}$$

Similarly as before, we use $\vec{m}$ to denote a generic element in $J^{r,t}(\vec{n},\vec{l})$. Let $\Phi^{r,t}(k)(\vec{n},\vec{l})(\vec{m})$ be the set of all paths in $\Phi^{r,t}(k)(\vec{n},\vec{l})$ whose number of horizontal moves at each signal corresponds to $\vec{m}$. From above analysis, we know $\big\{\Phi^{r,t}(k)(\vec{n},\vec{l})(\vec{m})\big\}_{\vec{m}\in\mathcal{J}^{r,t}(k)(\vec{n},\vec{l})}$ forms a partition of $\Phi^{r,t}(k)(\vec{n},\vec{l})$.

Through Step 1-3, we can write

$$\Phi^{r,t} = \bigcup_{k=0}^{\lfloor \frac{t-r}{2} \rfloor} \bigcup_{(\vec{n},\vec{l}) \in \mathcal{I}^{r,t}(k)} \bigcup_{\vec{m} \in \mathcal{J}^{r,t}(k)(\vec{n},\vec{l})} \Phi^{r,t}(k)(\vec{n},\vec{l})(\vec{m}).$$

Importantly, for all k, $(\vec{n},\vec{l})$ and $\vec{m}$, all histories in $\Phi^{r,t}(k)(\vec{n},\vec{l})(\vec{m})$ have the same probability under $P^{\sigma_1}_{\tilde{\xi}}$. In particular, if $h^{t+1} \in \Phi^{r,t}(k)(\vec{n},\vec{l})(\vec{m})$, then

$$P^{\sigma_1}_{\tilde{\xi}}(h^{t+1}) = \prod_{\hat{r}=0}^{r+k-1} \tilde{x}_{\hat{r}}^{n_{\hat{r}}} \tilde{y}_{\hat{r}+1}^{l_{\hat{r}+1}} \times \prod_{\bar{r}=0}^{r+k} \tilde{z}_{\bar{r}}^{m_{\bar{r}}}.$$

Therefore,

$$P^{\sigma_1}_{\tilde{\xi}}(\Phi^{r,t}) = \sum_{k=0}^{\lfloor \frac{t-r}{2} \rfloor} \sum_{(\vec{n},\vec{l}) \in \mathcal{I}^{r,t}(k)} \sum_{\vec{m} \in \mathcal{J}^{r,t}(k)(\vec{n},\vec{l})} \#\big(\Phi^{r,t}(k)(\vec{n},\vec{l})(\vec{m})\big) \times \prod_{\hat{r}=0}^{r+k-1} \tilde{x}_{\hat{r}}^{n_{\hat{r}}} \tilde{y}_{\hat{r}+1}^{l_{\hat{r}+1}} \times \prod_{\bar{r}=0}^{r+k} \tilde{z}_{\bar{r}}^{m_{\bar{r}}}. \tag{A.9}$$

To get an expression for $P^{\sigma_1}_{\tilde{\xi}}(\Phi^{r,t})$, it remains to calculate the number of paths in $\Phi^{r,t}(k)(\vec{n},\vec{l})(\vec{m})$. We do this in the next step.

Step 4: Calculating $\#\big(\Phi^{r,t}(k)(\vec{n},\vec{l})(\vec{m})\big)$.

Fix $0 \le k \le \lfloor \frac{t-r}{2} \rfloor$, $(\vec{n},\vec{l}) \in \mathcal{I}^{r,t}(k)$ and $\vec{m} \in \mathcal{I}^{r,t}(k)(\vec{n},\vec{l})$. If $r = 0$ and $k = 0$, then it is easy to see $(\vec{n},\vec{l}) = (\vec{0},\vec{0})$ and $\vec{m} = (t)$. Therefore $\#\big(\Phi^{r,t}(k)(\vec{n},\vec{l})(\vec{m})\big) = 1$. In what follows, we assume $r + k \ge 1$.

From each path $h^{t+1} = (r_0, \cdots, r_t) \in \Phi^{r,t}(k)(\vec{n},\vec{l})(\vec{m})$, we can obtain a new, shortened path, denoted by $\psi(h^{t+1})$, by removing all horizontal moves in h^{t+1} while keeping the number and order of upward and downward moves. Formally, define $t_0 = 0$ and inductively $t_j = \min\{t_{j-1} < s \le t | r_s \ne r_{t_{j-1}}\}$. This induction ends in finite steps and let $\bar{j}$ be the largest index. Because $h^{t+1} \in \Phi^{r,t}(k)(\vec{n},\vec{l})(\vec{m})$, it is straightforward to see $\bar{j} = \sum_{\tilde{r}=0}^{r+k}(n_{\tilde{r}} + l_{\tilde{r}})$ which is independent of h^{t+1}. Moreover, the resulting path $\psi(h^{t+1}) \equiv (r_0, r_{t_1}, \cdots, r_{t_{\bar{j}}})$ is a path in $\Phi^{r,\bar{j}}(k)(\vec{n},\vec{l})$. [3] The new path

[3]It is easy to see all paths in $\Phi^{r,\bar{j}}(k)(\vec{n},\vec{l})$ contain no horizontal moves because $\bar{j} = \sum_{\tilde{r}=0}^{r+k}(n_{\tilde{r}} + l_{\tilde{r}})$.

$(r_0, r_{t_1}, \cdots, r_{t_{\bar{j}}})$ preserves all the upward and downward moves in h^{t+1} and contains no horizontal moves. The above operation defines a mapping

$$\psi : \Phi^{r,t}(k)(\vec{n}, \vec{l})(\vec{m}) \to \Phi^{r,\bar{j}}(k)(\vec{n}, \vec{l}).$$

We use h to denote a typical path in $\Phi^{r,\bar{j}}(k)(\vec{n}, \vec{l})$. Then we have

$$\Phi^{r,t}(k)(\vec{n}, \vec{l})(\vec{m}) = \bigcup_{h \in \Phi^{r,\bar{j}}(k)(\vec{n}, \vec{l})} \psi^{-1}(\{h\}),$$

where ψ^{-1} is the pre-image of ψ. Figure A.1 presents an illustration of ψ and ψ^{-1}.

Lemma A.2.2. *For each $h \in \Phi^{r,\bar{j}}(k)(\vec{n}, \vec{l})$, we have*

$$\#\big(\psi^{-1}(\{h\})\big) = \prod_{\tilde{r}=0}^{r+k} \begin{pmatrix} m_{\tilde{r}} + n_{\tilde{r}} + l_{\tilde{r}} - 1 \\ n_{\tilde{r}} + l_{\tilde{r}} - 1 \end{pmatrix}.$$

Proof. Since the mapping ψ removes horizontal moves, ψ^{-1} adds them back. See Figure A.1. Notice along the path h, signal $\tilde{r}$ is passed $n_{\tilde{r}} + l_{\tilde{r}}$ times for all $0 \le \tilde{r} \le r+k$. If we are to add $m_{\tilde{r}}$ number of horizontal moves to h, there are

$$\begin{pmatrix} m_{\tilde{r}} + n_{\tilde{r}} + l_{\tilde{r}} - 1 \\ n_{\tilde{r}} + l_{\tilde{r}} - 1 \end{pmatrix}$$

number of ways to add, by Lemma A.1.1. Then the result follows. □

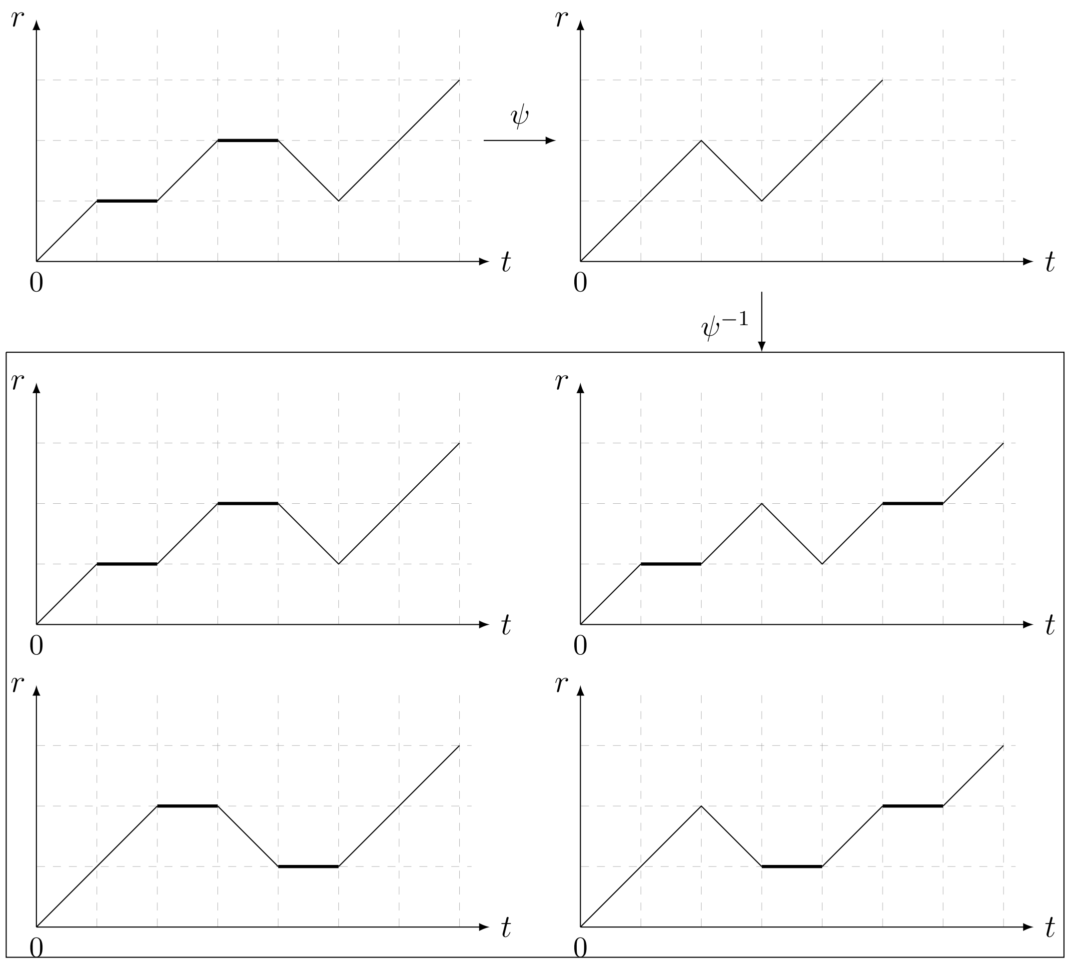
Notice $\#\big(\psi^{-1}(\{h\})\big)$ is independent of h. Thus

$$\#\Big(\Phi^{r,t}(k)(\vec{n}, \vec{l})(\vec{m})\Big) = \#\big(\Phi^{r,\bar{j}}(k)(\vec{n}, \vec{l})\big) \times \prod_{\tilde{r}=0}^{r+k} \begin{pmatrix} m_{\tilde{r}} + n_{\tilde{r}} + l_{\tilde{r}} - 1 \\ n_{\tilde{r}} + l_{\tilde{r}} - 1 \end{pmatrix}. \quad \text{(A.10)}$$

Therefore, what remains is to calculate $\#\big(\Phi^{r,\bar{j}}(k)(\vec{n}, \vec{l})\big)$. We do so in the following analysis.

Fix $(\vec{n}, \vec{l}) = (n_0, \cdots, n_{r+k-1}, l_1, \cdots, l_{r+k}) \in \mathcal{I}^{r,\bar{j}}(k)$. Recall $\bar{j} = \sum_{\tilde{r}=0}^{r+k}(n_{\tilde{r}} + l_{\tilde{r}})$. Now we construct a sequence of vectors $(\vec{n}^s, \vec{l}^s)_{s=0}^{r+k}$ as follows: $\forall\ 0 \le s \le r+k$,

Figure A.1: Illustration of ψ and ψ^{-1}.



The top-left panel illustrates a typical path $h \in \Phi^{r,t}(k)(\vec{n},\vec{l})(\vec{m})$ where it is readily to read off from the graph $r = 3$, $t = 7$, $k = 0$, $\vec{n} = (1,2,1,0)$, $\vec{l} = (0,0,1,0)$ and $\vec{m} = (0,1,1,0)$. The top-right panel is $\psi(h)$ which is obtained by removing the two horizontal moves (bold lines) in h while keeping the number and order of upward and downward moves. It is straightforward to see $\psi(h) = \Psi^{r,\bar{j}}(k)(\vec{n},\vec{l})$ where $\bar{j} = \sum_{\tilde{r}=0}^{r+k}(n_{\tilde{r}} + l_{\tilde{r}}) = 5$. The bottom four panels depict all paths in $\psi^{-1}(\psi(h))$ which are obtained by adding one horizontal move at $\tilde{r} = 1$ and one at $\tilde{r} = 2$ to $\psi(h)$, because $\vec{m} = (0,1,1,0)$. Because $\psi(h)$ reaches $\tilde{r} = 1$ and $\tilde{r} = 2$ two times respectively, for each $\tilde{r} = \{1,2\}$, there are two ways to add one horizontal move at $\tilde{r}$. Hence the total number of ways to add these horizontal moves is $2 \times 2 = 4$.

$$n^s_{\tilde{r}} \equiv \begin{cases} 1 & \text{if } 0 \le \tilde{r} < s \\ n_{\tilde{r}} & \text{if } s \le \tilde{r} \le r+k-1. \end{cases}$$

$$l^s_{\tilde{r}+1} \equiv \begin{cases} n^s_{\tilde{r}} - 1 & \text{if } 0 \le \tilde{r} \le r-1 \\ n^s_{\tilde{r}} & \text{if } r \le \tilde{r} \le r+k-1 \end{cases}$$

$$j^s \equiv \sum_{\tilde{r}=0}^{r+k} \left(n^s_{\tilde{r}} + l^s_{\tilde{r}}\right).$$

For example, when $s = r+k-1$, $n^s = (1, 1, \cdots, 1, n_{r+k-1}) \in \mathbb{N}^{r+k}$, and when $s = 0$, $(\vec{n}^s, \vec{l}^s) = (\vec{n}, \vec{l})$ and $j^s = \bar{j}$. Notice by construction, for every s, $(\vec{n}^s, \vec{l}^s) \in \mathcal{I}^{r,j^s}(k)$. Hence $\Phi^{r,j^s}(k)(\vec{n}^s, \vec{l}^s)$ is well defined. Because by construction $j^s \equiv \sum_{\tilde{r}=0}^{r+k}\left(n^s_{\tilde{r}} + l^s_{\tilde{r}}\right)$, all paths in $\Phi^{r,j^s}(k)(\vec{n}^s, \vec{l}^s)$ involve only upward and downward moves, no horizontal moves.

The main idea of the following analysis is to show all paths in $\Phi^{r,j^s}(k)(\vec{n}^s, \vec{l}^s)$ can be "constructed" from paths in $\Phi^{r,j^{s+1}}(k)(\vec{n}^{s+1}, \vec{l}^{s+1})$ and each path in $\Phi^{r,j^{s+1}}(k)(\vec{n}^{s+1}, \vec{l}^{s+1})$ generates a fixed number of paths in $\Phi^{r,j^s}(k)(\vec{n}^s, \vec{l}^s)$. This allows us to calculate $\#\big(\Phi^{r,\bar{j}}(k)(\vec{n}, \vec{l})\big)$ by induction. Given this main idea, we assume $n_{\tilde{r}} > 1$ for all $0 \le \tilde{r} \le r+k-1$. Otherwise, if $n_{\tilde{r}} = 1$ for some $\tilde{r}$, then it is easy to see by construction $\Phi^{r,j^{s+1}}(k)(\vec{n}^{s+1}, \vec{l}^{s+1}) = \Phi^{r,j^s}(k)(\vec{n}^s, \vec{l}^s)$ for $s = \tilde{r}$ and the question becomes trivial.

For any $0 \le s \le r+k-2$, consider a path $h = (r_0, \cdots, r_{j^s})$ in $\Phi^{r,j^s}(k)(\vec{n}^s, \vec{l}^s)$. We know along this path, there are $n^s_s = n_s$ periods, denoted by $t_1 < t_2 < \cdots < t_{n^s_s}$, at which an upward move from s to $s+1$ occurs, i.e. $r_\tau = s$ and $r_{\tau+1} = s+1$ for all $\tau \in \{t_1, \cdots, t_{n^s_s}\}$. We also know there are l^s_{s+1} periods, denoted by $t'_1 < t'_2 < \cdots < t'_{l^s_{s+1}}$, at which a downward move from $s+1$ to s occurs, i.e. $r_\tau = s+1$ and $r_{\tau+1} = s$ for all $\tau \in \{t'_1, \cdots, t'_{l^s_{s+1}}\}$. The following lemma states downward moves from $s+1$ to s and upward moves from s to $s+1$ are coupled: every such downward move must be followed by an immediate upward move, except possibly the very last downward move.

Lemma A.2.3. *We have* $t_1 < t'_1 < t_2 < t'_2 < \cdots < t'_{n_s-1} < t_{n_s}$ *and* $t'_\tau + 1 = t_{\tau+1}$ *for* $\tau = \{1, 2, \cdots, n_s - 1\}$*. Moreover, if we removes all these pairs of downward and upward moves at* $\{t'_\tau, t'_\tau + 1\}_{\tau=1}^{n_s-1}$ *from* h*, the resulting vector*

$$(r_0, \cdots, r_{t'_1}, r_{t'_1+3}, \cdots, r_{t'_2}, r_{t'_2+3}, \cdots, r_{t'_{n_s-1}}, r_{t'_{n_s-1}+3}, \cdots, r_{j^s})$$

is a path in $\Phi^{r,j^{s+1}}(k)(\vec{n}^{s+1}, \vec{l}^{s+1})$*.*

Proof. The inequality is straightforward. Consider any $\tau \in \{1, 2, \cdots, n_s - 1\}$. By construction, $r_{t'_\tau} = s+1$ and $r_{t'_\tau+1} = s$. Because h does not contain horizontal moves, $r_{t'_\tau+2}$ is either $s+1$ or $s-1$. If $s = 0$, then $r_{t'_\tau+2} = 1$ because there is no downward move at 0. Thus at $t'_\tau + 1$, an upward move from s to $s+1$ occurs, which implies $t'_\tau + 1 = t_{\tau+1}$ by definition of $t_{\tau+1}$ and the above inequality. Now consider $s > 0$ and assume $r_{t'_\tau+2} = s - 1$. This implies $t_{\tau+1} > t'_\tau + 2$. Because $r_{t_{\tau+1}} = s$, there must be an upward move from $s-1$ to s between periods $t'_\tau + 2$ and $t_{\tau+1}$. Moveover, because $r_{t_\tau} = s$, there must also be an upward move from $s - 1$ to s at period $t_\tau - 1$. Since $t_\tau - 1 < t'_\tau - 1 < t'_\tau + 2$, we know along the path h, there must be at least two upward moves from $s - 1$ to s, i.e. $n^s_{s-1} \geq 2$. But this contradicts the construction of n^s. Hence $r_{t'_\tau+2} = s + 1$, which implies $t'_\tau + 1 = t_{\tau+1}$. Lastly, if we remove periods t'_τ and $t'_\tau + 1$ from h, then the resulting vector is still a connected path because in h, $r_{t'_\tau} = r_{t'_\tau+2}$. After removing $n_s - 1$ pairs of these downward and upward moves, there will be only 1 upward move from s to $s + 1$. Moreover, the number of upward and downward moves at each other signal remain unchanged. Therefore, the resulting vector is a path in $\Phi^{r,j^{s+1}}(k)(\vec{n}^{s+1}, \vec{l}^{s+1})$ by construction. □

Lemma A.2.3 in fact defines a mapping

$$\gamma_s : \Phi^{r,j^s}(k)(\vec{n}^s, \vec{l}^s) \to \Phi^{r,j^{s+1}}(k)(\vec{n}^{s+1}, \vec{l}^{s+1})$$

such that for every $h \in \Phi^{r,j^s}(k)(\vec{n}^s, \vec{l}^s)$, $\gamma_s(h)$ is the path in $\Phi^{r,j^{s+1}}(k)(\vec{n}^{s+1}, \vec{l}^{s+1})$ after removing $n_s - 1$ pairs of downward move from $s+1$ to s and upward move from s to $s+1$ in h, as in Lemma A.2.3. Let γ_s^{-1} be the pre-image of γ_s. See Figure A.2 for an illustration of γ_s and γ_s^{-1}.

Lemma A.2.4. *Assume* $0 \le s \le r+k-2$. *For all* $h \in \Phi^{r,j^{s+1}}(k)(\vec{n}^{s+1}, \vec{l}^{s+1})$,

$$\#\big(\gamma_s^{-1}(h)\big) = \begin{pmatrix} n_s + l_{s+2} - 1 \\ l_{s+2} \end{pmatrix}.$$

Hence

$$\#\big(\Phi^{r,j^s}(k)(\vec{n}^s, \vec{l}^s)\big) = \begin{pmatrix} n_s + l_{s+2} - 1 \\ l_{s+2} \end{pmatrix} \times \#\big(\Phi^{r,j^{s+1}}(k)(\vec{n}^{s+1}, \vec{l}^{s+1})\big).$$

Proof. Since γ_s removes $n_s - 1$ pairs of downward-upward moves, γ_s^{-1} adds them back. Notice the signal $s+1$ is reached $n_s^{s+1} + l_{s+2}^{s+1} = 1 + l_{s+2}$ times along the path h. Hence the total number of ways to add $n_s - 1$ downward-upward pairs to h is

$$\#\big(\gamma_s^{-1}(h)\big) = \begin{pmatrix} (n_s - 1) + (1 + l_{s+2}) - 1 \\ (1 + l_{s+2}) - 1 \end{pmatrix} = \begin{pmatrix} n_s + l_{s+2} - 1 \\ l_{s+2} \end{pmatrix},$$

by Lemma A.1.1.

