Essays on Communication in Game Theory

Takakazu Honryo

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2012

ABSTRACT

Essays on Communication in Game Theory

Takakazu Honryo

This dissertation consists of essays on communication in game theory. The first chapter develops a model of dynamic persuasion. A sender has a fixed number of pieces of hard evidence that contain information about the quality of his proposal, each of which is either favorable or unfavorable. The sender may try to persuade a decision maker (DM) that she has enough favorable evidence by sequentially revealing at most one piece at a time. Presenting evidence is costly for the sender and delaying decisions is costly for the DM. I study the equilibria of the resulting dynamic communication game. The sender effectively chooses when to give up persuasion and the DM decides when to make a decision. Resolving the strategic tension requires probabilistic behavior from both parties. Typically, the DM will accept the sender's proposal even when she knows that the sender's evidence may be overall unfavorable. However, in a Pareto efficient equilibrium, the other type of error does not occur unless delays costs are very large. Furthermore, the sender's net gain from engaging in persuasion can be negative on the equilibrium path, even when persuasion is successful. we perform comparative statics in the costs of persuasion. I also characterize the DM's optimal stochastic commitment rule

and the optimal non-stochastic commitment rule; compared to the communication game, the former yields a Pareto improvement, whereas, the latter can leave even the DM either better or worse off.

The second chapter studies a unidimensional Hotelling-Downs model of electoral competition with the following innovation: a fraction of candidates have "competence", which is unobservable to voters. Competence means the ability to correctly observe a policy-relevant state of the world. This structure induces a signaling game between competent and incompetent candidates. We show that in equilibrium, proposing an extreme platform serves as a signal about competence, and has a strictly higher winning probability than that of the median platform. Polarization happens and the degree of it depends on how uncertain the state is and how much political candidates are office-motivated.

The third chapter examines the dynamic extension of Che, Dessein, and Kartik (2011). They study strategic communication by an agent who has non-verifiable private information about different alternatives. The agent does not internalize the principal's benefit from her outside option. They show that a pandering distortion arises in communication. This chapter studies the long-run consequence of their model when a new agent-principal pair is formed in each period, and principals in later periods may learn some information from predecessors' actions. I characterize the conditions under which effective communication between principal and agent can continue in perpetuity.

Contents

List of Figures

Acknowledgements

Navin Kartik has been my superb advisor at every single stage of my research project and at the job market. His influence is everywhere in my dissertation: some chapters are directly related to or based on his research. His course on strategic information transmission invited me to this exciting research area, and his smartness helped me a lot to cover my language barrier in research meetings. Because I am a big fan of his research, his comments on my research convinced me that I was on the right track, and made my Ph.D. life worry-free. He also saved the second chapter of my dissertation: I initially thought that the idea was not good enough, but he recommended that I make it a paper, and it eventually became a paper that I like a lot. His ability to give insightful comments, deepen analysis, and his broad knowledge that makes it possible to connect one research to another, showed me a model of a great professional scholar. Whether I should come and study at U.S. or not was a really big problem for me, but now, from ex-post point of view, my decision turned out to be totally right for having had him as my advisor.

I thank Yeon-Koo Che, who always actively holds reading group with students and involves himself in education, for his comments on all the chapters. I thank

Bogachan Celen for teaching me game theory from the very first step. It was really special time to go to his office and discuss game theory for hours. As a big fan of the idea of Storable vote, I thank Alessandra Casella for her advice, and in particular for reading my second chapter and gave me very detailed comments. I thank Qingmin Liu for his help on the first chapter, and his support during the job market. I hope to repay (even partially) my huge debt to them by contributing to enrichment of the field of micro economic theory.

I did not decide to pursue Ph.D. in economics without the influence of Makoto Yano. Although I chose a quite different research area from him, the quality of his research will be the one that I will always aim at.

I am grateful to Massimo Morelli, Bernard Salanie, and Satoru Takahashi for their comments on my dissertation chapters.

Among my friends, I particularly thank Hitoshi Shigeoka, Yoichi Sugita, and Kensuke Teshima. I am extremely fortunate to find three best friends, each have different research interest, in my grad school life, which made my Ph.D. time the happiest time in my entire life. Also, it is great that I will be a colleague of Yuya Takahashi at Mannheim, who decided to pursue Ph.D. in U.S. together. I would also like to thank my friends and colleagues at Columbia for their helpful discussions and for making my time in graduate school memorable. At the risk of forgetting some names, I thank Mariessa Herrman, Marcos Nakaguma, Maria Jose Prados, Izzet Yildiz, and Zhanna Zhanabekova.