Because $\#\big(\gamma_s^{-1}(h)\big)$ is independent of h and

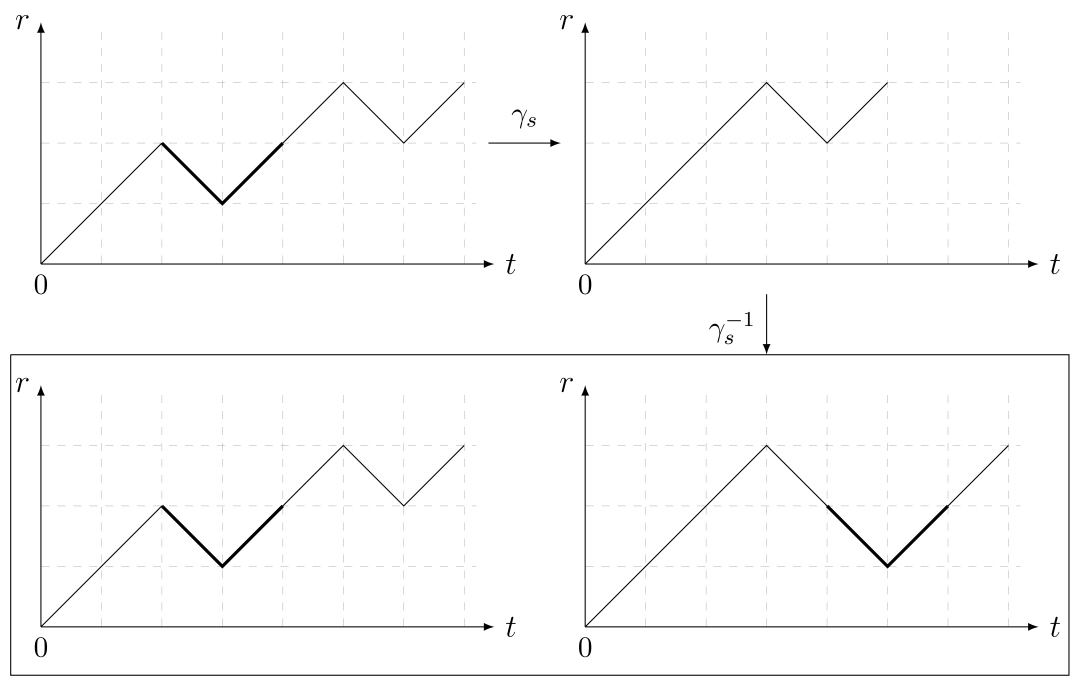
$$\Phi^{r,j^s}(k)(\vec{n}^s, \vec{l}^s) = \bigcup_{h \in \Phi^{r,j^{s+1}}(k)(\vec{n}^{s+1}, \vec{l}^{s+1})} \gamma_s^{-1}(h),$$

we have

$$\begin{aligned} \#\big(\Phi^{r,j^s}(k)(\vec{n}^s, \vec{l}^s)\big) &= \#\big(\gamma_s^{-1}(h)\big) \times \#\big(\Phi^{r,j^{s+1}}(k)(\vec{n}^{s+1}, \vec{l}^{s+1})\big) \\ &= \begin{pmatrix} n_s + l_{s+2} - 1 \\ l_{s+2} \end{pmatrix} \times \#\big(\Phi^{r,j^{s+1}}(k)(\vec{n}^{s+1}, \vec{l}^{s+1})\big). \end{aligned}$$

□

Figure A.2: Illustration of γ_s and γ_s^{-1}.



Assume $r = 3$, $k = 0$, $\vec{n} = (n_0, 2, 2)$, $\vec{l} = (n_0 - 1, 1, 0)$ for any $n_0 \geq 1$. Following the construction of $\{\vec{n}^s, \vec{l}^s, j^s\}$, we have $\vec{n}^2 = (1, 1, 2)$, $\vec{l}^2 = (0, 0, 1)$, $j^2 = 5$ and $\vec{n}^1 = (1, 2, 2)$, $\vec{l}^1 = (0, 1, 0)$, $j^1 = 7$.

The top-left panel illustrates a typical path $h \in \Phi^{r,j^1}(k)(\vec{n}^1, \vec{l}^1)$. The top-right panel is $\gamma_1(h)$ which is obtained by removing the pair of downward move from 2 to 1 and upward move from 1 to 2 (bold line in top-left panel). It is clear from the graph $\gamma_1(h)$ is a path in $\Phi^{r,j^2}(k)(\vec{n}^2, \vec{l}^2)$. The bottom two panels depict all paths in $\gamma_1^{-1}(\gamma_1(h))$ which are obtained by adding a pair of downward move from 2 to 1 and upward move from 1 to 2. Because $\gamma_1(h)$ reaches signal 2 two times at $t = 2$ and $t = 4$ respectively ($n_1^2 + l_3^2 = 2$), there are two ways to add this pair. The bottom left panel depicts the path if this pair is added to h at $t = 2$, which is the original h. The bottom right panel depicts the path if this pair is added to h at $t = 4$.

Recall these formula are derived under the assumption $n_s > 1$. Notice these formula also accommodate the case $n_s = 1$ because

$$\begin{pmatrix} 1 + l_{s+2} - 1 \\ l_{s+2} \end{pmatrix} = 1$$

and we have already known $\#\big(\Phi^{r,j^s}(k)(\vec{n}^s, \vec{l}^s)\big) = \#\big(\Phi^{r,j^{s+1}}(k)(\vec{n}^{s+1}, \vec{l}^{s+1})\big)$ when $n_s = 1$.

Now using simple induction yields

$$\begin{aligned}\#\big(\Phi^{r,\bar{j}}(k)(\vec{n}, \vec{l})\big) &= \#\big(\Phi^{r,j^0}(k)(\vec{n}^0, \vec{l}^0)\big) \\ &= \prod_{\tilde{r}=0}^{r+k-2} \begin{pmatrix} n_{\tilde{r}} + l_{\tilde{r}+2} - 1 \\ l_{\tilde{r}+2} \end{pmatrix} \times \#\big(\Phi^{r,j^{r+k-1}}(k)(\vec{n}^{r+k-1}, \vec{l}^{r+k-1})\big).\end{aligned}$$

But it is straightforward to see $\Phi^{r,j^{r+k-1}}(k)(\vec{n}^{r+k-1}, \vec{l}^{r+k-1})$ only contains a single path by construction. Hence

$$\#\big(\Phi^{r,\bar{j}}(k)(\vec{n}, \vec{l})\big) = \prod_{\tilde{r}=0}^{r+k-2} \begin{pmatrix} n_{\tilde{r}} + l_{\tilde{r}+2} - 1 \\ l_{\tilde{r}+2} \end{pmatrix}. \tag{A.11}$$

Combining Equations (A.10) and (A.11) yields

$$\#\Big(\Phi^{r,t}(k)(\vec{n}, \vec{l})(\vec{m})\Big) = \prod_{\tilde{r}=0}^{r+k-2} \begin{pmatrix} n_{\tilde{r}} + l_{\tilde{r}+2} - 1 \\ l_{\tilde{r}+2} \end{pmatrix} \times \prod_{\bar{r}=0}^{r+k} \begin{pmatrix} m_{\bar{r}} + n_{\bar{r}} + l_{\bar{r}} - 1 \\ n_{\bar{r}} + l_{\bar{r}} - 1 \end{pmatrix} \tag{A.12}$$

Step 5: Expression for $P^{\sigma_1}_{\tilde{\xi}}(R_t = r)$.

Combining Equations (A.9) and (A.12) yields $P^{\sigma_1}_{\tilde{\xi}}(R_t = r)$ for $t \geq r$:

$$\begin{aligned}&P^{\sigma_1}_{\tilde{\xi}}\big(R_t = r\big) \\ &= \sum_{k=0}^{\lfloor \frac{t-r}{2} \rfloor} \sum_{(\vec{n},\vec{l}) \in \mathcal{I}^{r,t}(k)} \Bigg[\prod_{\tilde{r}=0}^{r+k-2} \begin{pmatrix} n_{\tilde{r}} + l_{\tilde{r}+2} - 1 \\ l_{\tilde{r}+2} \end{pmatrix} \times \prod_{\hat{r}=0}^{r+k-1} \tilde{x}_{\hat{r}}^{n_{\hat{r}}} \tilde{y}_{\hat{r}+1}^{l_{\hat{r}+1}} \\ &\quad \times \Bigg(\sum_{\vec{m} \in \mathcal{J}^{r,t}(k)(\vec{n},\vec{l})} \prod_{\bar{r}=0}^{r+k} \begin{pmatrix} m_{\bar{r}} + n_{\bar{r}} + l_{\bar{r}} - 1 \\ n_{\bar{r}} + l_{\bar{r}} - 1 \end{pmatrix} \tilde{z}_{\bar{r}}^{m_{\bar{r}}} \Bigg) \Bigg]. \end{aligned} \tag{A.13}$$

As a final remark of this section, we note that though Equation (A.13) is derived under the normal type, it should be clear that the probability of reaching signal r in period t under the commitment type, i.e. $P_{\widehat{\xi}}(\Phi^{r,t})$, has the same form as Equation (A.13) except that $\{\tilde{x}_{\tilde{r}}, \tilde{y}_{\tilde{r}}, \tilde{z}_{\tilde{r}}\}_{\tilde{r}}$ are all replaced by the corresponding transition probabilities $\{\hat{x}_{\tilde{r}}, \hat{y}_{\tilde{r}}, \hat{z}_{\tilde{r}}\}$ under the commitment type.

A.3 Asymptotic behavior of posterior beliefs.

Recall $P_{\widehat{\xi}}$ is the the probability measure over R^∞ induced by the commitment type's strategy. In this section, we analyze the asymptotic behavior of the posterior likelihood ratio $\frac{\sum_{t=r}\delta^t P_{\widetilde{\xi}}^{\sigma_1}(R_t=r)}{\sum_{t=r}\delta^t P_{\widehat{\xi}}(R_t=r)}$.

We start by calculating $\sum_{t=r}\delta^t P_{\widetilde{\xi}}^{\sigma_1}(R_t=r)$. From Equation (A.13), we have

$$
\begin{aligned}
&\sum_{t\geq r}\delta^t P_{\widetilde{\xi}}^{\sigma_1}(R_t=r)\\
=\;&\sum_{t\geq r}\sum_{k=0}^{\lfloor\frac{t-r}{2}\rfloor}\sum_{(\vec{n},\vec{l})\in\mathcal{I}^{r,t}(k)}\delta^t\Bigg[\prod_{\tilde{r}=0}^{r+k-2}\binom{n_{\tilde{r}}+l_{\tilde{r}+2}-1}{l_{\tilde{r}+2}}\times\prod_{\hat{r}=0}^{r+k-1}\tilde{x}_{\hat{r}}^{n_{\hat{r}}}\tilde{y}_{\hat{r}+1}^{l_{\hat{r}+1}}\\
&\times\Bigg(\sum_{\vec{m}\in\mathcal{J}^{r,t}(k)(\vec{n},\vec{l})}\prod_{\bar{r}=0}^{r+k}\binom{m_{\bar{r}}+n_{\bar{r}}+l_{\bar{r}}-1}{n_{\bar{r}}+l_{\bar{r}}-1}\tilde{z}_{\bar{r}}^{m_{\bar{r}}}\Bigg)\Bigg]\\
=\;&\sum_{k=0}^{\infty}\sum_{t=r+2k}^{\infty}\sum_{(\vec{n},\vec{l})\in\mathcal{I}^{r,t}(k)}\Bigg[\prod_{\tilde{r}=0}^{r+k-2}\binom{n_{\tilde{r}}+l_{\tilde{r}+2}-1}{l_{\tilde{r}+2}}\times\prod_{\hat{r}=0}^{r+k-1}(\delta\tilde{x}_{\hat{r}})^{n_{\hat{r}}}(\delta\tilde{y}_{\hat{r}+1})^{l_{\hat{r}+1}}\\
&\times\Bigg(\sum_{\vec{m}\in\mathcal{J}^{r,t}(k)(\vec{n},\vec{l})}\prod_{\bar{r}=0}^{r+k}\binom{m_{\bar{r}}+n_{\bar{r}}+l_{\bar{r}}-1}{n_{\bar{r}}+l_{\bar{r}}-1}(\delta\tilde{z}_{\bar{r}})^{m_{\bar{r}}}\Bigg)\Bigg],
\end{aligned}
$$

where the second equality comes from i) interchanging the order of the first two summations in the first equality, and ii) the fact that for all $0\leq k\leq\lfloor\frac{t-r}{2}\rfloor$, $(\vec{n},\vec{l})\in\mathcal{I}^{r,t}(k)$ and $\vec{m}\in\mathcal{J}^{r,t}(k)(\vec{n},\vec{l})$, $\sum_{\bar{r}=0}^{r+k}(n_{\bar{r}}+l_{\bar{r}}+m_{\bar{r}})=t$ (see the definition of $\mathcal{J}^{r,t}(k)(\vec{n},\vec{l})$ in (A.8)). Using the definition of $\mathcal{I}^{r,t}(k)$ in (A.6), we can also interchange the second and the third summations in the above expression:

$$
\begin{aligned}
&\sum_{t\geq r}\delta^t P^{\sigma_1}_{\tilde{\xi}}(R_t=r)\\
=\;&\sum_{k=0}^{\infty}\sum_{(\vec{n},\vec{l})\in\bigcup_{\tau\geq r+2k}\mathcal{I}^{r,\tau}(k)}\left[\prod_{\tilde{r}=0}^{r+k-2}\begin{pmatrix} n_{\tilde{r}}+l_{\tilde{r}+2}-1\\ l_{\tilde{r}+2}\end{pmatrix}\times\prod_{\hat{r}=0}^{r+k-1}(\delta\tilde{x}_{\hat{r}})^{n_{\hat{r}}}(\delta\tilde{y}_{\hat{r}+1})^{l_{\hat{r}+1}}\right.\\
&\left.\times\Bigg(\sum_{t=\sum_{\underline{r}=0}^{r+k}(n_{\underline{r}}+l_{\underline{r}})}^{\infty}\;\sum_{\vec{m}\in\mathcal{J}^{r,t}(k)(\vec{n},\vec{l})}\prod_{\bar{r}=0}^{r+k}\begin{pmatrix} m_{\bar{r}}+n_{\bar{r}}+l_{\bar{r}}-1\\ n_{\bar{r}}+l_{\bar{r}}-1\end{pmatrix}(\delta\tilde{z}_{\bar{r}})^{m_{\bar{r}}}\Bigg)\right]. \qquad \text{(A.14)}
\end{aligned}
$$

The term in the last line of Equation (A.14) can then be simplified:

$$
\begin{aligned}
&\sum_{t=\sum_{\underline{r}}^{r+k}(n_{\underline{r}}+l_{\underline{r}})}^{\infty}\;\sum_{\vec{m}\in\mathcal{J}^{r,t}(k)(\vec{n},\vec{l})}\prod_{\bar{r}=0}^{r+k}\begin{pmatrix} m_{\bar{r}}+n_{\bar{r}}+l_{\bar{r}}-1\\ n_{\bar{r}}+l_{\bar{r}}-1\end{pmatrix}(\delta\tilde{z}_{\bar{r}})^{m_{\bar{r}}}\\
=\;&\sum_{t=\sum_{\underline{r}=0}^{r+k}(n_{\underline{r}}+l_{\underline{r}})}^{\infty}\;\sum_{\substack{\sum_{\bar{r}=0}^{r+k}m_{\bar{r}}=t-\sum_{\underline{r}=0}^{r+k}(n_{\underline{r}}+l_{\underline{r}})\\ m_{\bar{r}}\geq 0\ \forall\bar{r}}}\prod_{\bar{r}=0}^{r+k}\begin{pmatrix} m_{\bar{r}}+n_{\bar{r}}+l_{\bar{r}}-1\\ n_{\bar{r}}+l_{\bar{r}}-1\end{pmatrix}(\delta\tilde{z}_{\bar{r}})^{m_{\bar{r}}}\\
=\;&\sum_{s=0}^{\infty}\sum_{\substack{\sum_{\bar{r}=0}^{r+k}m_{\bar{r}}=s\\ m_{\bar{r}}\geq 0\ \forall\bar{r}}}\prod_{\bar{r}=0}^{r+k}\begin{pmatrix} m_{\bar{r}}+n_{\bar{r}}+l_{\bar{r}}-1\\ n_{\bar{r}}+l_{\bar{r}}-1\end{pmatrix}(\delta\tilde{z}_{\bar{r}})^{m_{\bar{r}}}\\
=\;&\prod_{\bar{r}=0}^{r+k}\Bigg(\sum_{m_{\bar{r}}=0}^{\infty}\begin{pmatrix} m_{\bar{r}}+n_{\bar{r}}+l_{\bar{r}}-1\\ n_{\bar{r}}+l_{\bar{r}}-1\end{pmatrix}(\delta\tilde{z}_{\bar{r}})^{m_{\bar{r}}}\Bigg)\\
=\;&\prod_{\bar{r}=0}^{r+k}\frac{1}{(1-\delta\tilde{z}_{\bar{r}})^{n_{\bar{r}}+l_{\bar{r}}}}, \qquad \text{(A.15)}
\end{aligned}
$$

where the first equality comes from definition of $\mathcal{J}^{r,t}(k)(\vec{n},\vec{l})$ in (A.8), the second from change of variable $s\equiv t-\sum(n_{\underline{r}}+l_{\underline{r}})$, third and forth from Lemma A.1.2. Plugging (A.15) into (A.14) and rearranging yield

$$
\begin{aligned}
&\sum_{t\geq r}\delta^t P^{\sigma_1}_{\tilde{\xi}}(R_t=r)\\
=\;&\sum_{k=0}^{\infty}\sum_{(\vec{n},\vec{l})\in\bigcup_{\tau\geq r+2k}\mathcal{I}^{r,\tau}(k)}\left[\prod_{\tilde{r}=0}^{r+k-2}\begin{pmatrix} n_{\tilde{r}}+l_{\tilde{r}+2}-1\\ l_{\tilde{r}+2}\end{pmatrix}\times\prod_{\hat{r}=0}^{r+k-1}\tilde{\theta}_{\hat{r}}^{n_{\hat{r}}}\tilde{\phi}_{\hat{r}+1}^{l_{\hat{r}+1}}\right], \qquad \text{(A.16)}
\end{aligned}
$$

where

$$\tilde{\theta}_{\tilde{r}} \equiv \frac{\delta \tilde{x}_{\tilde{r}}}{(1-\delta \tilde{z}_{\tilde{r}})} \quad \forall \tilde{r} \geq 0, \text{ and} \tag{A.17}$$

$$\tilde{\phi}_{\tilde{r}} \equiv \frac{\delta \tilde{y}_{\tilde{r}}}{(1-\delta \tilde{z}_{\tilde{r}})} \quad \forall \tilde{r} \geq 1. \tag{A.18}$$

Notice for fixed r and k, $(\vec{n},\vec{l}) \in \bigcup_{\tau \geq r+2k} \mathcal{I}^{r,\tau}(k)$ if and only if i) $n_{\tilde{r}} \geq 1$, for $0 \leq \tilde{r} \leq r+k-1$, and ii) $l_{\tilde{r}+1} = n_{\tilde{r}} - 1$ if $\tilde{r} \leq r-1$ and $l_{\tilde{r}+1} = n_{\tilde{r}}$ if $\tilde{r} \geq r$. Thus A.16 can then be written as

$$\begin{aligned}
&\sum_{t \geq r} \delta^t P^{\sigma_1}_{\tilde{\xi}}(R_t = r) \\
= & \sum_{k=0}^{\infty} \sum_{n_0,\cdots,n_{r+k-1} \geq 1} \left[\prod_{\tilde{r}=0}^{r-1} \tilde{\theta}_{\tilde{r}} \times \prod_{\hat{r}=0}^{r-2} \binom{(n_{\hat{r}}-1)+(n_{\hat{r}+1}-1)}{(n_{\hat{r}+1}-1)} (\tilde{\theta}_{\hat{r}} \tilde{\phi}_{\hat{r}+1})^{n_{\hat{r}}-1} \right. \\
& \left. \times \binom{(n_{r-1}-1)+n_r}{n_r} (\tilde{\theta}_{r-1}\tilde{\phi}_r)^{n_{r-1}-1} \times \prod_{\bar{r}=r}^{r+k-1} \binom{n_{\bar{r}}+n_{\bar{r}+1}-1}{n_{\bar{r}+1}} (\tilde{\theta}_{\bar{r}}\tilde{\phi}_{\bar{r}+1})^{n_{\bar{r}}} \right] \\
= & \sum_{k=0}^{\infty} \sum_{\substack{\hat{n}_r,\hat{n}_{r+1},\cdots,\hat{n}_{r+k-1} \geq 1 \\ \hat{n}_{r+k}=0}} \left[\prod_{\bar{r}=r}^{r+k-1} \binom{\hat{n}_{\bar{r}}+\hat{n}_{\bar{r}+1}-1}{\hat{n}_{\bar{r}+1}} (\tilde{\theta}_{\bar{r}}\tilde{\phi}_{\bar{r}+1})^{\hat{n}_{\bar{r}}} \right. \\
& \left. \times \Big(\prod_{\tilde{r}=0}^{r-1} \tilde{\theta}_{\tilde{r}} \times \sum_{\hat{n}_0,\cdots,\hat{n}_{r-1} \geq 0} \prod_{\hat{r}=0}^{r-1} \binom{\hat{n}_{\hat{r}}+\hat{n}_{\hat{r}+1}}{\hat{n}_{\hat{r}+1}} (\tilde{\theta}_{\hat{r}}\tilde{\phi}_{\hat{r}+1})^{\hat{n}_{\hat{r}}} \Big) \right],
\end{aligned} \tag{A.19}$$

where the second equality comes from change of variable $\hat{n}_{\tilde{r}} = n_{\tilde{r}} - 1$ if $0 \leq \tilde{r} \leq r-1$ and $\hat{n}_{\tilde{r}} = n_{\tilde{r}}$ when $\tilde{r} \geq r$. In order to simplify (A.19) further, define $\tilde{a}_0 \equiv \tilde{\theta}_0$ and inductively

$$\tilde{a}_{\tilde{r}} \equiv \frac{\tilde{\theta}_{\tilde{r}}}{1-\tilde{a}_{\tilde{r}-1}\tilde{\phi}_{\tilde{r}}} \tag{A.20}$$

for $\tilde{r} \geq 1$.

Lemma A.3.1. *For all $\tilde{r} \geq 0$, $\tilde{a}_{\tilde{r}} \leq \delta < 1$. Then we have for all $r \geq 1$,*

$$\Big(\prod_{\tilde{r}=0}^{r-1} \tilde{\theta}_{\tilde{r}} \times \sum_{\hat{n}_0,\cdots,\hat{n}_{r-1}\geq 0} \prod_{\hat{r}=0}^{r-1} \begin{pmatrix} \hat{n}_{\hat{r}} + \hat{n}_{\hat{r}+1} \\ \hat{n}_{\hat{r}+1} \end{pmatrix} (\tilde{\theta}_{\hat{r}}\tilde{\phi}_{\hat{r}+1})^{\hat{n}_{\hat{r}}}\Big)$$

$$= \prod_{\tilde{r}=0}^{r-1} \tilde{a}_{\tilde{r}} \times \Big(\frac{1}{1-\tilde{a}_{r-1}\tilde{\phi}_r}\Big)^{\hat{n}_r+1}. \tag{A.21}$$

Proof. From definitions of $\tilde{x}_{\tilde{r}}, \tilde{y}_{\tilde{r}}, \tilde{z}_{\tilde{r}}, \tilde{\theta}_{\tilde{r}}, \tilde{\phi}_{\tilde{r}}$ in (A.1)-(A.5) and (A.17)-(A.18), it is easy to see $\tilde{\theta}_0 \leq \delta$ and $\tilde{\theta}_{\tilde{r}} + \tilde{\phi}_{\tilde{r}} \leq \delta$ for $\tilde{r} \geq 1$. By construction, $\tilde{a}_0 = \tilde{\theta}_0 \leq \delta$. Now if $\tilde{a}_{\tilde{r}} \leq \delta$, then we have

$$\tilde{a}_{\tilde{r}+1} = \frac{\tilde{\theta}_{\tilde{r}+1}}{1-\tilde{a}_{\tilde{r}}\tilde{\phi}_{\tilde{r}+1}} < \frac{\tilde{\theta}_{\tilde{r}+1}}{1-\tilde{\phi}_{\tilde{r}+1}} \leq \delta.$$

Based on this, we now prove (A.21) by induction on r. When $r = 1$, we have

$$\theta_0 \times \sum_{n_0\geq 0} \begin{pmatrix} \hat{n}_0 + \hat{n}_1 \\ \hat{n}_1 \end{pmatrix} (\tilde{\theta}_0\tilde{\phi}_1)^{\hat{n}_0} = \theta_0 \times \Big(\frac{1}{1-\tilde{\theta}_0\tilde{\phi}_1}\Big)^{\hat{n}_1+1} = \tilde{a}_0 \times \Big(\frac{1}{1-\tilde{a}_0\tilde{\phi}_1}\Big)^{\hat{n}_1+1},$$

where the first equality comes from Lemma A.1.2. Hence (A.21) holds when $r = 1$.

Assume (A.21) holds for $r = r'$. When $r = r' + 1$, we have

$$
\begin{aligned}
& \Big(\prod_{\tilde r=0}^{r'} \tilde\theta_{\tilde r} \times \sum_{\hat n_0,\cdots,\hat n_{r'}\geq 0} \prod_{\hat r=0}^{r'} \begin{pmatrix} \hat n_{\hat r} + \hat n_{\hat r+1} \\ \hat n_{\hat r+1} \end{pmatrix} (\tilde\theta_{\hat r}\tilde\phi_{\hat r+1})^{\hat n_{\hat r}}\Big) \\
= \;& \tilde\theta_{r'} \times \sum_{\hat n_{r'}\geq 0} \Bigg[\begin{pmatrix} \hat n_{r'} + \hat n_{r'+1} \\ \hat n_{r'+1} \end{pmatrix} (\tilde\theta_{r'}\tilde\phi_{r'+1})^{\hat n_{r'}} \\
& \qquad \times \Big(\prod_{\tilde r=0}^{r'-1} \tilde\theta_{\tilde r} \times \sum_{\hat n_0,\cdots,\hat n_{r'-1}\geq 0} \prod_{\hat r=0}^{r'-1} \begin{pmatrix} \hat n_{\hat r} + \hat n_{\hat r+1} \\ \hat n_{\hat r+1} \end{pmatrix} (\tilde\theta_{\hat r}\tilde\phi_{\hat r+1})^{\hat n_{\hat r}}\Big)\Bigg] \\
= \;& \tilde\theta_{r'} \times \sum_{\hat n_{r'}\geq 0} \Bigg[\begin{pmatrix} \hat n_{r'} + \hat n_{r'+1} \\ \hat n_{r'+1} \end{pmatrix} (\tilde\theta_{r'}\tilde\phi_{r'+1})^{\hat n_{r'}} \times \prod_{\tilde r=0}^{r'-1} \tilde a_{\tilde r} \times \Big(\frac{1}{1-\tilde a_{r'-1}\phi_{r'}}\Big)^{n_{r'}+1}\Bigg] \\
= \;& \prod_{\tilde r=0}^{r'-1} \tilde a_{\tilde r} \times \frac{\tilde\theta_{r'}}{1-\tilde a_{r'-1}\phi_{r'}} \times \sum_{\hat n_{r'}\geq 0} \begin{pmatrix} \hat n_{r'} + \hat n_{r'+1} \\ \hat n_{r'+1} \end{pmatrix} \Big(\frac{\tilde\theta_{r'}\tilde\phi_{r'+1}}{1-\tilde a_{r'-1}\tilde\phi_{r'}}\Big)^{\hat n_{r'}} \\
= \;& \prod_{\tilde r=0}^{r'} \tilde a_{\tilde r} \times \Big(\frac{1}{1-\frac{\tilde\theta_{r'}\tilde\phi_{r'+1}}{1-\tilde a_{r'-1}\tilde\phi_{r'}}}\Big)^{\hat n_{r'+1}+1} \\
= \;& \prod_{\tilde r=0}^{r'} \tilde a_{\tilde r} \times \Big(\frac{1}{1-\tilde a_{r'}\tilde\phi_{r'+1}}\Big)^{\hat n_{r'+1}+1},
\end{aligned}
$$

where the second equality comes from induction hypothesis and the third comes from Lemma A.1.2. Hence the desired equation follows. □

Combining (A.19) and Lemma A.3.1, we finally have

$$
\sum_{t=r}^{\infty} \delta^t P_{\tilde\xi}^{\sigma_1}(R_t = r) = \prod_{\tilde r=0}^{r-1} \tilde a_{\tilde r} \times M^{\sigma_1}(r) \tag{A.22}
$$

where

$$
M^{\sigma_1}(r) \equiv
$$

$$
\sum_{k=0}^{\infty} \sum_{\substack{\hat n_r,\hat n_{r+1},\cdots,\hat n_{r+k-1}\geq 1 \\ \hat n_{r+k}=0}} \Big(\frac{1}{1-\tilde a_{r-1}\tilde\phi_r}\Big)^{\hat n_r+1} \times \Bigg[\prod_{\bar r=r}^{r+k-1} \begin{pmatrix} \hat n_{\bar r} + \hat n_{\bar r+1} - 1 \\ \hat n_{\bar r+1} \end{pmatrix} (\tilde\theta_{\bar r}\tilde\phi_{\bar r+1})^{\hat n_{\bar r}}\Bigg].
$$

It is important to notice that (A.22) is derived for arbitrary σ_1. In particular, it also holds for the commitment type's strategy $\widehat{\sigma}_1$ with $\widehat{\sigma}_1(r) = 1$ for all r. Let $\{\hat{x}_{\tilde{r}}, \hat{y}_{\tilde{r}}, \hat{z}_{\tilde{r}}, \hat{\theta}_{\tilde{r}}, \hat{\phi}_{\tilde{r}}, \hat{a}_{\tilde{r}}\}$ be the corresponding values in (A.1)-(A.5), (A.17)-(A.18) and (A.21) when $\widehat{\sigma}_1$ is used. Then we have

$$\sum_{t\geq r} \delta^t P_{\widehat{\xi}}(R_t = r) = \prod_{\tilde{r}=0}^{r-1} \hat{a}_{\tilde{r}} \times M^{\widehat{\sigma}_1}(r),$$

and consequently for any strategy σ_1 of the normal type

$$\frac{\sum_{t=r}^{\infty} \delta^t P_{\widetilde{\xi}}^{\sigma_1}(R_t = r)}{\sum_{t\geq r} \delta^t P_{\widehat{\xi}}(R_t = r)} = \prod_{\tilde{r}=0}^{r-1} \frac{\tilde{a}_{\tilde{r}}}{\hat{a}_{\tilde{r}}} \times \frac{M^{\sigma_1}(r)}{M_1^{\widehat{\sigma}}(r)}.$$

In the remaining of this section we show if the normal seller exerts low effort at infinitely many signals, i.e. $\#\{r|\sigma_1(r) = 0\} = \infty$, then $\lim_{r\to\infty} \sum_{t=r}^{\infty} \delta^t P_{\widetilde{\xi}}^{\sigma_1}(R_t = r)/\sum_{t\geq r} \delta^t P_{\widehat{\xi}}(R_t = r) = 0$. This is accomplished in two steps. First we will show $\lim_{r\to\infty} \prod_{\tilde{r}=0}^{r-1}(\tilde{a}_{\tilde{r}}/\hat{a}_{\tilde{r}}) = 0$. Second, we will show $1 \leq \sup_{\sigma_1,r} M^{\sigma_1}(r) < \infty$. The combination of these two will yield the desired result.

Define two functions: $\theta, \phi : [0,1] \to [0,1]$ such that

$$\theta(\varsigma) \equiv \frac{\delta(1-\varepsilon)\rho(h|\varsigma)}{1-\delta+\delta(1-\varepsilon)\rho(h|\varsigma)+\delta\varepsilon\rho(l|\varsigma)}, \tag{A.23}$$

$$\text{and} \quad \phi(\varsigma) = \frac{\delta\varepsilon\rho(l|\varsigma)}{1-\delta+\delta(1-\varepsilon)\rho(h|\varsigma)+\delta\varepsilon\rho(l|\varsigma)}. \tag{A.24}$$

Notice given σ_1, $\tilde{\theta}_{\tilde{r}}$ and $\tilde{\phi}_{\tilde{r}}$ defined in (A.17) and (A.18) are equal to $\theta\big(\sigma_1(\tilde{r})\big)$ and $\phi\big(\sigma_1(\tilde{r})\big)$ respectively. Similarly $\hat{\theta}_{\tilde{r}} = \theta\big(\widehat{\sigma}_1(\tilde{r})\big)$ and $\hat{\phi}_{\tilde{r}} = \phi\big(\widehat{\sigma}_1(\tilde{r})\big)$.

Lemma A.3.2. *1. The function θ is strictly increasing and function ϕ is strictly decreasing. Moreover for all $\varsigma \in [0,1]$, $\theta(\varsigma)\phi(\varsigma) < 1/4$.*

2. For all $a \in [0,1]$, the function $\varsigma \mapsto \theta(\varsigma)/(1-a\phi(\varsigma))$ is strictly increasing and the function $\varsigma \mapsto \phi(\varsigma)/(1-a\theta(\varsigma))$ is strictly decreasing.

Proof. These properties be directly verified from the definition of θ and ϕ. □

Lemma A.3.3. *Assume σ_1 exerts low effort at infinitely many signals, i.e. $\#\{r \in R|\sigma_1(r) = 0\} = \infty$, then*

$$\lim_{r\to\infty}\prod_{\tilde{r}=0}^{r-1}(\tilde{a}_{\tilde{r}}/\hat{a}_{\tilde{r}}) = 0.$$

Proof. First we show for arbitrary σ_1, $\tilde{a}_{\tilde{r}} \leq \hat{a}_{\tilde{r}}$ for all $\tilde{r} \geq 0$. By definition $\tilde{a}_0 = \tilde{\theta}_0 = \theta(\sigma_1(0)) \leq \theta(\widehat{\sigma}_1(0)) = \hat{\theta}_0 = \hat{a}_0$. Assume $\tilde{a}_{\tilde{r}} \leq \hat{a}_{\tilde{r}}$ for some $\tilde{r} \geq 0$. For $\tilde{r}+1$, we have

$$\begin{aligned}\tilde{a}_{\tilde{r}+1} = \frac{\tilde{\theta}_{\tilde{r}+1}}{1-\tilde{a}_{\tilde{r}}\tilde{\phi}_{\tilde{r}+1}} \leq \frac{\tilde{\theta}_{\tilde{r}+1}}{1-\hat{a}_{\tilde{r}}\tilde{\phi}_{\tilde{r}+1}} = \frac{\theta\big(\sigma_1(\tilde{r}+1)\big)}{1-\hat{a}_{\tilde{r}}\phi\big(\widehat{\sigma}_1(\tilde{r}+1)\big)} \\ \leq \frac{\theta\big(\widehat{\sigma}_1(\tilde{r}+1)\big)}{1-\hat{a}_{\tilde{r}}\phi\big(\widehat{\sigma}_1(\tilde{r}+1)\big)} = \frac{\hat{\theta}_{\tilde{r}+1}}{1-\hat{a}_{\tilde{r}}\hat{\phi}_{\tilde{r}+1}} = \hat{a}_{\tilde{r}+1},\end{aligned}$$

where the first inequality comes from induction hypothesis $\tilde{a}_{\tilde{r}} \leq \hat{a}_{\tilde{r}}$ and the second from Lemma A.3.2.

In addition, if σ_1 exerts low effort at some signal $\tilde{r}+1$. Then the above inequality implies

$$\tilde{a}_{\tilde{r}+1} \leq \frac{\theta\big(0\big)}{1-\hat{a}_{\tilde{r}}\phi\big(0\big)} < \frac{\theta\big(1\big)}{1-\hat{a}_{\tilde{r}}\phi\big(1\big)} = \hat{a}_{\tilde{r}+1},$$

or equivalently

$$\frac{\tilde{a}_{\tilde{r}+1}}{\hat{a}_{\tilde{r}+1}} \leq \frac{\theta(0)}{\theta(1)}\frac{1-\hat{a}_{\tilde{r}}\phi(1)}{1-\hat{a}_{\tilde{r}+1}\phi(0)} \leq \frac{\theta(0)}{\theta(1)}\frac{1-\delta\phi(1)}{1-\delta\phi(0)} < 1$$

where the second inequality comes from Lemma A.3.1 and strict inequality comes from Lemma A.3.2.

Now assume σ_1 exerts low effort at infinitely many signals. Let $L(r) \equiv \#\{0 \leq \tilde{r} \leq r|\sigma_1(r) = 0\}$. Then the above analysis implies

$$\lim_{r\to\infty}\prod_{\tilde{r}=0}^{r-1}\frac{\tilde{a}_{\tilde{r}}}{\hat{a}_{\tilde{r}}} \leq \lim_{r\to\infty}\left(\frac{\theta(0)}{\theta(1)}\frac{1-\delta\phi(1)}{1-\delta\phi(0)}\right)^{L(r-1)} = 0.$$

□

The next lemma shows the $M^{\sigma_1}(r)$ is uniformly bounded across all σ_1 and r.

Lemma A.3.4. *There exists $M > 0$ such that for all $r \geq 1$ and all strategy σ_1*

$$1 \leq M^{\sigma_1}(r) \leq M.$$

Proof. Fix a strategy σ_1 and r. Define for $f_0^{\sigma_1,r} = 1$ and $k \geq 1$,

$$f_k^{\sigma_1,r} \equiv \sum_{\substack{n_r,n_{r+1},\cdots,n_{r+k-1}\geq 1\\ n_{r+k}=0}} \left(\frac{1}{1-\tilde{a}_{r-1}\tilde{\phi}_r}\right)^{n_r} \times \left[\prod_{\bar{r}=r}^{r+k-1} \begin{pmatrix} n_{\bar{r}} + n_{\bar{r}+1} - 1 \\ n_{\bar{r}+1} \end{pmatrix} (\tilde{\theta}_{\bar{r}}\tilde{\phi}_{\bar{r}+1})^{n_{\bar{r}}}\right].$$

Then $M^{\sigma_1}(r) = \sum_{k\geq 0} f_k^{\sigma_1,r}$. The lower bound is obvious. We now show the existence of upper bound. Using a similar argument as in Lemma A.3.1, we can show that for $k \geq 1$,

$$f_k^{\sigma_1,r} = \prod_{\tilde{r}=r}^{r+k-1} \frac{\tilde{a}_{\tilde{r}}\phi_{\tilde{r}+1}}{1-\tilde{a}_{\tilde{r}}\phi_{\tilde{r}+1}}.$$

For any $K \geq 1$, applying Lemma A.1.3 to $\sum_{k=0}^{K} f_k^{\sigma_1,r}$ and using the recursive formulation of $\{\tilde{a}_{\tilde{r}}\}$ in (A.20) yield

$$\sum_{k=0}^{K} f_k^{\sigma_1,r} = \cfrac{1}{1-\cfrac{\tilde{a}_r\tilde{\phi}_{r+1}}{1-\cfrac{\tilde{\theta}_{r+1}\tilde{\phi}_{r+2}}{1-\cfrac{\tilde{\theta}_{r+2}\tilde{\phi}_{r+3}}{\cfrac{\ddots}{1-\cfrac{\tilde{\theta}_{r+K-1}\tilde{\phi}_{r+K-1}}{1-\tilde{\theta}_{r+K-1}\tilde{\phi}_{r+K}}}}}}}.$$

Define $\tilde{b}_1 = \tilde{\phi}_{r+K}$, $b_1 = \phi(0)$ where $\phi(\cdot)$ is defined in (A.24), and inductively

$$\tilde{b}_l = \frac{\tilde{\phi}_{r+K+1-l}}{1-\tilde{b}_{l-1}\tilde{\theta}_{r+K+1-l}} \quad \text{and} \quad b_l = \frac{\phi(0)}{1-b_{l-1}\theta(0)} \qquad l = 2,\cdots,K.$$

Notice that by construction, we have

$$\sum_{k=0}^{K} f_k^{\sigma_1,r} = \frac{1}{1-\tilde{a}_r\tilde{b}_K}.$$

Because $\tilde{a}_r \le \delta < 1$ by Lemma A.21, we know

$$\sum_{k=0}^{K} f_k^{\sigma_1, r} \le \frac{1}{1 - \tilde{b}_K}.$$

Using the fact that the mapping $\varsigma \mapsto \frac{\phi(\varsigma)}{1 - a\theta(\varsigma)}$ decreases for all $a \in [0,1]$ by Lemma A.3.2 and a similar argument as in Lemma A.3.1, we can show $\tilde{b}_K \le b_K$, which implies

$$\sum_{k=0}^{K} f_k^{\sigma_1, r} \le \frac{1}{1 - b_K}.$$

Notice $\{b_k\}$ is independent of σ_1 and r. Hence to show the existence of upper bound, it suffices to show $\sup_{K \ge 1} 1/(1 - b_K) < \infty$ or equivalently $\sup_{K \ge 1} b_K < 1$.