Finally, my parents, who came to New York City five times for sightseeing, have supported the entire period of my graduate study. I dedicate this dissertation to them in token of affection and gratitude.

CHAPTER 1

Dynamic Persuasion

1.1. Introduction

Persuasion is the act of influencing someone to undertake a particular action, or, more generally, to form a certain belief. Successful persuasion takes time and is costly for both parties: the speaker exerts effort to present convincing arguments or information and, in turn, the listener reflects upon or inspects these carefully. A typical process of persuasion may involve a back-and-forth interaction where the speaker gradually presents a series of arguments up until when the listener is either sufficiently convinced by the speaker or has decided that the speaker's case lacks merit.

This chapter is an attempt to understand some essential features of the dynamics of persuasion. As an example, consider an entrepreneur who is trying to convince a venture capitalist (VC) to invest in his startup. The VC only wants to invest if the startup is sufficiently likely to succeed. The onus is on the entrepreneur to explain and validate a number of different aspects of the project that justify investment. Of course, the VC will scrutinize each argument, possibly hiring third parties to do so. In a stylized way, the process may unfold as follows: the entrepreneur presents a set of facts about the project that the VC scrutinizes, and then the VC decides to either invest, walk away, or request further explanation; and the process repeats, with the entrepreneur deciding whether to comply or give up on persuading the VC.

Observe that this process has a dynamic element of "matching pennies" as in game theory. If the VC knows that the entrepreneur only brings profitable plans, she just rubber-stamps his proposals, rather than paying the various costs incurred to scrutinize the plans. On the other hand, once the entrepreneur thinks that the VC will not carefully scrutinize his proposal, he may bring plans lacking in precision. This, in turn, generates the VC's incentive to carefully scrutinize. This simple story, which has the flavor of the matching pennies game, shows that each player has an incentive to outfox his or her opponent.

Having this salient nature of persuasion in mind, this chapter describes the dynamic process of persuasion in a formal game theoretic model. A sender (persuader, speaker) may try to persuade a decision maker (receiver, listener) that she has enough favorable evidence for his proposal by sequentially communicating evidence by paying the communication cost. He can also remain silent, which incurs no cost for him. At each period, the decision maker chooses whether to require another piece of evidence that delays her decision making, or not. Hence she chooses to require evidence as long as she can expect that there is an informational gain from doing so. We show that the equilibrium involves probabilistic decision making from both parties. The decision maker may make a decision before she gets enough information from the sender, so she may make the wrong decision.

Our model succeeds in providing some essential features of the dynamics of persuasion. Each time the sender communicates a piece of evidence, the decision maker updates her belief about the sender's proposal, and she accepts the proposal with a strictly positive probability. As the game proceeds, the decision maker accumulates more and more information and the probability that she makes the wrong decision decreases. A good sender, who has enough good evidence, continues to persuade by showing evidence until his proposal is accepted. However, he may pay too much communication costs so he may lose ex-post even though he was successful at persuasion. A bad sender tries to persuade the decision maker with some probability. When the decision maker judges that the cost of requiring a further piece of evidence exceeds the additional informational benefit, she accepts the proposal for sure and the process of persuasion terminates.

The fact that the equilibrium involves probabilistic decision-making stems from the game's similarity to the matching pennies game. If the decision maker does not accept the sender's proposal until a certain amount of evidence is shown, then the sender never tries to persuade unless he has enough good evidence. However this implies that the first piece of good evidence already screens out the bad sender and the decision maker loses the incentive to check the rest of the evidence. If the decision maker cannot make a "commitment to listen", they should use mixed strategy in order to get around this strategic tension, as in the matching pennies game.

General characterization of the equilibrium demonstrates the following results. There is a lower bound of the probability of immediate acceptance every time the sender communicates evidence (either good or bad). Actually, this lower bound is the acceptance probability that makes the sender's communication cost equal his immediate expected gain. Also, silence never meets immediate acceptance, which tells us that only a costly message has a persuasive power.[1] Finally, and rather obviously, the decision maker accepts the proposal for sure only after the sender shows a good piece of evidence.