We show this by induction. Clearly $b_0 < 1$. Assume $b_K < 1$ then

$$b_{K+1} = \frac{\phi(0)}{1 - b_K \theta(0)} \le \frac{\phi(0)}{1 - \theta(0)} < 1,$$

where the last inequality comes from the definition of θ and ϕ in (A.23) and (A.18). Therefore $\sup_{K \ge 1} b_K \le \frac{\phi(0)}{1 - \theta(0)} < 1$, completing the proof. □

A.4 Proof of Lemma 1.3.1

Let $\mathscr{S}_1 = \mathscr{S}_2 = \mathscr{B} = [0,1]^R$ be the strategy space of the seller, the strategy space of the buyers and the set of all possible posterior beliefs, respectively. Endow them with the usual product topology. For each $\sigma_1 \in \mathscr{S}_1$, recall $\nu^{\sigma_1} = (\nu^{\sigma_1}(0), \nu^{\sigma_1}(1), \cdots) \in \mathscr{B}$ is the associated posterior beliefs.

Lemma A.4.1. *For any t and r, the mapping $\sigma_1 \mapsto P_{\tilde{\xi}}^{\sigma_1}(R_t = r)$ from $\mathscr{S}_1$ to $[0,1]$ is continuous.*

Proof. If $t < r$, then $P_{\tilde{\xi}}^{\sigma_1}(R_t = r) = 0$ for all σ_1 and the result follows.

Now fix $t \ge r$. From Equation (A.13), we know for any σ_1, the value $P_{\tilde{\xi}}^{\sigma_1}(R_t = r)$ only depends on $\sigma_1(0), \cdots, \sigma_1(r + \lfloor \frac{t-r}{2} \rfloor)$. Hence the mapping $\sigma_1 \mapsto P_{\tilde{\xi}}^{\sigma_1}(R_t = r)$ can be viewed as a mapping from $[0,1]^{r+\lfloor (t-r)/2 \rfloor}$ to $[0,1]$ and we know it is continuous from Equation (A.13) again. Therefore the mapping $\sigma_1 \mapsto P^{\sigma_1}_{\widetilde{\xi}}(R_t = r)$ from $\mathscr{S}_1$ to $[0,1]$ is continuous. $\square$

Lemma A.4.2. *For any $r \geq 0$, the mapping $\sigma_1 \mapsto \nu^{\sigma_1}(r)$ from $\mathscr{S}_1$ to $[0,1]$ is continuous.*

Proof. Fix $r \geq 0$. Notice for all $\sigma_1 \in \mathscr{S}_1$,

$$\frac{1-\nu^{\sigma_1}(r)}{\nu^{\sigma_1}(r)} = \frac{1-b_0}{b_0} \frac{\sum_{t=r} \delta^t P^{\sigma_1}_{\widetilde{\xi}}(R_t = r)}{\sum_{t=r} \delta^t P_{\widehat{\xi}}(R_t = r)},$$

and the denominator of the right hand side is independent of σ_1. Hence to show the continuity of $\sigma_1 \mapsto \nu^{\sigma_1}(r)$, it suffices to show $\sigma_1 \mapsto \sum_{t \geq r} \delta^t P^{\sigma_1}_{\widetilde{\xi}}(R_t = r)$ is continuous. Because for every $t \geq r$, the mapping $\sigma_1 \mapsto P^{\sigma_1}_{\widetilde{\xi}}(R_t = r)$ is continuous by previous lemma and $\sup_{\sigma_1, t} P^{\sigma_1}_{\widetilde{\xi}}(R_t = r) \leq 1$, the discounted sum of all these mappings is also continuous, i.e. $\sigma_1 \mapsto \sum_{t \geq r} \delta^t P^{\sigma_1}_{\widetilde{\xi}}(R_t = r)$ is continuous. $\square$

Proof of Lemma 1.3.1. We first consider buyers' problem. For each $\sigma_1 \in \mathscr{S}_1$, let

$$Br_2(\sigma_1) \equiv \left\{ \sigma_2 \in \mathscr{S}_2 \middle| \sigma_2(r) \in \operatorname*{arg\,max}_{\varsigma_2 \in [0,1]} u_b\Big(\nu^{\sigma_1}(r) + (1-\nu^{\sigma_1}(r))\sigma_1(\widetilde{\xi}, r), \varsigma_2\Big),\ \forall r \geq 0 \right\}$$

be the set of best responses of the buyers. It is easy to see the correspondence $Br_2 : \mathscr{S}_1 \rightrightarrows \mathscr{S}_2$ is nonempty and convex valued. Because of Lemma A.4.2, for each r the mapping $(\sigma_1, \varsigma_2) \mapsto u_b\Big(\nu^{\sigma_1}(r) + (1-\nu^{\sigma_1}(r))\sigma_1(\widetilde{\xi}, r), \varsigma_2\Big)$ is continuous. Therefore the correspondence Br_2 is upper hemicontinuous.

Now we turn to seller's problem. Let $B(R \times \mathscr{S}_2)$ be the space of all bounded continuous functions from $R \times \mathscr{S}_2$ to $\mathbb{R}$, endowed with the supremum norm.[4] Define $f : [0,1] \times R \times \mathscr{S}_2 \to \mathbb{R}$ as $f(\varsigma_1, r, \sigma_2) = (1-\delta) u_s(\varsigma_1, \sigma_2(r))$. Clearly f is continuous.

[4] We endow R with the discrete topology.

Let $T : B(R \times \mathscr{S}_2) \to B(R \times \mathscr{S}_2)$ be the contraction mapping:

$$\begin{aligned} TV(r, \sigma_2) &= \max_{\varsigma_1 \in [0,1]} f(\varsigma_1, r, \sigma_2) + \delta \Big[(1-\varepsilon)\rho(h|\varsigma_1) V(r+1, \sigma_2) \\ &\qquad + \big(1 - (1-\varepsilon)\rho(h|\varsigma_1) - \varepsilon\rho(l|\varsigma_1)\big) V(r, \sigma_2) \\ &\qquad + \varepsilon\rho(l|\varsigma_1) V(\max(r-1, 0), \sigma_2) \Big]. \end{aligned}$$

Let V^* be the unique fixed point of T and $\Gamma : R \times \mathscr{S}_2 \rightrightarrows [0,1]$ be the associated policy correspondence. Clearly $V^*(\,\cdot\,, \sigma_2)$ is the value function to the seller when he plays a best response. It is also easy to see a strategy σ_1 is a best response to σ_2 if and only if $\sigma_1(r) \in \Gamma(r, \sigma_2)$ for all r. Let $Br_1(\sigma_2)$ be the set of all best responses of the seller given σ_2. Then we have $Br_1(\sigma_2) = \prod_{r \geq 0} \Gamma(r, \sigma_2)$. Because the correspondence Γ is nonempty, convex valued and upper hemicontinuous, so is Br_1.

Finally, define $Br : \mathscr{S}_1 \times \mathscr{S}_2 \rightrightarrows \mathscr{S}_1 \times \mathscr{S}_2$ as $Br(\sigma_1, \sigma_2) \equiv Br_1(\sigma_2) \times Br_2(\sigma_1)$. Then Br is nonempty, convex valued and upper hemicontinous. Since $\mathscr{S}_1 \times \mathscr{S}_2$ is compact, by Glicksberg (1952)'s generalization of Kakutani's fixed-point theorem, the correspondence Br has a fixed point (σ_1^*, σ_2^*). This profile together with the associated posterior beliefs $\nu^* = \nu^{\sigma_1^*}$ is a stationary public equilibrium. $\square$

A.5 Proof of Theorem 1.3.1

Define $\Delta = u_s(L, s) - u_s(H, s)$ and $\Delta' = u_s(L, c) - u_s(L, s)$. By Assumptions 1 and 3, both $\Delta > 0$ and $\Delta' > 0$.

Lemma A.5.1. *Let (σ_1, σ_2) be any stationary strategy profile and $V : R \to \mathbb{R}$ be the associated seller's value function. Define $\Delta_r = V(r+1) - V(r)$ for all $r \geq 0$. Then if $\varepsilon < \beta$, we have*

$$|\Delta_r| \leq \frac{1-\delta}{\delta} \frac{\Delta'}{\beta - \varepsilon} \quad \forall r \geq 0.$$

Proof. To simplify notations, let $u(r) = u_s(\sigma_1(r), \sigma_2(r))$ and $x_r = \rho(h|\sigma_1(r))$ for all

$r \geq 0$. We show by induction

$$|\Delta_r| \leq \frac{(1-\delta)\Delta'}{\delta(1-\varepsilon)\beta} \sum_{\tilde{r}=0}^{r} \Big(\frac{\varepsilon(1-\beta)}{(1-\varepsilon)\beta}\Big)^{\tilde{r}}.$$

When $r = 0$, we have

$$V(0) = (1-\delta)u(0) + \delta\Big[(1-\varepsilon)x_0 V(1) + (1-(1-\varepsilon)x_0)V(0)\Big].$$

Rearranging yields

$$\Delta_0 = \frac{(1-\delta)\big(V(0)-u(0)\big)}{\delta(1-\varepsilon)x_0}.$$

Thus

$$|\Delta_0| \leq \frac{(1-\delta)\Delta'}{\delta(1-\varepsilon)\beta}.$$

Now assume the desired inequality holds for $\tilde{r} = 0, \cdots, r-1$. For r, we have

$$\begin{aligned} V(r) = (1-\delta)u(r) + \delta\Big[&(1-\varepsilon)x_r V(r+1) \\ &+ \big(1-(1-\varepsilon)x_r - \varepsilon(1-x_r)\big)V(r) + \varepsilon(1-x_r)V(r-1)\Big]. \end{aligned}$$

Rearranging yields

$$\Delta_r = \frac{(1-\delta)\big(V(r)-u(r)\big)}{\delta(1-\varepsilon)x_r} + \frac{\varepsilon(1-x_r)}{(1-\varepsilon)x_r}\Delta_{r-1}.$$

Thus we have

$$|\Delta_r| \leq \frac{(1-\delta)\Delta'}{\delta(1-\varepsilon)\beta} + \Big(\frac{\varepsilon(1-\beta)}{(1-\varepsilon)\beta}\Big)|\Delta_{r-1}| \leq \frac{(1-\delta)\Delta'}{\delta(1-\varepsilon)\beta} \sum_{\tilde{r}=0}^{r} \Big(\frac{\varepsilon(1-\beta)}{(1-\varepsilon)\beta}\Big)^{\tilde{r}},$$

where the second inequality comes from induction hypothesis. Hence the desired inequality follows. When $\varepsilon < \beta$ we then have

$$|\Delta_r| \leq \frac{(1-\delta)\Delta'}{\delta(1-\varepsilon)\beta} \frac{1}{1-\frac{\varepsilon(1-\beta)}{(1-\varepsilon)\beta}} = \frac{1-\delta}{\delta}\frac{\Delta'}{\beta-\varepsilon}.$$

□

Lemma A.5.2. *Assume (σ_1^*, σ_2^*) is a stationary public equilibrium. Let $V : R \to \mathbb{R}$ be the associated seller's value function. Define $\Delta_r = V(r+1) - V(r)$ for all $r \geq 0$. For all $r \geq 1$, If $\sigma_1^*(r) = 0$ and $\Delta_{r-1} > 0$, then $\Delta_r > 0$.*

Proof. If $\sigma_1^*(r) = 0$, then we know $\sigma_2^*(r) = 0$ since it is complete information and buyers play myopic best response. Thus we know

$$V(r) = (1-\delta)u_s(L,s) + \delta\Big[(1-\varepsilon)\beta V(r+1) + \big(1-(1-\varepsilon)\beta - \varepsilon(1-\beta)\big)V(r) + \varepsilon(1-\beta)V(r-1)\Big].$$

Rearranging yields

$$V(r) = \frac{(1-\delta)u_s(L,s) + \delta\varepsilon(1-\beta)V(r-1) + \delta(1-\varepsilon)\beta V(r+1)}{1-\delta+\delta(1-\varepsilon)\beta+\delta\varepsilon(1-\beta)}.$$

Hence $V(r)$ is a convex combination of $u_s(L,s)$, $V(r-1)$ and $V(r+1)$. Notice $u_s(L,s)$ is the minmax value of the stage game, hence $V(r) > V(r-1) \geq u_s(L,s)$ where the first inequality comes from assumption $\Delta_{r-1} > 0$. This directly implies $V(r+1) > V(r)$, or equivalently $\Delta_r > 0$, completing the proof. □

Proof of Theorem 1.3.1. Define

$$\bar{\varepsilon} \equiv \frac{\beta\Delta}{(\alpha-\beta)\Delta' + \Delta}.$$

We show when $\varepsilon < \bar{\varepsilon}$, Theorem 1.3.1 holds.

En route to a contradiction, assume (σ_1^*, σ_2^*) is a stationary public equilibrium in $\Gamma(0,\alpha,\beta,\delta,\varepsilon)$ and $\sigma_1^*(r) > 0$ for some r. Let V be the associated value function and $\Delta_r \equiv V(r+1) - V(r)$ for $r \geq 0$ as above. Because $\sigma_1^*(r) > 0$, incentive at r requires

$$(1-\varepsilon)\Delta_r + \varepsilon\Delta_{r-1} \geq \frac{(1-\delta)\Delta}{\delta(\alpha-\beta)}.$$

Hence

$$(1-\varepsilon)\Delta_r \geq \frac{(1-\delta)\Delta}{\delta(\alpha-\beta)} - \varepsilon\Delta_{r-1} \geq \frac{1-\delta}{\delta}\Big[\frac{\Delta}{\alpha-\beta} - \frac{\varepsilon}{\beta-\varepsilon}\Delta'\Big] > 0, \tag{A.25}$$

where the second inequality comes from Lemma A.5.1 and the third from $\varepsilon < \bar{\varepsilon}$.

We now show value function V is strictly increasing from r on and converges to infinity as r tends to infinity. This leads to contradiction, because value function must be bounded by stage game payoffs.

Step 1: we show if $\sigma_1^*(r') = 0$ for all $r' = r+1, r+2, \cdots, r+k$, then V must be strictly increasing on the domain $\{r, r+1, \cdots, r+k+1\}$. From inequality (A.25), we know $\Delta_r > 0$ and hence $V(r+1) - V(r) > 0$. If $\sigma_1^*(r+1) = 0$, then by Lemma A.5.2, we know $\Delta_{r+1} > 0$ or equivalently $V(r+2) > V(r+1)$. If $\sigma_1^*(r+2) = 0$, then we apply Lemma A.5.2 again and show $V(r+3) > V(r+2)$. This arguments continues for $r' = r+3, \cdots, r+k$ and the result follows.

Step 2: we show it is impossible that $\sigma_1^*(r') = 0$ for all $r' > r$. If this is the case, then from above argument, we know V is strictly increasing from r on and $\lim_{r' \to \infty} V(r') > u_s(L, s)$. But this is impossible because

$$V(r') \leq (1-\delta) \sum_{\tau=0}^{r'-r} \delta^\tau u_s(L, s) + \delta^{r'-r} u_s(L, c)$$

and when r' is arbitrarily large, $V(r')$ should be arbitrarily close to $u_s(L, s)$.

Step 3: We show V must be strictly increasing from r on and converges to infinity. From Steps 1 and 2, we know from r on, there must be infinitely many signals at which the seller exerts high effort with positive probability. Let them be $r = r_0 < r_1 < r_2 < \cdots$. Notice (a) Step 1 implies V must strictly be increasing on the domain $\{r_k + 1, \cdots, r_{k+1}\}$ for all $k \geq 0$ and (b) at each of these signals, inequality (A.25) holds. This directly implies the desired result, completing the proof. $\square$

A.6 Proof of Theorem 1.3.2

Assume $\alpha > 2\beta$. It is easy to see there exists $\underline{\delta}, \overline{\varepsilon} \in (0,1)$ such that when $\delta > \underline{\delta}$ and $\varepsilon < \overline{\varepsilon}$ the following inequality holds

$$\frac{1-\varepsilon}{1-\delta+\delta\varepsilon(1-\beta)+\delta(1-\varepsilon)\beta} > \frac{1}{\delta(\alpha-\beta)}.$$

We now show if α, β, δ and ε are in the above range and $b_0 < \kappa$, then there is no stationary public equilibrium in which the seller always exerts high effort. En route to a contradiction, $(\sigma_1^*, \sigma_2^*, \nu^*)$ is stationary public equilibrium in which the normal seller always exerts low effort, i.e. $\sigma_1^*(r) = 0$ for all $r \geq 0$.

We begin with analysis of the associated posterior beliefs.

Lemma A.6.1. *There exist* $0 < \underline{\lambda} < \overline{\lambda} < 1$ *such that*

$$\frac{\nu^*(r)}{1-\nu^*(r)} = \frac{b_0}{1-b_0}\frac{1-\overline{\lambda}}{1-\underline{\lambda}}\left(\frac{\overline{\lambda}}{\underline{\lambda}}\right)^r \quad \forall r \geq 0.$$

Hence ν^* *strictly increases and* $\nu^*(r) \to 1$ *as* $r \to \infty$.

Proof. Notice for all $r \geq 1$ and $t \geq r$, we have

$$\begin{aligned} P_{\tilde{\xi}}^{\sigma_1^*}(R_t = r) &= (1-\varepsilon)\beta P_{\tilde{\xi}}^{\sigma_1^*}(R_{t-1} = r-1) \\ &\quad + \left[\varepsilon\beta + (1-\varepsilon)(1-\beta)\right] P_{\tilde{\xi}}^{\sigma_1^*}(R_{t-1} = r-1) \\ &\quad + \varepsilon(1-\beta) P_{\tilde{\xi}}^{\sigma_1^*}(R_{t-1} = r+1). \end{aligned}$$

This implies for all $r \geq 1$

$$\begin{aligned}
\sum_{t\geq r} \delta^t P_{\tilde{\xi}}^{\sigma_1^*}(R_t = r) &= (1-\varepsilon)\beta \sum_{t\geq r} \delta^t P_{\tilde{\xi}}^{\sigma_1^*}(R_{t-1} = r-1) \\
&\quad + \big[\varepsilon\beta + (1-\varepsilon)(1-\beta)\big] \sum_{t\geq r} \delta^t P_{\tilde{\xi}}^{\sigma_1^*}(R_{t-1} = r) \\
&\quad + \varepsilon(1-\beta) \sum_{t\geq r} \delta^t P_{\tilde{\xi}}^{\sigma_1^*}(R_{t-1} = r+1) \\
&= \delta(1-\varepsilon)\beta \sum_{t\geq r-1} \delta^t P_{\tilde{\xi}}^{\sigma_1^*}(R_t = r-1) \\
&\quad + \delta\big[\varepsilon\beta + (1-\varepsilon)(1-\beta)\big] \sum_{t\geq r} \delta^t P_{\tilde{\xi}}^{\sigma_1^*}(R_t = r) \\
&\quad + \delta\varepsilon(1-\beta) \sum_{t\geq r+1} \delta^t P_{\tilde{\xi}}^{\sigma_1^*}(R_t = r+1),
\end{aligned}$$

where the second equality comes from the fact $P_{\tilde{\xi}}^{\sigma_1^*}(R_t = r) = 0$ for all $t < r$. For each $r \geq 0$, define

$$q_r \equiv \sum_{t\geq r} \delta^t P_{\tilde{\xi}}^{\sigma_1^*}(R_t = r).$$

Then from the above equality, the sequence $\{q_r\}_{r\geq 0}$ must satisfy the following second order difference equation

$$\delta(1-\varepsilon)\beta q_{r+2} - \big[1 - \delta\varepsilon\beta - \delta(1-\varepsilon)(1-\beta)\big] q_{r+1} + \delta\varepsilon(1-\beta) q_r = 0 \quad \forall r \geq 0.$$

Thus we know there exist constants C_1 and C_2 such that

$$q_r = C_1 \lambda_1^r + C_2 \lambda_2^r, \quad \forall r \geq 0,$$

where

$$\begin{aligned}
\lambda_1 =& (2\delta(1-\varepsilon)\beta)^{-1} \times \Bigg(1 - \delta + \delta(1-\varepsilon)\beta + \delta\varepsilon(1-\beta) \\
&- \sqrt{(1-\delta)^2 + 2\delta(1-\delta)\big[(1-\varepsilon)\beta + \varepsilon(1-\beta)\big] + \delta^2\big[(1-\varepsilon)\beta - \varepsilon(1-\beta)\big]^2}\Bigg)
\end{aligned}$$

and

$$
\begin{aligned}
\lambda_1 =& (2\delta(1-\varepsilon)\beta)^{-1} \times \Bigg(1-\delta+\delta(1-\varepsilon)\beta+\delta\varepsilon(1-\beta) \\
&+ \sqrt{(1-\delta)^2+2\delta(1-\delta)\big[(1-\varepsilon)\beta+\varepsilon(1-\beta)\big]+\delta^2\big[(1-\varepsilon)\beta-\varepsilon(1-\beta)\big]^2}\Bigg)
\end{aligned}
$$

are the two roots of the corresponding characteristic polynomial

$$\delta(1-\varepsilon)\beta\lambda^2 - \big[1-\delta\varepsilon\beta-\delta(1-\varepsilon)(1-\beta)\big]\lambda+\delta\varepsilon(1-\beta)=0.$$

Then it is easy to see that both λ_1 and λ_2 are real, and $0<\lambda_1<1<\lambda_2$. Notice

$$\sum_{r\geq 0} q_r = \sum_{r\geq 0}\sum_{t\geq r}\delta^t P_{\tilde{\xi}}^{\sigma_1^*}(R_t=r) = \sum_{t\geq 0}\delta^t \sum_{0\leq r\leq t} P_{\tilde{\xi}}^{\sigma_1^*}(R_t=r) = \sum_{t\geq 0}\delta^t = \frac{1}{1-\delta}$$

Hence $q_r\to 0$ as $r\to\infty$. Because $q_r = C_1\lambda_1^r + C_2\lambda_2^r$ and $\lambda_2>1$, we must have $C_2=0$. Therefore, $q_r = C_1\lambda_1^r$ for all $r\geq 0$. From the above equation, we know $C_1=(1-\lambda_1)/(1-\delta)$. Define $\underline{\lambda}\equiv\lambda_1$. We then have

$$\sum_{t\geq r}\delta^t P_{\tilde{\xi}}^{\sigma_1^*}(R_t=r) = \frac{1-\underline{\lambda}}{1-\delta}\underline{\lambda}^r \quad \forall r\geq 0.$$

Notice the above analysis also applies to $\sum_{t\geq r}\delta^t P_{\hat{\xi}}(R_t=r)$ with β being replaced by α. If we replace all the β's in the expression of $\underline{\lambda}$ by α and define this value to be $\overline{\lambda}$, we then have

$$\sum_{t\geq r}\delta^t P_{\hat{\xi}}(R_t=r) = \frac{1-\overline{\lambda}}{1-\delta}\overline{\lambda}^r \quad \forall r\geq 0.$$

Moreover, it is straightforward to check $\overline{\lambda}>\underline{\lambda}$ because $\alpha>\beta$. Therefore, we have

$$\frac{\nu^*(r)}{1-\nu^*(r)} = \frac{b_0}{1-b_0}\frac{1-\overline{\lambda}}{1-\underline{\lambda}}\left(\frac{\overline{\lambda}}{\underline{\lambda}}\right)^r \quad \forall r\geq 0$$

as desired. □

Proof of Theorem 1.3.2. From previous lemma, we know

$$\frac{\nu^*(0)}{1-\nu^*(0)} = \frac{b_0}{1-b_0}\frac{1-\overline{\lambda}}{1-\underline{\lambda}} < \frac{b_0}{1-b_0} < \frac{\kappa}{1-\kappa},$$

where the fist inequality comes from $\overline{\lambda} > \underline{\lambda}$ and the second inequality from the assumption $b_0 < \kappa$. Equivalently, $\nu^*(0) < \kappa$. Because ν^* strictly increases and $\nu^*(r) \to 1$, we know there exists $\overline{r} \geq 0$ such that $\nu^*(r) < \kappa$ for $r \leq \overline{r}$ and $\nu^*(r) > \kappa$ for $r > \overline{r}$. Because the normal seller always shirks and buyers play myopic best response, we know $\sigma_2^*(r) = 0$ for $r \leq \overline{r}$ and $\sigma_2^*(r) = 1$ for $r > \overline{r}$, i.e. buyers choose the standard product if the signal is less than or equal to $\overline{r}$, and choose the customized product if the signal is above $\overline{r}$. Because by assumption, the normal seller always exerts low effort, if the current signal is $r \leq \overline{r}$, the seller's current period payoff is $u_s(L, s)$. If instead the current signal is $r > \overline{r}$, the current period payoff is $u_s(L, c)$.

Let $V : R \to \mathbb{R}$ be the seller's value function induced by (σ_1^*, σ_2^*). Because by Assumption 1 and 3, $u_s(L, c) > u_s(H, c) > u_s(L, s)$, it is easy to see V is increasing. Then we have

$$\begin{aligned} V(\overline{r}) &= (1-\delta)u_s(L,s) + \delta\Big[(1-\varepsilon)\beta V(\overline{r}+1) + \big(\varepsilon\beta + (1-\varepsilon)(1-\beta)\big)V(\overline{r}) \\ &\quad +\varepsilon(1-\beta)V(\max(\overline{r}-1,0))\Big] \\ &\leq (1-\delta)u_s(L,s) + \delta\Big[(1-\varepsilon)\beta V(\overline{r}+1) + \big(1-(1-\varepsilon)\beta\big)V(\overline{r})\Big], \end{aligned}$$

and

$$\begin{aligned} V(\overline{r}+1) &= (1-\delta)u_s(L,c) + \delta\Big[(1-\varepsilon)\beta V(\overline{r}+2) + \big(\varepsilon\beta + (1-\varepsilon)(1-\beta)\big)V(\overline{r}+1) \\ &\quad +\varepsilon(1-\beta)V(\overline{r})\Big] \\ &\geq (1-\delta)u_s(L,c) + \delta\Big[\big(1-\varepsilon(1-\beta)\big)V(\overline{r}+1) + \varepsilon(1-\beta)V(\overline{r})\Big]. \end{aligned}$$

where the first inequality comes from $V(\overline{r}) \geq V(\max\{\overline{r}-1, 0\})$ and the second from $V(\overline{r}+2) \geq V(\overline{r}+1)$. Combining these two inequalities yields

$$V(\overline{r}+1) - V(\overline{r}) \geq \frac{1-\delta}{1-\delta+\delta\varepsilon(1-\beta)+\delta(1-\varepsilon)\beta}\big(u_s(L,c) - u_s(L,s)\big).$$

Because by Assumptions 1 and 3 $u_s(L, c) - u_s(H, c) \geq u_s(L, s) - u(H, s)$ and $u_s(H, c) >$

$u(L,s)$, we have $u_s(L,c) - u_s(L,s) > u_s(L,s) - u_s(H,s)$. Therefore,

$$\begin{aligned} & (1-\varepsilon)\big[V(\overline{r}+1) - V(\overline{r})\big] + \varepsilon\big[V(\overline{r}) - V(\max\{\overline{r}-1,0\})\big] \\ \geq\ & (1-\varepsilon)\big[V(\overline{r}+1) - V(\overline{r})\big] \\ \geq\ & \frac{(1-\varepsilon)(1-\delta)}{1-\delta+\delta\varepsilon(1-\beta)+\delta(1-\varepsilon)\beta}\big(u_s(L,s) - u_s(H,s)\big). \end{aligned}$$

Since $\frac{1-\varepsilon}{1-\delta+\delta\varepsilon(1-\beta)+\delta(1-\varepsilon)\beta} > \frac{1}{\delta(\alpha-\beta)}$, we have

$$(1-\varepsilon)\big[V(\overline{r}+1) - V(\overline{r})\big] + \varepsilon\big[V(\overline{r}) - V(\max\{\overline{r}-1,0\})\big] > \frac{1-\delta}{\delta(\alpha-\beta)}\big(u_s(L,s) - u_s(H,s)\big).$$

Rearranging terms, we can see this inequality is equivalent to

$$\begin{aligned} & (1-\delta)u_s(H,s) + \delta\Big[(1-\varepsilon)\alpha V(\overline{r}+1) + \big(\varepsilon\alpha + (1-\varepsilon)(1-\alpha)\big)V(\overline{r}) \\ & +\varepsilon(1-\alpha)V(\max(\overline{r}-1,0))\Big] \\ >\ & (1-\delta)u_s(L,s) + \delta\Big[(1-\varepsilon)\beta V(\overline{r}+1) + \big(\varepsilon\beta + (1-\varepsilon)(1-\beta)\big)V(\overline{r}) \\ & +\varepsilon(1-\beta)V(\max(\overline{r}-1,0))\Big]. \end{aligned}$$

Therefore, exerting high effort at signal $\overline{r}$ is a profitable deviation for the seller. This contradicts to the assumption that $(\sigma_1^*, \sigma_2^*, \nu^*)$ is an equilibrium, completing the proof. □

A.7 Proof of Theorem 1.3.3

The next lemma roughly states that there is no stationary public equilibrium in which the normal type seller always mimics the commitment type. A similar argument also appears in Mailath and Samuelson (2001).

Lemma A.7.1. *There exists a $K > 0$ such that in all stationary public equilibrium σ^*, for any K consecutive signals $r, r+1, \cdots, r+K-1$, there exists at least $0 \leq k \leq K-1$ at which the normal type seller makes low effort, i.e. $\sigma_1^*(r_k) = 0$.*

Proof. Define $\triangle \equiv u_s(L,s) - u_s(H,s)$ to be the minimal gain of the seller from exerting low effort. By Assumption 1, $\triangle > 0$. Define K to be any integer such that

$$\frac{\triangle(1-\delta)}{\delta(\alpha-\beta)}K > u_s(L,c) - u_s(H,s).$$

Fix any stationary public equilibrium σ^*. Let $\{V(r)\}_r$ be the associated value function of the seller. En route to a contradiction, assume there are K consecutive signals $r, \cdots, r+K-1$ such that $\sigma^*(r+k) > 0$ for all $0 \le k \le K-1$. Then At each of these signal, the seller weakly prefers making high effort. Therefore a necessary condition is for $0 \le k \le K-1$

$$\begin{aligned}
& (1-\delta)u_s(H,\sigma_2^*(r+k)) + \delta(1-\varepsilon)\alpha V(r+k+1) \\
& +\delta(\varepsilon\alpha + (1-\varepsilon)(1-\alpha))V(r+k) + \delta\varepsilon(1-\alpha)V(r+k-1) \\
\ge\ & (1-\delta)u_s(L,\sigma_2^*(r+k)) + \delta(1-\varepsilon)\beta V(r+k+1) \\
& +\delta(\varepsilon\beta + (1-\varepsilon)(1-\beta))V(r+k) + \delta\varepsilon(1-\beta)V(r+k-1).
\end{aligned}$$

This implies

$$\begin{aligned}
& \delta(1-\varepsilon)(\alpha-\beta)\big[V(r+k+1) - V(r+k)\big] \\
& +\delta\varepsilon(\alpha-\beta)\big[V(r+k) - V(r+k-1)\big] \\
\ge\ & (1-\delta)\big[u_s(L,\sigma_2^*(r+k)) - u_s(H,\sigma_2^*(r+k))\big] \\
\ge\ & (1-\delta)\triangle, \quad \forall 0 \le k \le K-1,
\end{aligned}$$

where the last inequality comes from Assumption 1. Equivalently

$$(1-\varepsilon)\big[V(r+k+1) - V(r+k)\big] + \varepsilon\big[V(r+k) - V(r+k-1)\big] \ge \frac{(1-\delta)\triangle}{\delta(\alpha-\beta)}, \quad \forall 0 \le k \le K-1.$$

Summing up these inequalities yields

$$\begin{aligned}
&\frac{\triangle(1-\delta)}{\delta(\alpha-\beta)}K \\
\leq\ & \sum_{k=0}^{K-1}\Big((1-\varepsilon)\big[V(r+k+1)-V(r+k)\big]+\varepsilon\big[V(r+k)-V(r+k-1)\big]\Big) \\
=\ & (1-\varepsilon)\big[V(r+K)-V(r)\big]+\varepsilon\big[V(r+K-1)-V(r-1)\big] \\
\leq\ & u_s(L,c)-u_s(H,s),
\end{aligned}$$

where the last inequality comes from the face $u_s(H,s) \leq V(r) \leq u_s(L,c)$ by Assumptions 1 and 3. This contradicts the definition of K. □

Combining Lemmas A.7.1, A.3.1 and A.3.4, we have the following property about asymptotic behavior of posterior beliefs in any stationary public equilibrium.

Lemma A.7.2. *In any stationary public equilibrium σ^*, the associated posterior belief $\nu^*(r) \to 1$ as $r \to \infty$.*

Proof. It suffices to show $\frac{1-\nu^*(r)}{\nu^*(r)} \to 0$. From Equation (A.22) and Lemma A.3.4, we know

$$\frac{1-\nu^*(r)}{\nu^*(r)} = \frac{1-b_0}{b_0}\frac{\sum_{t=r}^{\infty}\delta^t P_{\tilde{\xi}}^{\sigma_1^*}(R_t=r)}{\sum_{t=r}^{\infty}\delta^t P_{\widehat{\xi}}(R_t=r)} = \frac{1-b_0}{b_0}\frac{M^{\sigma_1^*}(r)}{M^{\widehat{\sigma}_1}(r)}\frac{\prod_{\tilde{r}=0}^{r-1}\tilde{a}_{\tilde{r}}}{\prod_{\tilde{r}=0}^{r-1}\hat{a}_{\tilde{r}}} \leq \frac{1-b_0}{b_0}M\frac{\prod_{\tilde{r}=0}^{r-1}\tilde{a}_{\tilde{r}}}{\prod_{\tilde{r}=0}^{r-1}\hat{a}_{\tilde{r}}}.$$

By Lemma A.7.1, in σ_1^* there are infinitely many signals at which the normal type exerts low effort. Lemma A.3.3 then implies the right hand side converges to 0, as desired. □

Proof of Theorem 1.3.3. Fix a stationary public equilibrium $(\sigma_1^*, \sigma_2^*, \nu^*)$. By Lemma A.7.2, we know $\lim_r \nu^*(r) \to 1$. Hence there exists $r^\dagger$ such that $\nu^*(r) > \kappa$ for all $r \geq r^\dagger$. Then all buyers who see signal $r \geq r^\dagger$ believes that the seller will exert effort with probability at least $\nu^*(r) > \kappa$. Because all buyers are short-lived, they just play their best response. Therefore by Assumption 2, after seeing signal $r \geq r^\dagger$, all entering buyer chooses the customized product, i.e. $\sigma_2^*(r) = 1$ for all $r \geq r^\dagger$.

We now turn to the behavior of the seller. Let $V(\cdot)$ be the associated value function of the normal type of the seller. Let $K > 0$ be an integer such that

$$\delta^K \big[u_s(L,c) - u_s(H,s)\big] < \frac{1-\delta}{\delta} \frac{u_s(L,c) - u_s(H,c)}{\alpha - \beta}.$$

Then at any signal $r \geq r^\dagger + K$, we know all buyers who enter in the next K periods will observe signal higher than or equal to $r^\dagger$ because the signal can decrease by at most 1 for every buyer. Therefore all of the next K entering buyers will choose the customized product. Thus by exerting low effort for the next K buyers, the normal seller can guarantee himself a payoff at least

$$(1-\delta) \sum_{k=0}^{K-1} \delta^k u_s(L,c) + \delta^K u_s(H,s).$$

This in turn implies

$$V(r+1) - V(r) \leq \delta^K \big[u_s(L,c) - u_s(H,s)\big]$$

for all $r \geq r^\dagger + K$ because all $V(r+1)$ is bounded above by $u_s(L,c)$ which is the highest possible stage payoff to the seller from Assumptions 1 and 3. Then for all $r \geq r^\dagger + K + 1$, we have

$$\begin{aligned} & (1-\varepsilon)\big[V(r+1) - V(r)\big] + \varepsilon\big[V(r) - V(r-1)\big] \\ \leq & \delta^K \big[u_s(L,c) - u_s(H,s)\big] \\ < & \frac{1-\delta}{\delta} \frac{u_s(L,c) - u_s(H,c)}{\alpha - \beta}. \end{aligned}$$

But this is equivalent to

$$\begin{aligned} & (1-\delta)u_s(L,c) + \delta(1-\varepsilon)\alpha V(r+1) \\ & + \delta(\varepsilon\alpha + (1-\varepsilon)(1-\alpha))V(r) + \delta\varepsilon(1-\alpha)V(r-1) \\ > & (1-\delta)u_s(H,c) + \delta(1-\varepsilon)\beta V(r+1) \\ & + \delta(\varepsilon\beta + (1-\varepsilon)(1-\beta))V(r) + \delta\varepsilon(1-\beta)V(r-1), \end{aligned}$$

which implies that the normal seller strictly prefers L to H. Hence we have $\sigma_1^*(r) = 0$ for all $r \geq \underline{r} \equiv r^\dagger + K + 1$. This completes the proof. $\square$

A.8 Proofs for Section 1.4

Proof of Lemma 1.4.1. Suppose by contradiction there exists such a μ. Without loss of generality, assume $\{i \in \mathbb{N} | \mu(\tau_i < \infty) > 0\} = \mathbb{Z}$. If all agents have identical prior belief about when they enter, they must have the same probability of entering in period 1, i.e. $\mu_1^i = \mu_1^1$ for all $i \geq 0$. Because by definition, $\cup_i(\tau_i = 1) = \Sigma$, we have $\sum_{i\geq 1} \mu_1^i = \sum_{i\geq 1} \mu(\tau_i = 1) = \mu(\Sigma) = 1$. This implies $\mu_1^1 > 0$ and $\sum_{i\geq 0} \mu_1^i = +\infty$, a contradiction. □

Proof of Lemma 1.4.2. We postpone the proof in Lemma A.8.3 below. □

Proof of Proposition 1.4.1. Fix any $\delta \in (0, 1)$. We explicitly construct a random entry model μ that satisfies EPS and S with continuation probability δ.

For each vector $\theta_n = (n-1, n-2, \cdots, 0) \in \Sigma_n$, define $\mu(\{\theta_n\}) \equiv (1-\delta)\delta^{n-1}$. It is easy to see μ is a probability measure over Σ. By construction, $\mu(\Sigma_n) = \mu(\{\theta_n\}) = (1-\delta)\delta^{n-1}$ and hence S is satisfied. Moreover, for each i and t, agent i enters in period t if and only if agent $i+t$ enters in period 1 if and only if θ_{i+t} is realized. Thus $(\tau_i = t) = \{\theta_{i+t}\}$ and $\mu_t^i = (1-\delta)\delta^{i+t}$. Therefore, $\mu_t^i = \delta^{i-j}\mu_t^j$ for all i, j and t. By Lemma 1.4.2, μ satisfies EPS. □

Proof of Proposition 1.4.2. Let μ be a random entry model that satisfy EPS and S with continuation probability δ. Let $\{\mu_i^t\}_{i,t}$ be the corresponding prior beliefs about entry time.

Because μ satisfies EPS, by Lemma 1.4.2 there exist $\{c^i\}$ such that

$$\mu_t^i = c^i \mu_t^1 \quad \forall i, t \geq 0.$$

For any $t \geq 0$, we then have

$$\sum_i \mu_t^i = \big(\sum_i c^i\big)\mu_t^1.$$

But notice $\sum_i \mu_t^i = \sum_i \mu(\tau_i = t)$ is the probability of at least t entries in model μ, which is equal to $\sum_{n\geq t}(1-\delta)\delta^{n-1}$ by S. Therefore we have

$$\mu_t^1 = \frac{\delta^{t-1}}{\sum_i c^i} \quad \forall t \geq 0.$$

Plugging this expression into Equation (1.4) yields

$$\nu_1^P(\xi|r) = \frac{b(\xi)\sum_{t=1}^{\infty}\delta^t P_\xi\big(R_t = r\big)}{\sum_{\xi'} b(\xi')\sum_{t=1}^{\infty}\delta^t P_{\xi'}\big(R_t = r\big)}.$$

Because of EPS, every agent must have the same posterior belief as agent 1 does. This leads to the desired result. □

Remark A.8.1. As mentioned in footnote 11, an alternative way of modeling infinite entries is to consider random entry models that last forever. Formally, let $\Sigma_\infty \equiv \{(i_1, i_2, \cdots) \in \mathbb{Z}^\infty | i_s \neq i_t, \text{ if } s \neq t\}$. A random entry model that last forever is just a probability measure over Σ_∞. As in the context, we can similarly define prior beliefs about entering time for each agent. Such a random entry model is naturally stationary: it continues after every period with probability 1. However, there is no such random entry model that satisfies EPS, as stated in the following lemma.

Lemma A.8.1. *There is no random entry model μ over Σ_∞ that satisfies EPS.*

Proof. Suppose by contradiction μ satisfies EPS. Without loss of generality, assume agent 1 enters in period 1 with positive probability. A careful examination of the proof in Lemma 1.4.2 shows Lemma 1.4.2 also holds in this case. Hence for each i there exists c^i such that $\mu_t^i = c^i \mu_t^1$ for all $t \geq 0$. Hence $1 = \sum_i \mu_t^i = \sum_i c^i \mu_t^1$ for all t. This implies $\mu_t^1 = 1/(\sum_i c^i)$ for all t. But $1 \geq \mu(\tau_1 < +\infty) = \sum_t \mu_t^1 = \sum_t 1/(\sum_i c^i) = \infty$, a contradiction. □

Remark A.8.2. As mentioned in footnote 13, the notion of EPS in Definition 1.4.2 seems too strong for the application of the reputation game because it requires common posterior beliefs even for stochastic processes that can not be generated by any strategy of the seller. We examine this problem here.

To do so, we first consider strategies of the seller in the reputation game. Because the seller is long-lived, a private history of the seller consists of all past signals, actions and qualities. A strategy of the seller is a mapping from the set of all possible private histories to the interval $[0, 1]$. We now consider a particular class of strategies of the seller that only depend on current signal and current period (calendar time). Denote by $\mathscr{S}_1^*$ this class of strategies. It is clear that the set of all stationary public strategies of the seller is a subset of $\mathscr{S}_1^*$. Every strategy $\sigma_1 \in \mathscr{S}_1^*$, together with prior belief b_0, the strategy of the commitment type and the evolution of the public signals, induces a probability measure over $\Xi \times R^\infty$. Denote by $\mathcal{Q}^*$ the set of all probability measures generated by such strategies, which is clear a subset of $\mathcal{P}$. We proceed to show that any random entry model that induces common posterior beliefs for all processes in $\mathcal{Q}^*$ must also induces common posterior beliefs for all processes in $\mathcal{P}$. This in turn will imply the notion of EPS in Definition 1.4.2 is not as strong as it seems.