Although there is a plethora of equilibria, we characterize the set of Pareto equilibria that are not Pareto dominated by other equilibria and, furthermore, the best equilibrium for the decision maker, which is unique. Generally, it is possible to have an equilibrium that involves intuitively inessential stages such as the sender remains silent, which incurs no cost for him, and the decision maker just waits for the sender to start talking. We show that any Pareto efficient equilibrium excludes such redundant stages. Specifically, we show that in a Pareto efficient equilibrium, the sender never communicates bad evidence, sender's silence meets immediate rejection, and the decision maker's acceptance probability immediately after seeing a piece of good evidence is either maximized or minimized among all possible ways of constructing an equilibrium. It is also shown that in a Pareto efficient equilibrium, once the process of persuasion starts, the decision maker does

[1]Hence, even if we endow the sender with a set of cheap messages as available message, these only have the same role as silence.

not make the error of rejecting a good proposal. We further show that in the best equilibrium the decision maker requires the largest amount of good evidence in order to accept for certainty: intuitively, increasing the amount of good evidence necessary for persuasion discourages bad senders from trying to persuade.

The uniqueness of the best equilibrium for all parameter values enables us to pin down a reasonable benchmark on which we conduct comparative static analysis. We particularly examine the effects of two players' costs of communication on their expected payoffs and expected duration of persuasion. We show that a decrease in the costs of communication for the decision maker (delay costs) benefits her through two effects. The first one is the direct effect. The second one, which is indirect effect, benefits the decision maker by discouraging the bad sender from trying to persuade. It also reduces the sender's expected payoff because it increases the length of time for acceptance. With respect to the effects of the cost of communication for the sender, on top of some intuitive results, a decrease in it also lengthens the expected time of acceptance.

While in the main analysis, we consider that the decision maker cannot make any form of commitment, we also characterize her optimal commitment problem. First, we show that the optimal commitment mechanism takes a stochastic form, in which the decision maker attaches the highest probability of acceptance to each node that prevents the bad type sender from trying to persuade. Furthermore, it can be shown that this does not harm the sender relative to the (best) equilibrium, which means that the optimal stochastic commitment can be a Pareto improvement. This is because the commitment makes it possible to avoid the case in which the bad sender tries, but fails, to persuade the decision maker. In this case, which necessarily happens with a positive probability in the equilibrium, both pay wasteful communication costs.

In order to consider the case that it is hard to make stochastic commitment, we also examine a limited commitment method in which the decision maker can make only the non-stochastic commitment of requiring a predetermined amount of evidence. We show that even this limited commitment is beneficial for the decision maker relative to the best equilibrium when the sender's communication cost is low. However, interestingly, playing the best equilibrium is better when the sender's cost is high. This result comes from the fact that the equilibrium of the game may make the sender pay more communication cost ex-post than the gain from persuasion, which allows the decision maker to extract more information from him. In contrast, in the non-stochastic commitment, the sender is perfectly knowledgeable about the outcome of the persuasion at the beginning and, hence, it is impossible to make him show a large amount of evidence.

1.1.1. Related Literature

Our model is most closely related to the literature on strategic communication with verifiable messages, which is also called persuasion games. The most important benchmark was developed by Grossman (1981) and Milgrom (1981). They study a persuasion model in which the sender is not required to tell the truth in a precise manner, and show that we have complete unraveling of information. Shin (1994) studies a persuasion game in which the decision maker does not know how precise the sender's information is, and shows that unraveling of information breaks down. Verrechia (1983) incorporates a cost of information transmission for the sender to those models, and also shows that it prevents complete unravelling of information.[2] In the current study, complete unraveling of information does not happen because the decision maker is not willing to pay the cost of communication up to the point that full information is obtained.[3] Forges and Koessler (2008) characterize the sets of equilibrium payoffs achievable with unmeditated communication in persuasion games with multi-stages. Hörner and Skrzypacz (2011) study a dynamic model of verifiable information transmission in which a seller can transmit information gradually as the buyer makes payment for it.

In using a setting in which the sender gets a collection of binary signals about the state, this chapter is related to Dziuda (2007) and Quement (2010). Dziuda (2007) offers a model in which a sender tries to persuade the decision maker to make a particular action by revealing verifiable information. In her model, the persuader

[2]Kartik, Ottaviani, and Squintani (2007) and Kartik (2009) study a model in which the sender's information is not verifiable but he bears a cost of lying and, hence, information is costly to falsify.