The following lemma states that though $\mathcal{Q}^*$ is only a small subset of $\mathcal{P}$, it is in fact very rich.

Lemma A.8.2. *There exists a sequence of strategies $\{\sigma_1^k\}_{k\geq 0} \subset \mathscr{S}_1^*$ such that for all $k \geq 1$,*

(i). $P_{\tilde{\xi}}^{\sigma_1^k}(R_t = r) = P_{\tilde{\xi}}^{\sigma_1^0}(R_t = r)$ for all $t \neq k$, $r \geq 0$, and

(ii). $P_{\tilde{\xi}}^{\sigma_1^k}(R_k = 0) \neq P_{\tilde{\xi}}^{\sigma_1^0}(R_k = 0)$,

where $P_{\tilde{\xi}}^{\sigma_1^k}$ is the probability measure over R^∞ induced by σ_1^k.

Proof. We show this result by construction. Let σ_1^0 be strategy which always mixes between high and low effort with equal probability. For each $k \geq 1$, we consider a strategy σ_1^k that imitates σ_1^0 in periods $0, 1, \cdots, k-2, k+1, k+2, \cdots$ but differs from σ_1^0 in period $k-1$ and k. Let $\varsigma_{k-1,0}, \cdots, \varsigma_{k-1,k-1}$ and $\varsigma_{k,0}, \cdots, \varsigma_{k,k}$ denote the corresponding actions (to be specified) in periods $k-1$ and k, i.e. $\varsigma_{t,r}$ is the probability of making high effort if the current signal is r in period t, for $t = k-1, k$.[5] It is clear

[5]Recall from the evolution of the public signals, in any period t, only signals $0, 1, \cdots, t$ can be

$\{\sigma_1^k\}_{k\geq 0} \subset \mathscr{S}_1^*$ by construction because each strategy σ_1^k only depends on current signal and current period.

It is straightforward to see the distributions of signals in period $0, 1, \cdots, k-1$ are identical under $P_{\tilde{\xi}}^{\sigma_1^0}$ and $P_{\tilde{\xi}}^{\sigma_1^k}$. Moreover the distributions of signals in period k and $k+1$ under $P_{\tilde{\xi}}^{\sigma_1^k}$ now can be calculated recursively:

$$\begin{aligned} P_{\tilde{\xi}}^{\sigma_1^k}(R_k = 0) &= \big[1-(1-\varepsilon)\rho(h|\varsigma_{k-1,0})\big] P_{\tilde{\xi}}^{\sigma_1^0}(R_{k-1} = 0) \\ &\quad + \varepsilon\rho(l|\varsigma_{k-1,1}) P_{\tilde{\xi}}^{\sigma_1^0}(R_{k-1} = 1), \end{aligned} \tag{A.26}$$

$$\begin{aligned} P_{\tilde{\xi}}^{\sigma_1^k}(R_k = r) &= \big[(1-\varepsilon)\rho(h|\varsigma_{k-1,r-1})\big] P_{\tilde{\xi}}^{\sigma_1^0}(R_{k-1} = r-1) \\ &\quad + \big[\varepsilon\rho(h|\varsigma_{k-1,r}) + (1-\varepsilon)\rho(l|\varsigma_{k-1,r})\big] P_{\tilde{\xi}}^{\sigma_1^0}(R_{k-1} = r) \\ &\quad + \varepsilon\rho(l|\varsigma_{k-1,r+1}) P_{\tilde{\xi}}^{\sigma_1^0}(R_{k-1} = r+1) \quad \forall r \geq 1, \end{aligned} \tag{A.27}$$

$$\begin{aligned} P_{\tilde{\xi}}^{\sigma_1^k}(R_{k+1} = 0) &= \big[1-(1-\varepsilon)\rho(h|\varsigma_{k,0})\big] P_{\tilde{\xi}}^{\sigma_1^k}(R_k = 0) \\ &\quad + \varepsilon\rho(l|\varsigma_{k,1}) P_{\tilde{\xi}}^{\sigma_1^k}(R_{k-1} = 1), \end{aligned} \tag{A.28}$$

$$\begin{aligned} P_{\tilde{\xi}}^{\sigma_1^k}(R_{k+1} = r) &= \big[(1-\varepsilon)\rho(h|\varsigma_{k,r-1})\big] P_{\tilde{\xi}}^{\sigma_1^k}(R_k = r-1) \\ &\quad + \big[\varepsilon\rho(h|\varsigma_{k,r}) + (1-\varepsilon)\rho(l|\varsigma_{k,r})\big] P_{\tilde{\xi}}^{\sigma_1^k}(R_k = r) \\ &\quad + \varepsilon\rho(l|\varsigma_{k,r+1}) P_{\tilde{\xi}}^{\sigma_1^k}(R_k = r+1) \quad \forall r \geq 1. \end{aligned} \tag{A.29}$$

We are now ready to specify $\varsigma_{k-1,0}, \cdots, \varsigma_{k-1,k-1}$ and $\varsigma_{k,0}, \cdots, \varsigma_{k,k}$. Let $\varsigma_{k-1,r} = \frac{1}{2}$ for all $1 \leq r \leq k-1$ and $\varsigma_{k-1,0}$ be some arbitrary number different from but close to $\frac{1}{2}$. It is then straightforward to see that $P_{\tilde{\xi}}^{\sigma_1^k}(R_k = 0) \neq P_{\tilde{\xi}}^{\sigma_1^0}(R_k = 0)$ from (A.26). Hence condition (ii) of the lemma is satisfied. Given the determined $\varsigma_{k-1,0}, \cdots, \varsigma_{k-1,k-1}$ (and hence $\{P_{\tilde{\xi}}^{\sigma_1^k}(R_k = r)\}_{r=0}^k$ from (A.26) and (A.27)), it is then easy to see from (A.28) and (A.29), condition (i) of the lemma specifies $k+1$ linear equations

$$P_{\tilde{\xi}}^{\sigma_1^k}(R_{k+1} = r) = P_{\tilde{\xi}}^{\sigma_1^0}(R_{k+1} = r) \quad \forall r = 0, \cdots, k \tag{A.30}$$

in $k+1$ unknowns $\varsigma_{k,0}, \cdots, \varsigma_{k,k}$. By a careful examination of this system of linear equations, it is easy to see that the corresponding coefficient matrix is always of full

reached with positive probability. Hence the actions at signal $r > t$ in period t is irrelevant.

rank. Therefore, this system of linear equations has a solution $\varsigma_{k,0}, \cdots, \varsigma_{k,k}$. Moreover, if $\varsigma_{k-1,0}$ is chosen to be close enough to $\frac{1}{2}$, then we know from (A.26) and (A.27), the signal distribution in period k under $P_{\widetilde{\xi}}^{\sigma_1^k}$ and that under $P_{\widetilde{\xi}}^{\sigma_1^0}$ is close. This implies the solutions $\varsigma_{k,0}, \cdots, \varsigma_{k,k}$ of the system of linear equations in (A.30) are also close to $\frac{1}{2}$, making sure $\varsigma_{k,0}, \cdots, \varsigma_{k,k}$ are well defined probabilities.

In sum we have shown under σ_1^k, condition (i) of the lemma is satisfied for all periods $t = 0, \cdots, k-1, k+1$. Notice by construction σ_1^k coincide with σ_1^0 from period $k+1$ on. Therefore, condition (i) is also satisfied for all periods $t = k+2, k+3, \cdots$. This completes the proof. □

We say a random entry model μ satisfies $\mathcal{Q}^*$-EPS, if (1.5) holds for all $P \in \mathcal{Q}^*$. Finally we have

Lemma A.8.3. *The following statements are equivalent:*

(i) μ satisfies $\mathcal{Q}^*$-EPS.

(ii) For all $i, j \in \mathbb{Z}$, there exists c^{ij} such that $\mu_t^i = c^{ij}\mu_t^j$ for all $t \geq 0$.

(iii) μ satisfies EPS.

Proof. $(i) \implies (ii)$: Assume μ satisfies $\mathcal{Q}^*$-EPS. Let $\{\sigma_1^k\}_{k\geq 0}$ be the sequence of strategies specified in Lemma A.8.2.

Because μ satisfies $\mathcal{Q}^*$-EPS, we have for all i, j and $P \in \mathcal{Q}^*$,

$$\nu_i^{P\otimes\mu}(\xi|0) = \nu_j^{P\otimes\mu}(\xi|0) \quad \text{and} \quad \nu_i^{P\otimes\mu}(\xi'|0) = \nu_j^{P\otimes\mu}(\xi'|0).$$

This implies i and j's posterior likelihood ratios between $\widetilde{\xi}$ and $\widehat{\xi}$ are identical. From Equation 1.4, this can be written as

$$\frac{\sum_{t=0}^{\infty} \mu_t^i P_{\widetilde{\xi}}\Big(R_t = 0\Big)}{\sum_{t=0}^{\infty} \mu_t^i P_{\widehat{\xi'}}\Big(R_t = 0\Big)} = \frac{\sum_{t=0}^{\infty} \mu_t^j P_{\widetilde{\xi}}\Big(R_t = 0\Big)}{\sum_{t=0}^{\infty} \mu_t^j P_{\widehat{\xi'}}\Big(R_t = 0\Big)} \quad \forall P \in \mathcal{Q}^*. \tag{A.31}$$

Define

$$c^{ij} \equiv \frac{\sum_{t=0}^{\infty} \mu_t^i P_{\widehat{\xi}}\Big(R_t = 0\Big)}{\sum_{t=0}^{\infty} \mu_t^j P_{\widehat{\xi}}\Big(R_t = 0\Big)}.$$

When restricting attention to $P_{\tilde{\xi}} \in \{P_{\tilde{\xi}}^{\sigma_1^k}\}_{k\geq 0}$ and rearranging, (A.31) can be written as

$$\sum_{t=0}^{\infty}(\mu_t^i - c^{ij}\mu_t^j)P_{\tilde{\xi}}^{\sigma_1^k}\Big(R_t = 0\Big) = 0 \quad \forall k \geq 0.$$

Using the fact that for each $k \geq 1$, the two sequences $\{P_{\tilde{\xi}}^{\sigma_1^k}(R_t = 0)\}_{t\geq 0}$ and $\{P_{\tilde{\xi}}^{\sigma_1^0}(R_t = 0)\}_{t\geq 0}$ only differ when $t = k$, we immediately know that $\mu_t^i = c^{ij}\mu_t^j$ for all $t \geq 0$.

$(ii) \Longrightarrow (iii)$: Straightforward from Equation (1.4).

$(iii) \Longrightarrow (i)$: Straightforward by definition. □

Appendix B

Proofs for Chapter 2

B.1 Proof of Proposition 2.3.1

One important property of relative entropy is the chain rule. Let P and Q be two distributions over the product $X \times Y$ (see for example Cover and Thomas (2006) Chapter 2 and Gossner (2011b)). The chain rule states that the relative entropy of P and Q can be expanded as the sum of a relative entropy and a conditional relative entropy:

$$d(P\|Q) = d(P_X\|Q_X) + E_{P_X}(d(P(\,\cdot\,|x)\|Q(\,\cdot\,|x))),$$

where P_X (resp., Q_X) is the marginal distribution of P (resp., Q) over X and $P(\,\cdot\,|x)$ (resp., $Q(\,\cdot\,|x)$) is the conditional probability of P (resp., Q) over Y given x.

Fix a commitment type $\xi(\hat{\alpha}_1) \in \hat{\Xi}$. Suppose $\sigma = (\sigma_1, \sigma_2)$ is a Nash equilibrium of the reputation game with exogenous learning. Let P^σ be the probability measure over $\Xi \times (A_1 \times A_2 \times Z_1 \times Z_2 \times Y)^\infty$ induced by σ, μ and $\{\rho(\,\cdot\,|\xi)\}_{\xi\in\Xi}$, as in section 2. Let $\widehat{P}^\sigma$ be the conditional probability of P^σ given the event that player 1 is the commitment type $\xi(\hat{\alpha}_1)$. The measure $\widehat{P}^\sigma$ determines how the play evolves if player 1 is of type $\xi(\hat{\alpha}_1)$.

Let $\sigma_1' \in \Sigma_1$ be the strategy for player 1 in which the normal type of player

1 mimics the behavior of the commitment type $\xi(\hat{\alpha}_1)$, i.e. $\sigma_1'(\xi_0, h_1^t) = \hat{\alpha}_1$ for all $h_1^t \in \bigcup_{t\geq 0} H_{1t}$. Let $\sigma' = (\sigma_1', \sigma_2)$. The probability measure $\widetilde{P}^{\sigma'}$ (recall from section 2, $\widetilde{P}^{\sigma'} = P^{\sigma'}\big(\,\cdot\,\big|\{\xi_0\}\times(A_1\times A_2\times Z_1\times Z_2\times Y)^\infty\big)$) describes how the normal type of player 1 expects the play to evolve if he *deviates* to the commitment strategy of $\xi(\hat{\alpha}_1)$. The only difference between $\widetilde{P}^{\sigma'}$ and $\widehat{P}^{\sigma}$ is the distributions of player 2's exogenous signals. Because we assume the realizations of player 2's exogenous signals only depend on the type of player 1 and are independent of the play, for all $h^t \in (A_1 \times A_2 \times Z_1 \times Z_2 \times Y)^t$ we have

$$\widetilde{P}^{\sigma'}(h^t) = \widehat{P}^{\sigma}(h^t)\prod_{\tau=0}^{t-1}\frac{\rho(y^\tau|\xi_0)}{\rho(y^\tau|\xi(\hat{\alpha}_1))},$$

where $y^0, y^1, \cdots, y^{t-1}$ are the exogenous signals contained in the history h^t. Notice by Assumption 2.2.1, $\rho(y|\xi(\hat{\alpha}_1)) > 0$ whenever $\rho(y|\xi_0) > 0$. Hence the right hand side of the above equality is well defined.

Let P_2^{σ}, $\widetilde{P}_2^{\sigma'}$ and $\widehat{P}_2^{\sigma}$ be the marginal distributions of P^{σ}, $\widetilde{P}^{\sigma'}$ and $\widehat{P}^{\sigma}$ respectively on player 2's histories $(Z_2 \times Y)^\infty$, and let $\{P_{2t}^{\sigma}\}_{t\geq 1}$, $\{\widetilde{P}_{2t}^{\sigma'}\}_{t\geq 1}$ and $\{\widehat{P}_{2t}^{\sigma}\}_{t\geq 1}$ be the corresponding finite dimensional distributions. In period -1 before the play, player 2 believes that P_{2t}^{σ} is the distributions of his signals (both endogenous and exogenous) in the first t periods. However, if player 1 is the normal type and he deviates to the commitment strategy of $\xi(\hat{\alpha}_1)$, $\widetilde{P}_{2t}^{\sigma'}$ is the true distribution of player 2's signals in the first t periods. The following lemma gives an upper bound on the prediction errors in player 2's first t periods signals.

Lemma B.1.1. *For all $t \geq 1$,*

$$d\big(\widetilde{P}_{2t}^{\sigma'}\|P_{2t}^{\sigma}\big) \leq -\log\mu\big(\xi(\hat{\alpha}_1)\big) + t\lambda_{\xi(\hat{\alpha}_1)}.$$

Proof. We show this by a simple calculation:

$$
\begin{aligned}
d(\widetilde{P}_{2t}^{\sigma'} \| P_{2t}^{\sigma}) &\equiv \sum_{h_2^t \in H_{2t}} \widetilde{P}_{2t}^{\sigma'}(h_2^t) \log \frac{\widetilde{P}_{2t}^{\sigma'}(h_2^t)}{P_{2t}^{\sigma}(h_2^t)} \\
&= \sum_{h_2^t \in H_{2t}} \widetilde{P}_{2t}^{\sigma'}(h_2^t) \log \left[\frac{\widehat{P}_{2t}^{\sigma}(h_2^t)}{P_{2t}^{\sigma}(h_2^t)} \prod_{\tau=0}^{t-1} \frac{\rho(y^\tau | \xi_0)}{\rho(y^\tau | \xi(\hat{\alpha}_1))} \right] \\
&= \sum_{h_2^t \in H_{2t}} \widetilde{P}_{2t}^{\sigma'}(h_2^t) \log \frac{\widehat{P}_{2t}^{\sigma}(h_2^t)}{P_{2t}^{\sigma}(h_2^t)} + \sum_{h_2^t \in H_{2t}} \widetilde{P}_{2t}^{\sigma'}(h_2^t) \log \Big(\prod_{\tau=0}^{t-1} \frac{\rho(y^\tau | \xi_0)}{\rho(y^\tau | \xi(\hat{\alpha}_1))} \Big).
\end{aligned}
$$

Notice the second term is the relative entropy of the distributions on player 2's exogenous signals in the first t periods when player 1 is the normal type and when he is the commitment type $\xi(\hat{\alpha}_1)$. Because the exogenous signals are conditionally independent across time, the chain rule implies the second term is exactly $t\lambda_{\xi(\hat{\alpha}_1)}$. Moreover, since $\widehat{P}_{2t}^{\sigma}$ is obtained by conditioning P_{2t}^{σ} on the event that player 1 is the commitment type $\xi(\hat{\alpha}_1)$, we have

$$
\frac{\widehat{P}_{2t}^{\sigma}(h_2^t)}{P_{2t}^{\sigma}(h_2^t)} \leq \mu\big(\xi(\hat{\alpha}_1)\big) \quad \forall h_2^t \in H_{2t}.
$$

Therefore the first term is no greater than $-\log \mu(\hat{\theta})$. These two observations imply the desired result. □

For any private history $h_2^t \in \bigcup_{t \geq 0} H_{2t}$, $P_{2,t+1}^{\sigma}$ (resp., $\widetilde{P}_{2,t+1}^{\sigma'}$) induces player 2's one step ahead prediction on his *endogenous* signals $z_2^t \in Z_2$, denoted by $p_{2t}^{\sigma}(\cdot | h_2^t)$ (resp., $\widetilde{p}_{2t}^{\sigma'}(\cdot | h_2^t)$).[1] In the equilibrium, at the information set h_2^t, player 2 believes that his endogenous signals will realize according to $p_{2t}^{\sigma}(\cdot | h_2^t)$. But if player 2 had known that player 1 was the normal type and played like the commitment type $\xi(\hat{\alpha}_1)$, then player 2 would predict his endogenous signals according to $\widetilde{p}_{2t}^{\sigma'}(\cdot | h_2^t)$.

[1] If h_2^t has probability 0 under P^{σ}, i.e. it is not reached in the equilibrium σ, then the one step ahead prediction is not well defined. But this does not matter because we will consider the average (over h_2^t) one step prediction errors.

For any $t \geq 1$, let $\widetilde{E}^{\sigma'}_{2t}[\,\cdot\,]$ denote the expectation over H_{2t} with respect to the probability measure $\widetilde{P}^{\sigma'}_{2t}$. The following lemma is a direct application of the chain rule.

Lemma B.1.2. *For all $t \geq 0$,*

$$\widetilde{E}^{\sigma'}_{2t}\left[d\left(\widetilde{p}^{\sigma'}_{2t}(\,\cdot\,|h^t_2)\|p^{\sigma}_{2t}(\,\cdot\,|h^t_2)\right)\right] \leq d\left(\widetilde{P}^{\sigma'}_{2,t+1}\middle\|P^{\sigma}_{2,t+1}\right) - d\left(\widetilde{P}^{\sigma'}_{2t}\middle\|P^{\sigma}_{2t}\right),$$

where $d\left(\widetilde{P}^{\sigma'}_{2,0}\middle\|P^{\sigma}_{2,0}\right) \equiv 0$.

Proof. Let $q_{2,t+1}(\,\cdot\,|h^t_2, z^t_2)$ (resp., $\widetilde{q}_{2,t+1}(\,\cdot\,|h^t_2, z^t_2)$) be the one step ahead prediction on his *exogenous* signals if he had observed his past private history h^t_2 and current period endogenous signal z^t_2, induced by $P^{\sigma}_{2,t+1}$ (resp., $\widetilde{P}^{\sigma'}_{2,t+1}$). Because Assumption 2.2.1 and Lemma B.1.1 implies $d\left(\widetilde{P}^{\sigma'}_{2t}\middle\|P^{\sigma}_{2t}\right) < \infty$ for all $t \geq 1$, applying chain rule twice yields

$$\begin{aligned}
& d\left(\widetilde{P}^{\sigma'}_{2,t+1}\middle\|P^{\sigma}_{2,t+1}\right) - d\left(\widetilde{P}^{\sigma'}_{2t}\middle\|P^{\sigma}_{2t}\right) \\
= \;& \widetilde{E}^{\sigma'}_{2t}\left[d\left(\widetilde{p}^{\sigma'}_{2t}(\,\cdot\,|h^t_2)\|p^{\sigma}_{2t}(\,\cdot\,|h^t_2)\right)\right] + E^{\dagger}_{2,t+1}\left[d\left(\widetilde{q}_{2,t+1}(\,\cdot\,|h^t_2, z^t_2)\|q_{2,t+1}(\,\cdot\,|h^t_2, z^t_2)\right)\right],
\end{aligned}$$

where $E^{\dagger}_{2,t+1}$ is with respect to the marginal distribution of $\widetilde{P}^{\sigma'}_{2,t+1}$ over $(Z_2 \times Y)^t \times Z_2$. The desired result is obtained by noting that the last term in the above expression is nonnegative because relative entropy is always nonnegative. □

Let $d^{\delta,\sigma}_{\xi(\hat{\alpha}_1)}$ be the expected average discounted sum of player 2's one step ahead prediction errors if player 1 is the normal type and he deviates to mimicking the commitment type $\xi(\hat{\alpha}_1)$

$$\begin{aligned}
d^{\delta,\sigma}_{\xi(\hat{\alpha}_1)} \;&\equiv\; \widetilde{E}^{\sigma'}\left[(1-\delta)\sum_{t=0}^{\infty}\delta^t d\left(\widetilde{p}^{\sigma'}_{2t}(\,\cdot\,|h^t_2)\|p^{\sigma}_{2t}(\,\cdot\,|h^t_2)\right)\right] \\
&=\; (1-\delta)\sum_{t=0}^{\infty}\delta^t \widetilde{E}^{\sigma'}_{2t}\left[d\left(\widetilde{p}^{\sigma'}_{2t}(\,\cdot\,|h^t_2)\|p^{\sigma}_{2t}(\,\cdot\,|h^t_2)\right)\right],
\end{aligned}$$

where δ is player 1's discount factor.