[3]Che and Kartik (2009) build a model of verifiable information, but the sender has to pay the cost of information acquisition. They analyze the problem of who to ask for advice, given the fact that full information revelation from the sender does not happen.

may be either a strategic agent or a truth teller. Quement (2010) constructs a model in which the sender has either a small amount of evidence or a large amount of evidence. Their question is whether the strategic sender has an incentive to reveal unfavorable signals or not, and they show that the sender may do so. Although our model does not pay a particular attention for the question of whether the sender communicates unfavorable signals or not, it also has an equilibrium in which the sender communicates unfavorable signals (and it is proved to be inefficient). This is because even sending an unfavorable signal incurs the cost and thus it signals that the sender is a good type, who has confidence of being able to persuade, ultimately.

There are some studies that investigate a problem of persuasion as a mechanism design problem. In Glazer and Rubinstein (2004), the decision maker is allowed to check one piece of evidence of the sender's proposal, and they study mechanisms that maximize the probability that the decision maker accepts the sender's request, if and only if it is justified. Sher (2010) generalizes the Glazer and Rubinstein's model in a way that both static and dynamic persuasion can be considered, and characterizes the relation between them. Kamenica and Gentzkow (2010) demonstrate that a sender can induce his favorite action from the decision maker by ingeniously designing the signal structure, by which they can make Bayesian updating of information. This chapter also addresses a similar problem of how the decision maker should design her acceptance rule by examining the optimal commitment problem.

Analytically, this study is closely related to a variant of the games of attrition where players use mixed strategies to resolve the dynamic strategic tension.[4] Hendricks, Weiss, and Wilson (1988) study a war of attrition in a complete information model. Kreps and Wilson (1982b), and Ordover and Rubinsten (1986) consider models of attrition with asymmetric information. Although they build models on zero-sum payoff structure while we do not, their models are analytically similar to our model in the sense that some of the players must play mixed strategies to have a gradual revelation of types in an equilibrium. An important difference is that in their studies, one of the informed players has a dominant strategy for the duration of the game and, as a consequence, all nodes of the game are reached with a positive probability. However, in our study, no player has a dominant strategy and all players' incentives are endogenously determined in the game. One more important difference is that in the variant of war of attrition, duration works as an indirect signal about player's private information, which can be cost of fighting, cost of failing the agreement, time preference, and so on, through showing how much they can "burn money". In our model, in contrast, private information is gradually revealed by the process of the decision maker directly asking the sender. Baliga and Ely (2010) consider a model in which a principal uses torture to extract information from an informed agent. In equilibrium, the informed agent

[4]A notable difference is that we formulate the game in discrete time, rather than continuous time that is standard in game of attrition.

reveals information gradually, initially resisting and facing torture, but eventually he concedes.

This chapter is also related to the literature on cheap-talk communication in dynamic models. Sobel (1985) develops a dynamic cheap talk model[5] in which the sender is either a friend or an enemy of the decision maker, and examines the problem of how long the sender should spend on constructing his reputation and when he should deceive the decision maker. Aumann and Hart (2003), and Krishna and Morgan (2004) show that multiple exchanges of messages can convey more information than a single message. Eso and Fong (2008) study a model with multiple senders where the decision maker can choose when to make her decision. They show that the threat of costly delay can induce instantaneous full revelation of information.

This chapter is organized as follows. Section 1.2 introduces the basic structure of the model. In section 1.3, we provide analysis on the simplest example of the model. In Section 1.4, we provide general cauterization of equilibrium. Section 1.5, we do comparative static analysis. In section 1.6, we examine commitment problems. Proof of the theorems can be found in the Appendix.

[5]For a benchmark model of the cheap-talk game, see Crawford and Sobel (1982).

1.2. Model

There are two players: a sender (persuader) and a decision maker, or DM hereafter.[6] The sender has a proposal that he would like the DM to accept. The quality of the sender's proposal θ (the state) is either 1 or -1; there is a common prior over the state. The sender does not observe the state θ,[7] but he receives $N \in \mathbb{Z}$ pieces of evidence that contain information about the quality of the proposal. Each piece of evidence is either good (G) or bad (B). The vector of evidence $e \in \{G, B\}^N$ is drawn from a distribution $g(e|\theta)$. Given the environment, we also have the probability distribution over θ, conditional on the realization of e. We assume that the pieces of evidence are interchangeable in the sense that the $\mathbb{E}[\theta|e]$ depends only on the number of pieces of good evidence in e. Given the assumption, we have the expected value of θ, conditional on the realization of j pieces of good evidence among N, and denote it by $\mathbb{E}[\theta|j]$. We assume that $\mathbb{E}[\theta|j]$ is increasing with j, which means that more good evidence makes the prospect of the proposal better. In order to exclude trivial cases, we assume that $\mathbb{E}[\theta|0] < 0$ and $\mathbb{E}[\theta|N] > 0$.

Denote by ξ the threshold number of good pieces of evidence that makes the expected value of θ higher than zero, that is,

$$\mathbb{E}[\theta|\xi - 1] < 0 \leq \mathbb{E}[\theta|\xi].$$

[6]Throughout, we use female pronouns for the decision maker and male pronouns for the sender.

[7]In our model, it does not matter at all whether we assume that the sender observes θ or not.

Furthermore, denote by f the unconditional density over the realization of the number of good pieces of evidence; the sender's *type*. The DM wants to accept the sender's proposal if $\theta = 1$ and reject if $\theta = -1$ and, hence, she cares about the sender's type. Everything, except the realization of the sender's type is common knowledge.

To illustrate our setting, as an example, think of the following simple scenario which is taken from the literature on strategic voting.[8] The prior probability that the state is 1 is $1/2$. There are two pieces of evidence, i.e., $N = 2$. Each piece of evidence is independent from each other. Conditional on $\theta = 1$, the probability that a realization of a piece of evidence is G is $p > 1/2$, and conditional on $\theta = -1$, it is $1 - p$. Then, it follows that

$$\mathbb{E}[\theta|0] = \frac{1-2p}{p^2 + (1-p)^2}, \ \mathbb{E}[\theta|1] = 0, \text{ and } \mathbb{E}[\theta|2] = \frac{2p-1}{p^2 + (1-p)^2},$$

Hence, $\mathbb{E}[\theta|2] > \mathbb{E}[\theta|1] > \mathbb{E}[\theta|0]$ and ξ is one. Also, $f(0) = f(2) = \frac{1}{2}\{p^2 + (1-p)^2\}$ and $f(1) = 1 - p^2 - (1-p)^2$. Our setting allows more general cases, relative to this example, in a sense that we do not necessarily assume that each piece of evidence is independent from all others.

[8] See Feddersen and Pesendorfer (1998) for an example.

1.2.1. Dynamic Game of Persuasion

In the process of the game, the decision maker's turn and the sender's turn alternate. At each turn of the DM, she has three choices: whether she accepts, rejects, or continues, which is interpreted as requiring a piece of evidence from the sender. At each turn of the sender, he has three choices: communicating the DM about a good piece of evidence, a bad piece of evidence, or being silent. It is assumed that the sender cannot reveal more than one piece of evidence at a time, which is understood to be a technological constraint of communication. We can also think that it is extremely costly to communicate multiple evidence at a time. The game is terminated once the DM chooses to accept or reject.

The formal description of the model is as follows. Time is discrete and extends from 0 to ∞ that is denoted by $t \in T = \{0, 1, 2, .., \infty\}$. Before everything starts, Nature draws $\theta \in \{-1, 1\}$ and, conditional on the realization, it chooses the sender type, the number of pieces of good evidence the sender has. The number j is the sender's private knowledge. In our model, it is assumed that the sender is not informed about the realization of θ, although it does not matter at all for the analysis. At period 0, the decision maker chooses one from $\{A, R, C\}$, where A, R, and C correspond to accept, reject, and continue (require a piece of evidence), respectively. If C is chosen, the game proceeds to period 1. In period 1, first the sender chooses $m_1 \in \{G, B, S\}$ under the condition that he can choose G (B) only when $j \geq 1$ ($j \leq N - 1$). Here, G and B mean to show a good or bad piece of evidence, respectively, and S means that the sender remains silent.[9] He can show a good (bad) piece of evidence only when he has at least one of it. Then, the communication takes place and the DM chooses one from $\{A, R, C\}$ and, in the case that C is chosen, the game proceeds to period 2. Now, in the beginning of period 2, the sender chooses $m_2 \in \{G, B, S\}$ under the condition that m_2 can be G only when $j \geq 2$ if $m_1 = G$ and $j \geq 1$ if $m_1 \neq G$. We have the symmetric condition for B as well. The rest of the game is described in a similar manner. The game terminates once the DM chooses either A or R.