The next lemma, combining Lemma B.1.1 and Lemma B.1.2, provides an upper bound for $d_{\xi(\hat{\alpha}_1)}$.

Lemma B.1.3.

$$d^{\delta,\sigma}_{\xi(\hat{\alpha}_1)} \leq -(1-\delta)\mu\big(\xi(\hat{\alpha}_1)\big) + \lambda_{\xi(\hat{\alpha}_1)}.$$

Proof.

$$\begin{aligned}
d^{\delta,\sigma}_{\xi(\hat{\alpha}_1)} &\leq (1-\delta)\sum_{t=0}^{\infty}\delta^t\Big(d\big(\widetilde{P}^{\sigma'}_{2,t+1}\|P^{\sigma}_{2,t+1}\big) - d\big(\widetilde{P}^{\sigma'}_{2t}\|P^{\sigma}_{2t}\big)\Big)\\
&= (1-\delta)\sum_{t=0}^{\infty}\delta^t d\big(\widetilde{P}^{\sigma'}_{2,t+1}\|P^{\sigma}_{2,t+1}\big) - (1-\delta)\sum_{t=0}^{\infty}\delta^t d\big(\widetilde{P}^{\sigma'}_{2t}\|P^{\sigma}_{2t}\big)\\
&= (1-\delta)^2\sum_{t=1}^{\infty}\delta^{t-1} d\big(\widetilde{P}^{\sigma'}_{2t}\|P^{\sigma}_{2t}\big)\\
&\leq (1-\delta)^2\sum_{t=1}^{\infty}\delta^{t-1}\Big[-\log\mu\big(\xi(\hat{\theta})\big) + t\lambda_{\xi(\hat{\alpha}_1)}\Big]\\
&= -(1-\delta)\log\mu\big(\xi(\hat{\theta})\big) + \lambda_{\xi(\hat{\alpha}_1)},
\end{aligned}$$

where the first inequality comes from Lemma B.1.2 and the second inequality from Lemma B.1.1. □

An important feature of Lemma B.1.3 is that the upper bound on the expected prediction error is independent of P^{σ} and $\widetilde{P}^{\sigma'}$, which allows us to bound player 1's payoff in *any* Nash equilibrium.

Proof of Proposition 2.3.1. In equilibrium, at any information set $h_2^t \in \bigcup_{t\geq 0} H_{2t}$ that is reached with positive probability, $\sigma_2(h_2^t)$ is a best response to $E\big(\sigma_1(\xi, h_1^t)\big|h_2^t\big)$ and his one step ahead prediction on his endogenous signals is $p^{\sigma}_{2t}(\,\cdot\,|h_2^t)$. If player 1 is the normal type and he deviates to mimicking $\xi(\hat{\alpha}_1)$, the one step ahead prediction is $\widetilde{p}^{\sigma'}_{2t}(\,\cdot\,|h_2^t)$. Thus at any h_2^t with positive probability under $\widetilde{P}^{\sigma'}$, player 2 plays a $d\big(\widetilde{p}^{\sigma'}_{2t}(\,\cdot\,|h_2^t)\|p^{\sigma}_{2t}(\,\cdot\,|h_2^t)\big)$-entropy confirming best response to $\hat{\alpha}_1$.[2] Because σ is a Nash equilibrium, the deviation is not profitable. Hence in equilibrium, the payoff to the

[2]Because $\widetilde{P}^{\sigma'}$ is absolutely continuous with respect to P^{σ}.

normal type is at least as high as

$$
\begin{aligned}
& \widetilde{E}^{\sigma'}\Big[(1-\delta)\sum_{t=0}^{\infty}\delta^t u_1(a^t)\Big] \\
= \;& (1-\delta)\sum_{t=0}^{\infty}\delta^t \widetilde{E}_{2t}^{\sigma'}\Big[u_1\big(\hat{\alpha}_1, \sigma_2(h_2^t)\big)\Big] \\
\geq \;& (1-\delta)\sum_{t=0}^{\infty}\delta^t \widetilde{E}_{2t}^{\sigma'}\Big[\underline{V}_{\xi(\hat{\alpha}_1)}\Big(d\big(\widetilde{p}_{2t}^{\sigma'}(\,\cdot\,|h_2^t)\|p_{2t}^{\sigma}(\,\cdot\,|h_2^t)\big)\Big)\Big] \\
\geq \;& (1-\delta)\sum_{t=0}^{\infty}\delta^t \widetilde{E}_{2t}^{\sigma'}\Big[V_{\xi(\hat{\alpha}_1)}\Big(d\big(\widetilde{p}_{2t}^{\sigma'}(\,\cdot\,|h_2^t)\|p_{2t}^{\sigma}(\,\cdot\,|h_2^t)\big)\Big)\Big] \\
\geq \;& V_{\xi(\hat{\alpha}_1)}\Big(-(1-\delta)\log\mu\big(\xi(\hat{\alpha}_1)\big) + \lambda_{\xi(\hat{\alpha}_1)}\Big),
\end{aligned}
$$

where the second inequality comes from the definition of $V_{\xi(\hat{\alpha}_1)}$ and the last inequality from Jesen's inequality and Lemma B.1.3. Since the Nash equilibrium σ and the commitment type $\xi(\hat{\alpha}_1)$ are arbitrary, the result follows. □

Appendix C

Proofs for Chapter 3

C.1 Proof of Lemma 3.2.1 and 3.2.2

Proof of Lemma 3.2.1. Just notice that the characterization of the buyers' equilibrium strategies, i.e. (3.1), uniquely pins down the equilibrium strategy of the buyers σ_B^* in any Bayesian perfect equilibrium. Moreover, it is easy to see this strategy σ_B^* is a Bayesian strategy using market belief as the state variable: for any p_t, r_t and h^t, h^s, if $\pi(h^t) = \pi(h^s)$ then $\sigma_B^*(h^t, p_t, r_t) = \sigma_B^*(h^s, p_t, r_t)$. This implies that the monopolist in fact faces a Bayesian decision problem with market belief as the state variable. The solution to this Bayesian decision problem is precisely given by Bellman equation (3.3), which is formulated in terms of posterior belief cut-offs. Then it is obvious that the monopolist's strategy is a best response to σ_B^* after any history if and only if the induced cut-off after any history solves maximization problem given expected future payoff. This establishes the equivalence between the two different formulations. $\square$

Proof of Lemma 3.2.2. Monotonicity is straightforward.

Convexity: It is easy to see that the monopolist's value function (3.3) also solves the following Bellman equation

$$
\begin{aligned}
V(\pi) = \max_{r\in[\underline{r},\overline{r}]} &(1-\delta)\big[1-\pi F^h(r)-(1-\pi)F^l(r)\big] \\
&\times \Big[\frac{\pi r}{\pi r+(1-\pi)(1-r)}h+\frac{(1-\pi)(1-r)}{\pi r+(1-\pi)(1-r)}l-v\Big] \\
&+(1-\delta)\big[1-\pi F^h(r)-(1-\pi)F^l(r)\big]V\Big(\frac{\pi(1-F^h(r))}{\pi(1-F^h(r))+(1-\pi)(1-F^l(r))}\Big) \\
&+(1-\delta)\big[\pi F^h(r)+(1-\pi)F^l(r)\big]V\Big(\frac{\pi F^h(r)}{\pi F^h(r)+(1-\pi)F^l(r)}\Big).
\end{aligned}
$$

Hence for convexity, we only need to show for any convex $\tilde{V}$, the function $\hat{V}$ defined as

$$
\begin{aligned}
\tilde{V}(\pi) = \max_{r\in[\underline{r},\overline{r}]} &(1-\delta)\big[1-\pi F^h(r)-(1-\pi)F^l(r)\big] \\
&\times \Big[\frac{\pi r}{\pi r+(1-\pi)(1-r)}h+\frac{(1-\pi)(1-r)}{\pi r+(1-\pi)(1-r)}l-v\Big] \\
&+(1-\delta)\big[1-\pi F^h(r)-(1-\pi)F^l(r)\big]\hat{V}\Big(\frac{\pi(1-F^h(r))}{\pi(1-F^h(r))+(1-\pi)(1-F^l(r))}\Big) \\
&+(1-\delta)\big[\pi F^h(r)+(1-\pi)F^l(r)\big]\hat{V}\Big(\frac{\pi F^h(r)}{\pi F^h(r)+(1-\pi)F^l(r)}\Big) \quad \forall\pi\in[0,1].
\end{aligned}
$$

is also convex. Because the maximum of a class of convex functions is also convex, to show $\tilde{V}$ is convex, it suffices to show for each $r\in[\underline{r},\overline{r}]$, the objective function in the above equation is convex in π. The convexity of the mapping $\pi\mapsto(1-\delta)\big[1-\pi F^h(r)-(1-\pi)F^l(r)\big]\Big[\frac{\pi r}{\pi r+(1-\pi)(1-r)}h+\frac{(1-\pi)(1-r)}{\pi r+(1-\pi)(1-r)}l-v\Big]$ is straightforward. The convexity of the mappings $\pi\mapsto\big[1-\pi F^h(r)-(1-\pi)F^l(r)\big]\hat{V}\Big(\frac{\pi(1-F^h(r))}{\pi(1-F^h(r))+(1-\pi)(1-F^l(r))}\Big)$ comes from

$$
\begin{aligned}
\frac{\pi^\lambda F^h(r)}{\pi^\lambda F^h(r)+(1-\pi^\lambda)F^l(r)} =&\lambda\frac{\pi F^h(r)+(1-\pi)F^l(r)}{\pi^\lambda F^h(r)+(1-\pi^\lambda)F^l(r)}\frac{\pi F^h(r)}{\pi F^h(r)+(1-\pi)F^l(r)} \\
&+(1-\lambda)\frac{\tilde{\pi}F^h(r)+(1-\tilde{\pi})F^l(r)}{\pi^\lambda F^h(r)+(1-\pi^\lambda)F^l(r)}\frac{\tilde{\pi}F^h(r)}{\tilde{\pi}F^h(r)+(1-\tilde{\pi})F^l(r)}
\end{aligned}
$$

where π and $\tilde{\pi}$ are arbitrary market beliefs, $\lambda\in(0,1)$ and $\pi^\lambda=\lambda\pi+(1-\lambda)\tilde{\pi}$.

Lipschitz continuity: Notice $V(\pi) \geq V_m(\pi)$ for all $\pi \in [0,1]$ where

$$V_m(\pi) \equiv \max_{r\in[\underline{r},\overline{r}]} \big(1 - \pi F^h(r) - (1-\pi)F^l(r)\big) \\ \times \Big[\frac{\pi r}{\pi r + (1-\pi)(1-r)} h + \frac{(1-\pi)(1-r)}{\pi r + (1-\pi)(1-r)} l - v\Big]$$

is the monopolist's myopice value function. Moreover, we have $V(1) = V_m(1)$. Because V is increasing and convex, to show V is Lipschitz continuous, it suffices to show V_m is Lipschitz continuous. □

C.2 Proofs for Section 3.3

The proof of Lemma 3.3.1 is based on Smith, Sørensen, and Tian (2015) and adapted to the current setting. Let V be the monopolist's value function. Let $V^d(\pi)$ be V's right derivative if $\pi \in [0,1)$ and V's left derivative if $\pi = 1$. By Lemma 3.2.2, we know V^d exists. Moreover, we must have

$$V(\pi) = \max_{\tilde{\pi}} V^d(\tilde{\pi})(\tilde{\pi} - \pi) + V(\tilde{\pi}) \quad \forall \pi \in [0,1]. \tag{C.1}$$

For all $\pi \in (0,1)$ and $q \in (\underline{q}(\pi), \overline{q}(\pi))$, define

$$\widetilde{V}(q,\pi) \equiv (1-\delta)\big[1 - G(q|\pi)\big]\big[qh + (1-q)l - v\big] \\ + \delta\big[1 - G(q|\pi)\big]V\big(\pi^P(q,\pi)\big) + \delta G(q|\pi)V\big(\pi^N(q,\pi)\big).$$

The following lemma shows that $\widetilde{V}(\cdot,\pi)$ is absolutely continuous for all $\pi \in (0,1)$.

Lemma C.2.1. *For any $\pi \in (0,1)$ and $q \in (\underline{\pi}, \overline{\pi})$, define*

$$\widetilde{V}_1(q,\pi) \equiv (1-\delta)\big[1 - G(q|\pi)\big](h-l) - (1-\delta)g(q|\pi)[qh + (1-q)l - v] \\ - \delta\Big[V^d(\pi^P(q,\pi))(q - \pi^P(q,\pi)) + V(\pi^P(q,\pi))\Big]g(q|\pi) \\ + \delta\Big[V^d(\pi^N(q,\pi))(q - \pi^N(q,\pi)) + V(\pi^N(q,\pi))\Big]g(q|\pi).$$

Then $\widetilde{V}(q,\pi) = \int_{\underline{q}(\pi)}^{q} \widetilde{V}_1(\tilde{q},\pi)\mathrm{d}\tilde{q}$.

Proof. From Equation (C.1), we know

$$\begin{aligned}(1-G(q|\pi))V\big(\pi^P(q,\pi)\big) =&(1-G(q|\pi))\max_{\tilde\pi}\Big[V^d(\tilde\pi)(\pi^P(q,\pi)-\tilde\pi)+V(\tilde\pi)\Big]\\ =&\max_{\tilde\pi}(1-G(q|\pi))\Big[V^d(\tilde\pi)(\pi^P(q,\pi)-\tilde\pi)+V(\tilde\pi)\Big]\\ =&\max_{\tilde\pi}\Big[V^d(\tilde\pi)\int_q^{\overline{q}}\tilde q g(\tilde q|\pi)\mathrm{d}\tilde q - V^d(\tilde\pi)\tilde\pi(1-G(q|\pi))\\ &+(1-G(q|\pi))V(\tilde\pi)\Big]\\ =&\max_{\tilde\pi}\int_q^{\overline{q}(\pi)}\big[V^d(\tilde\pi)(\tilde q-\tilde\pi)+V(\tilde\pi)\big]g(\tilde q|\pi)\mathrm{d}\tilde q.\end{aligned}$$

By Corollary 4 and Theorem 2 in Milgrom and Segal (2002), we know the term $(1-G(q|\pi))V\big(\pi^P(q,\pi)\big)$ is absolutely continuous and its almost everywhere derivative is $-\big[V^d(\pi^P(q,\pi))(q-\pi^P(q,\pi))+V(\pi^P(q,\pi))\big]g(q|\pi)$ Similarly, we can show that $G(q|\pi)V(\pi^N(q,\pi))$ is also absolutely continuous and its almost everywhere derivative is $\big[V^d(\pi^N(q,\pi))(q-\pi^N(q,\pi))+V(\pi^N(q,\pi))\big]g(q|\pi)$. This proves the lemma. □

Proof of Lemma 3.3.1. Let q^* be a solution to the monopolist's Bellman equation (3.3). Assume $q^*(\underline{\pi}) > q^*(\overline{\pi})$ for some $\underline{\pi} < \overline{\pi}$. Then it is clear that $[q^*(\overline{\pi}), q^*(\underline{\pi})] \subset \mathrm{supp}G(\,\cdot\,|\underline{\pi}) \cap \mathrm{supp}G(\,\cdot\,|\overline{\pi})$. Fix $q \in [q^*(\overline{\pi}), q^*(\underline{\pi})]$. Because the information structure satisfies belief monotonicity, we know $(1-G(q|\overline{\pi}))/g(q|\overline{\pi}) \geq (1-G(q|\underline{\pi}))/g(q|\underline{\pi})$. Moreover, Lemma 7 in Smith, Sørensen, and Tian (2015) shows that

$$V^d(\pi^N(q,\overline{\pi}))(q-\pi^N(q,\overline{\pi}))+V(\pi^N(q,\overline{\pi})) \geq V^d(\pi^N(q,\underline{\pi}))(q-\pi^N(q,\underline{\pi}))+V(\pi^N(q,\underline{\pi}))$$

and

$$V^d(\pi^P(q,\overline{\pi}))(q-\pi^P(q,\overline{\pi}))+V(\pi^P(q,\overline{\pi})) \leq \big[V^d(\pi^P(q,\underline{\pi}))(q-\pi^P(q,\underline{\pi}))+V(\pi^P(q,\underline{\pi})).$$

Therefore we have

$$\widetilde{V}_1(q,\overline{\pi}) \geq \frac{g(q|\overline{\pi})}{g(q|\underline{\pi})}\widetilde{V}_1(q,\underline{\pi}) \quad \forall q \in [q^*(\overline{\pi}), q^*(\underline{\pi})].$$

Because the information structure satisfies belief monotonicity, Proposition 1 and Proposition 2 in Quah and Strulovici (2009) show that $q^*(\overline{\pi})$ is a solution to the Bellman equation 3.3 at $\underline{\pi}$ and $q^*(\underline{\pi})$ is a solution to the Bellman equation 3.3 at $\overline{\pi}$, completing the proof. □

Proof of Proposition 3.3.1. By Lemma 3.3.1 and Theorem 2.4.3 in Topkis (1998), we know there exists an increasing selection from the optimal cut-off correspondence. Let it be q^*. Then define $\sigma_M^*(h^t) = q^*(\pi(h^t))h + (1 - q^*(h^t))l - v$ for all h^t. Then from Lemma 3.2.1, it is easy to check that (σ_M^*, σ_B^*) is a Bayesian perfect equilibrium where σ_B^* satisfies (3.1). □

C.3 Proof of Proposition 3.4.1

The whole proof of Proposition 3.4.1 is divided into several lemmas.

Lemma C.3.1. *If $h > v > l$ and information structure is unbounded, then in any equilibrium, both C^P and C^N are degenerate.*

Proof. Because $\underline{r} = 0$ and $\overline{r} = 1$, the support of private posterior beliefs given any market belief $\pi \in (0,1)$ is always $[0,1]$, i.e. $\text{supp}G(\cdot|\pi) = [0,1]$ for all $\pi \in (0,1)$. Hence if $\pi \notin \{0,1\}$ and $\pi \in C^N$, by definition of C^N we know $q^*(\pi) = 1$ and the current period payoff to the monopolist is 0. Because buyer's behavior does not provide any information, the market belief will remain the same in the next period. This implies the monopolist will charge $q^*(\pi) = 1$ in all later periods, and thus $V(\pi) = 0$. But by charging a price $q' < 1$, the monopolist can always guarantee himself a strictly positive payoff, a contradiction. Now assume $\pi \notin (0,1)$ and $\pi \in C^P$. By definition of C^P, we know $q^*(\pi) = 0$. Then because $v > l$, the actual price charged by the monopolist is $l - v < 0$. By the same argument as above, we know $V(\pi) = l - v < 0$, a contradiction. □

Lemma C.3.2. *If $h > v > l$ and information structure is bounded, then in any equilibrium, C^N is non-degenerate.*

Proof. We show there exists $\underline{\pi} >$ such that $[0, \underline{\pi}] \subset C^N$. Choose any $\varepsilon > 0$ such that $v - \frac{\delta}{1-\delta}\varepsilon > l$. Given $\varepsilon > 0$, choose any $\underline{\pi} \in (0, 1)$ such that

$$h\frac{\underline{\pi}\overline{r}}{\underline{\pi}\overline{r} + (1-\underline{\pi})(1-\overline{r})} + l\frac{(1-\underline{\pi})(1-\overline{r})}{\underline{\pi}\overline{r} + (1-\underline{\pi})(1-\overline{r})} - v < -\frac{\delta}{1-\delta}\epsilon,$$

and

$$V\Big(\frac{\underline{\pi}\overline{r}}{\underline{\pi}\overline{r} + (1-\underline{\pi})(1-\overline{r})}\Big) < \epsilon.$$

Such $\underline{\pi}$ exists because $v - \frac{\delta}{1-\delta}\varepsilon > l$ and V is continuous with $V(0) = 0$. Then for any $\pi \in (0, \underline{\pi})$,

$$\begin{aligned} V(\pi) &= (1-\delta)\big[1 - G(q^*(\pi)|\pi)\big]\big[hq^*(\pi) + l(1 - q^*(\pi)) - v\big] \\ &+ \delta\big[1 - G(q^*(\pi)|\pi)\big]V\big(\pi^P(q^*(\pi), \pi))\big) + \delta G(q^*(\pi)|\pi)V\big(\pi^N(q^*(\pi), \pi)\big) \\ &\le (1-\delta)\big[1 - G(q^*(\pi)|\pi)\big]\big[h\overline{q}(\pi) + l(1 - \overline{q}(\pi)) - v\big] \\ &+ \delta\big[1 - G(q^*(\pi)|\pi)\big]V\big(\overline{q}(\pi))\big) + G(q^*(\pi)|\pi)V\big(\pi\big) \end{aligned}$$

where $\overline{q}(\pi) = \frac{\pi\overline{r}}{\pi\overline{r}+(1-\pi)(1-\overline{r})} \ge q^*(\pi)$ and the inequality comes from $\pi^P(q^*(\pi), \pi) \le \overline{q}(\pi)$, $\pi^N(q^*(\pi), \pi) \le \pi$ and the monotonicity of V. Because $\pi \in (0, \underline{\pi})$, by construction we know $V(\overline{q}(\pi)) < \varepsilon$. Then the above inequality inequality implies

$$\begin{aligned} &(1 - G(q^*(\pi)|\pi))V(\pi) \\ \le&(1-\delta)\big[1 - G(q^*(\pi)|\pi)\big]\big[h\overline{q}(\pi) + l(1 - \overline{q}(\pi)) - v\big] \\ &+ \delta\big[1 - G(q^*(\pi)|\pi)\big]\epsilon. \end{aligned}$$

If $1 - G(q^*(\pi)|\pi)) > 0$, this inequality reduces to

$$V(\pi) \le (1-\delta)\big[h\overline{q}(\pi) + l(1 - \overline{q}(\pi)) - v\big] + \delta\varepsilon < -\delta\varepsilon + \delta\varepsilon = 0.$$

But we know this is impossible, because the monopolist can always guarantee a payoff at least 0 by always setting the cut-off to $\pi\overline{r}/(\pi\overline{r} + (1-\pi)(1-\overline{r}))$. Hence we must have $1 - G(q^*(\pi)|\pi)) = 0$. This means

$$q^*(\pi) = \frac{\pi \underline{r}}{\pi \underline{r} + (1-\pi)(1-\underline{r})},$$

and thus $\pi \in C^N$. This proves there exists $\underline{\pi}$ such that $[0, \underline{\pi}] \subset C^N$. □

Lemma C.3.3. *In any equilibrium q^*, if $\underline{q}(\pi) = \frac{\pi \underline{r}}{\pi \underline{r} + (1-\pi)(1-\underline{r})}$ does not maximize the monopolist's myopic payoff for market belief $\pi \in (0,1)$, then $q^*(\pi) \neq \underline{q}(\pi)$.*

Proof. En route to a contradiction, assume there exists an equilibrium q^* such that $q^*(\pi) = \underline{q}(\pi)$ for some $\pi \in (0,1)$ but there exists $q \in \text{supp}G(\cdot|\pi)$ such that

$$\big[1 - G(q|\pi)\big]\big[qh + (1-q)l - v\big] > \underline{q}(\pi)h + (1 - \underline{q}(\pi))l - v.$$

Because $q^*(\pi) = \underline{q}(\pi)$, there will be no learning in current and later periods. Hence $V(\pi) = \underline{q}(\pi)h + (1 - \underline{q}(\pi))l - v$. But we have

$$\begin{aligned}
V(\pi) &= \underline{q}(\pi)h + (1 - \underline{q}(\pi))l - v \\
&< (1-\delta)\big[1 - G(q|\pi)\big]\big[qh + (1-q)l - v\big] + \delta(\underline{q}[\pi)h + (1 - \underline{q}(\pi))l - v\big] \\
&= (1-\delta)\big[1 - G(q|\pi)\big]\big[qh + (1-q)l - v\big] + \delta V(\pi) \\
&\leq (1-\delta)\big[1 - G(q|\pi)\big]\big[qh + (1-q)l - v\big] \\
&\quad + \delta\big[1 - G(q|\pi)\big]V\big(\pi^P(q,\pi)\big) + \delta G(q|\pi)V\big(\pi^N(q,\pi)\big),
\end{aligned}$$

where the second inequality comes from the fact that V is convex. This show that q is a profitable deviation at π, contradicting the assumption that q^* is an equilibrium. □

Lemma C.3.4. *If $h > v > l$, information structure is bounded and $f^l(\underline{r}) = 0$, then $C^P = \{1\}$.*

Proof. By Lemma C.3.3, we can show this by showing that for any $\pi \in (0,1)$, $\underline{q}(\pi) = \pi \underline{r} / \big[\pi \underline{r} + (1-\pi)(1-\underline{r})\big]$ is not the monopolist's myopic best response. The myopic payoff to the monopolist for any $q \in \text{supp}G(\cdot|\pi)$ is

$$\big[1 - G(q|\pi)\big]\big[qh + (1-q)l - v\big].$$

Using the fact that

$$G(q|\pi) = \int_{\frac{\pi r}{\pi r + (1-\pi)(1-r)} \le q} \big[\pi f^h(r) + (1-\pi) f^L(r)\big] \mathrm{d}r, \tag{C.2}$$

$f^h(r)/f^l(r) = \frac{r}{1-r}$ and f^l is continuous, the derivative of the myopic payoff with respect to q is

$$\big[1 - G(q|\pi)\big]\big[h - l\big] - g(q|\pi)\big[qh + (1-q)l - v\big].$$

When $q = \underline{q}(\pi)$, this boils down to

$$\big[h - l\big] - g(\underline{q}(\pi)|\pi)\big[\underline{q}(\pi)h + (1 - \underline{q}(\pi))l - v\big].$$

Using Equation (C.2), we can easily show $g(\underline{q}(\pi)|\pi) = 0$ if $f^l(\underline{r}) = 0$. Hence the derivative of the myopic payoff at $\underline{q}(\pi)$ is strictly positive, proving that $\underline{q}(r)$ is not a myopic best response. □

Lemma C.3.5. *If $h > v > l$, information structure is bounded and $f^l(\underline{r}) > 0$, then C^P is non-degenerate.*

Proof. We show that there exists $\overline{\pi} \in (0, 1)$ such that $[\overline{\pi}, 1] \subset C^P$. For any $\pi \in (0, 1)$ and $r \in [\underline{r}, \overline{r}]$, let $q(r, \pi)$ be the posterior belief about the high quality given market belief π and private signal r, i.e.

$$q(r, \pi) \equiv \frac{\pi r}{\pi r + (1-\pi)(1-r)}.$$

The whole proof of this lemma can be divided into four steps.

Step 1: For any $r \in (\underline{r}, \overline{r})$, there exists $\pi_r \in (0, 1)$ such that

$$\begin{aligned}
&(1-\delta)\big[q(\underline{r}, \pi)h + (1 - q(\underline{r}, \pi)l - v\big] + \delta V(\pi) \\
>&(1-\delta)\big[1 - G(q(r', \pi)|\pi)\big]\big[q(r', \pi)h + (1 - q(r', \pi))l - v\big] \\
&+ \delta\big[1 - G(q(r', \pi)|\pi)\big] V\big(\pi^P(q(r', \pi), \pi)\big) + \delta G(q(r', \pi)|\pi) V\big(\pi^N(q(r', \pi), \pi)\big)
\end{aligned}$$

for all $r > r'$, $\pi > \pi_r$.

Because V is increasing, by simple algebra, it is easy to see that a sufficient condition for this claim is that for any $r \in (\underline{r}, \overline{r})$, there exists $\pi_r \in (0,1)$ such that

$$\begin{aligned}&(1-\delta)G(q(r,\pi)|\pi)\big[q(\underline{r},\pi)h + (1-q(\underline{r},\pi)l - v\big]\\ >&(1-\delta)\big[1-G(q(r,\pi)|\pi)\big]\big[q(\overline{r},\pi) - q(\underline{r},\pi)\big](h-l)\\ &+\delta\big[1-G(q(r,\pi)|\pi)\big]\big[V\big(\pi^P(q(\overline{r},\pi),\pi)\big) - V\big(\pi\big)\big].\end{aligned}$$

But this is obviously true because the left hand side converges to $(1-\delta)F^h(r)(h-v) > 0$ and the right hand side converges to 0 as π goes to 1.

Step 2: For any $r \in (\underline{r}, \overline{r})$ and $L > 0$, there exists π_r^L such that for all $\pi > \pi_r^L$ and $r' \in [\underline{r}, r]$

$$g\big(q(r',\pi)|\pi\big) > L$$

where $g(\,\cdot\,|\pi)$ is the density of $G(\,\cdot\,|\pi)$.

Using Equation (C.2), we can show

$$g(q|\pi) = \frac{\pi(1-\pi)^2}{(1-q)\big[(1-\pi)q + \pi(1-q)\big]^2} f^L\Big(\frac{(1-\pi)q}{(1-\pi)q+\pi(1-q)}\Big).$$

Hence

$$g\big(q(r',\pi)|\pi\big) = \frac{\big(\pi r' + (1-\pi)(1-r')\big)^2}{\pi(1-\pi)(1-r')} f^l(r') \geq \frac{M}{\pi(1-\pi)},$$

where

$$M = \min_{r'\in[\underline{r},r]} \frac{\big(\pi r' + (1-\pi)(1-r')\big)^2}{(1-r')} f^l(r').$$

Because $f^l(\underline{r}) > 0$, f^l has full support and is continuous, $M > 0$. Therefore, for any $L > 0$, there exists π_r^L such that

$$g\big(q(r',\pi)|\pi\big) > L \quad \forall r' \in [\underline{r}, r].$$

Step 3: For any $r \in (0,1)$ and any $\varepsilon > 0$, there exists π_r^ε such that

$$\pi^P\big(q(r',\pi),\pi\big) - \pi \leq \varepsilon G\big(q(r',\pi)|\pi\big) \quad \forall r' \in [0,r], \pi \in [\pi_r^\varepsilon, 1].$$

Notice

$$\begin{aligned}\pi^P\big(q(r',\pi),\pi\big)-\pi &= \pi^P\big(q(r',\pi),\pi\big)-\pi^P\big(q(\underline{r},\pi),\pi\big)\\ &= \int_{q(\underline{r},\pi)}^{q(r',\pi)} \frac{\partial \pi^P(\tilde{q},\pi)}{\partial \tilde{q}}\mathrm{d}\tilde{q}\\ &= \int_{q(\underline{r},\pi)}^{q(r',\pi)} \frac{\pi^P(\tilde{q},\pi)-\tilde{q}}{1-G(\tilde{q}|\pi)}g(\tilde{q}|\pi)\mathrm{d}\tilde{q}.\end{aligned}$$

We have for all $\tilde{q}\in[q(\underline{r},\pi),q(r,\pi)]$,

$$\frac{\pi^P(\tilde{q},\pi)-\tilde{q}}{1-G(\tilde{q}|\pi)} \le \frac{q(\overline{r},\pi)-q(\underline{r},\pi)}{1-G\big(q(r,\pi)|\pi\big)} \to 0,$$

as $\pi\to 1$. Therefore, for any $r\in(0,1)$ and $\epsilon>0$, there exists $\pi_r^{\varepsilon}\in(0,1)$ such that

$$\pi^P\big(q(r',\pi),\pi\big)-\pi \le \varepsilon\int_{q(\underline{r},\pi)}^{q(r',\pi)} g(\tilde{q}|\pi)\mathrm{d}\tilde{q} = \varepsilon G\big(q(r',\pi),\pi\big), \quad \forall r'\in[\underline{r},r], \pi\in[\pi_r^{\varepsilon},1].$$

Step 4: There exists $\overline{\pi}\in(0,1)$ such that in any equilibrium q^*, $q^*(\pi)=q(\underline{r},\pi)$ for all $\pi\in[\overline{\pi},1]$.

Fix any arbitrary $r\in(\underline{r},\overline{r})$. By Lemma 3.2.2, we know there exists some $a>0$ such that $|V(\pi)-V(\pi')|\le a|\pi'-\pi|$ for all $\pi',\pi\in[0,1]$. Pick any $\varepsilon>0$ such that $h-v-\frac{\delta}{1-\delta}a\varepsilon>0$. Pick any $L>0$ such that $L(h-v-\frac{\delta}{1-\delta}a\varepsilon)>2(h-l)$. Let $\overline{\pi}=\max\{\pi_r,\pi_r^L,\pi_r^{\varepsilon},\tilde{\pi}\}$ where $\pi_r,\pi_r^L,\pi_r^{\varepsilon}$ are from Steps 1, 2 and 3 respectively and $\tilde{\pi}\in(0,1)$ is chosen so that $q(\underline{r},\pi)h+(1-q(\underline{r},\pi))l-v-\frac{\delta}{1-\delta}a\varepsilon>\frac{1}{2}(h-v-\frac{\delta}{1-\delta}a\varepsilon)$ for all $\pi>\tilde{\pi}$.

From Step 1, we know for any $\pi>\overline{\pi}$, $q^*(\pi)<q(r,\pi)$. We now show $q^*(\pi)=q(\underline{r},\pi)$. Fix $\pi\in[\overline{\pi},1]$ and pick any $r'\in(\underline{r},r]$. By construction, we know

$$g\big(\tilde{q}|\pi)\big)\Big[q(\underline{r},\pi)h+(1-q(\underline{r},\pi)l-v-\frac{\delta}{1-\delta}a\varepsilon\Big]>h-1 \quad \forall\tilde{q}\in[q(\underline{r},\pi),q(r',\pi)].$$

Therefore

$$\int_{q(\underline{r},\pi)}^{q(r',\pi)}\Big[g\big(\tilde{q}|\pi)\big)\big[q(\underline{r},\pi)h+(1-q(\underline{r},\pi))l-v-\frac{\delta}{1-\delta}a\varepsilon\big]-(h-1)\Big]\mathrm{d}\tilde{q}>0,$$

or equivalently

$$\begin{aligned}
&(1-\delta)G\big(q(r',\pi)|\pi\big)\big[q(\underline{r},\pi)h+(1-q(\underline{r},\pi))l-v\big] \\
>&(1-\delta)\big[q(r',\pi)-q(\underline{r},\pi)\big](h-l)+\delta G\big(q(r',\pi)|\pi\big)a\varepsilon \\
\geq&(1-\delta)\big[1-G\big(q(r',\pi)|\pi\big)\big]\big[q(r',\pi)-q(\underline{r},\pi)\big](h-l) \\
&+\delta\big[1-G\big(q(r',\pi)|\pi\big)\big]a\big[\pi^P\big(q(r',\pi),\pi\big)-\pi\big].
\end{aligned}$$

But this will imply

$$\begin{aligned}
&(1-\delta)\big[q(\underline{r},\pi)h+(1-q(\underline{r},\pi))l-v\big]+\delta V(\pi) \\
>&(1-\delta)\big[1-G\big(q(r',\pi)|\pi\big)\big]\big[q(r',\pi)h+(1-q(r',\pi))l-v\big] \\
&+\delta\big[1-G\big(q(r',\pi)|\pi\big)\big]V\big(\pi^P(q(r',\pi),\pi)\big)+\delta G\big(q(r',\pi)|\pi\big)V\big(\pi^N(q(r',\pi),\pi)\big),
\end{aligned}$$

Because this inequality holds for all $r' \in (\underline{r}, r]$, we know $q^*(\pi) = q(\underline{r}, \pi)$, completing the proof. □

Bibliography

Banerjee, A. V. (1992): "A simple model of herd behavior," *The Quarterly Journal of Economics*, pp. 797–817.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, pp. 992–1026.

Bolton, G., B. Greiner, and A. Ockenfels (2013): "Engineering trust: reciprocity in the production of reputation information," *Management Science*, 59(2), 265–285.

Bose, S., G. Orosel, M. Ottaviani, and L. Vesterlund (2006): "Dynamic monopoly pricing and herding," *The RAND Journal of Economics*, 37(4), 910–928.

——— (2008): "Monopoly pricing in the binary herding model," *Economic Theory*, 37(2), 203–241.

Boylan, R. T. (1992): "Laws of large numbers for dynamical systems with randomly matched individuals," *Journal of Economic Theory*, 57(2), 473–504.

Cai, H., Y. Chen, and H. Fang (2009): "Observational larning: evidence from a randomized natural field experiment.," *American Economic Reivew*, 99, 864–882.

Cover, T., and J. Thomas (2006): *Elements of information theory.* Wiley-interscience.

Cripps, M. W., G. J. Mailath, and L. Samuelson (2004): "Imperfect monitoring and impermanent reputations," *Econometrica*, 72, 407–432.

——— (2007): "Disappearing Private Reputations in Long-run Relationships," *Journal of Economic Theory*, 134, 287–316.

Dellarocas, C., and C. A. Wood (2008): "The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias," *Management Science*, 54(3), 460–476.

Ekmekci, M. (2011): "Sustainable reputations with rating systems," *Journal of Economic Theory*, 146, 479–503.

Ekmekci, M., O. Gossner, and A. Wilson (2012a): "Impermanent types and permanent reputations," *Journal of Economic Theory*, 147(1), 162–178.

Ekmekci, M., O. Gossner, and A. Wilson (2012b): "Impermanent Types and Permanent Reputations," *Journal of Economic Theory*, 147(1), 162–178.

Feller, W. (1968): *An Introduction to Probability Theory and Its Applications, Volume 1.* Wiley, New York, 3rd edn.

Fudenberg, D., and D. K. Levine (1989): "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57, pp. 759–778.

——— (1992): "Maintaining a Reputation when Strategies are Imperfectly Observed," *The Review of Economic Studies*, 59, pp. 561–579.

Glicksberg, I. L. (1952): "A further generalization of the Kakutani fixed point theorem, with application to Nash equilibrium points," *Proceedings of the American Mathematical Society*, 3(1), 170–174.

Gossner, O. (2011a): "Simple bounds on the value of a reputation," *Econometrica*, 79(5), 1627–1641.

——— (2011b): "Simple Bounds on The Value of a Reputation," *Econometrica*, 79(5), pp. 1627–1641.

Guarino, A., H. Harmgart, and S. Huck (2011): "Aggregate information cascades," *Games and economic behavior*, 73, 167–185.

Herrera, H., and J. Hörner (2013): "Biased social learning," *Games and Economic Behavior*, 80, 131–146.

Jehiel, P., and L. Samuelson (2012): "Reputation with analogical reasoning*," *The Quarterly Journal of Economics*, 127(4), 1927–1969.

Kets, W. (2011): "Robustness of equilibria in anonymous local games," *Journal of Economic Theory*, 146(1), 300–325.

Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson (1982): "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of economic theory*, 27, 245–252.

Kreps, D. M., and R. Wilson (1982): "Reputation and imperfect information," *Journal of economic theory*, 27(2), 253–279.

Liu, Q. (2011): "Information acquisition and reputation dynamics," *The Review of Economic Studies*, 78(4), 1400–1425.

Liu, Q., and A. Skrzypacz (2014): "Limited records and reputation bubbles," *Journal of Economic Theory*, 151, 2–29.

Mailath, G. J., and L. Samuelson (2001): "Who wants a good reputation?," *Review of Economic Studies*, 68, 415–441.

——— (2006a): *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, New York.

——— (2006b): *Repeated Games and Reputations: Long-run Relationships*. Oxford University Press, USA.

Milgrom, P., and J. Roberts (1982): "Predation, reputation, and entry deterrence," *Journal of economic theory*, 27(2), 280–312.

Milgrom, P., and I. Segal (2002): "Envelope theorems for arbitrary choice sets," *Econometrica*, 70(2), 583–601.

Milgrom, P., and C. Shannon (1994): "Monotone comparative statics," *Econometrica: Journal of the Econometric Society*, pp. 157–180.

Monzón, I., and M. Rapp (2014): "Observational learning with position uncertainty," *Journal of Economic Theory*, 154, 375–402.

Moscarini, G., and M. Ottaviani (2001): "Price competition for an informed buyer," *Journal of Economic Theory*, 101(2), 457–493.

Nosko, C., and S. Tadelis (2015): "The limits of reputation in platform markets: an empirical analysis and field experiment," NBER Working Paper No. 20830.

Osborne, M. J., and A. Rubinstein (1994): *A Course In Game Theory*. MIT Press, Cambridge, Massachusetts.

Phelan, C. (2006): "Public trust and government betrayal," *Journal of Economic Theory*, 130(1), 27–43.

Quah, J. K.-H., and B. Strulovici (2009): "Comparative statics, informativeness, and the interval dominance order," *Econometrica*, 77(6), 1949–1992.

Smith, L., and P. Sørensen (2000): "Pathological outcomes of observational learning," *Econometrica*, 68(2), 371–398.

Smith, L., P. Sørensen, and J. Tian (2015): "Informational herding, optimal experimentation, and contrarianism," working paper.

Topkis, D. M. (1998): *Supermodularity and complementarity*. Princeton university press.

Wall, H. S. (1967): *Analytic Theory of Continued Fractions*. Chelsea Publishing Company, New York.

Wiseman, T. (2009): "Reputation and exogenous private learning," *Journal of Economic Theory*, 144(3), 1352–1357.