Message history at period t is a sequence of messages communicated up to period t, and it is denoted with superscript by m^t. The set of all histories at period t is $M^t = \times_t \{G, B, S\}$, and the set of all histories is $M = \cup M^t$. Then, define function $N_G : M \to \{1, 2, .., N\}$, $N_B : M \to \{1, 2, .., N\}$ and $N_S : M \to \{1, 2, .., N\}$ as the number of G, B, and S along message history m^t, respectively[10]. Obviously, we have $N_G(m^t) + N_B(m^t) + N_S(m^t) = t$. In the following analysis, the set of available message for type j sender after message history m^t is denoted by $M(m^t, j)$, that is $S \subset M(m^t, j)$ for all (m^t, j) and

$$G \in M\left(m^t, j\right) \text{ iff } j > N_G\left(m^t\right) \text{ and } B \in M\left(m^t, j\right) \text{ iff } N - j > N_B\left(m^t\right).$$

[9] We can also change the model by allowing sender to send a cheap message from a finite set of cheap messages, without adding any change to the results.

[10] More precisely, $N_G(m^t) = |\{k | m_k = G,\ k \leq t\}|$, $N_B(m^t) = |\{k | m_k = B,\ k \leq t\}|$, and $N_S(m^t) = |\{k | m_k = S,\ k \leq t\}|$.

Therefore, $M(m^t, j)$ cannot contain G (B) if the sender runs out of good (bad) pieces of evidence to communicate on the history m^t.

In our model, persuasion is costly for both players. We can simply think that it is costly because it takes up valuable time and there are cognitive costs that they have to pay to make the DM understand the sender's explanation. Specifically, we want to think that communication is costly for the DM because it delays his decision making, and it is costly for the sender because formulating or explaining evidence to the DM is costly due to cognitive costs. In this sense, we will use the term "communication cost" in a broad sense that includes delay cost.

We can also take the interpretation of Dewatripont and Tirole (2005)'s observation, which states that information is neither hard nor soft initially, but the degree of softness is endogenously changed. Only by combining the mutual effort of the two sides they turn the information into the hard type. If we take this interpretation, we assume that the degree of softness is zero-one[11]. To make things simple, we assume that the cost of communicating a piece of evidence is fixed for both sides. Thus, the communication technology for our model is specified as follows:

Communication cost for the DM.

[11]In Dewatripont and Tirole (2005), in contrast, the level of effort, which can increase the probability of being able to make information hard, is chosen by both sides. They examine the problem of moral hazard in team with that setting.

The (one time) cost of communication for the DM is represented by a function $\eta : \{G, B, S\} \to \mathbb{R}$, where

$$1 > \eta(G) = \eta(B) > 0 \text{ and } \eta(S) > 0.$$

Communication cost for the sender.

The (one time) cost of communication for the sender is represented by a function $\delta : \{G, B, S\} \to \mathbb{R}$, where

$$\delta(G) = \delta(B) > 0 \text{ and } \delta(S) = 0.$$

These say that communicating a piece of evidence G or B is costly for the DM as well as the sender and, in particular, communicating a piece of good evidence incurs strictly positive cost for both. Silence is also costly for the DM,[12] while it is not for the sender. Although it is possible to work on a model of positive silence cost for the sender, the assumption simplifies some of the mathematical expressions that appear later. Assumptions of $\eta(G) = \eta(B)$ and $\delta(G) = \delta(B)$ are purely for notational simplicity, and it is straightforward to extend the model by relaxing those assumptions.

We simply denote $\eta(G)$ (hence, also $\eta(B)$) by η, $\eta(S)$ by η_S, and $\delta(G)$ $(\delta(B))$ by δ. The communication costs that two players have to pay depend on how many

[12] It is also possible to choose a model setting in which silence does not incur cost for the DM. We chose the current setting because it generates inessential multiplicity of equilibrium.

times pieces of evidence or silence are communicated, multiplied by the communication cost. To shorten the notation, we define the functions that represent the costs of communication along the message history m^t, as follows:

$$C_{DM}\left(m^t\right) = \eta\{N_G\left(m^t\right) + N_B\left(m^t\right)\} + \eta_S N_S\left(m^t\right)$$

for the decision maker and

$$C_S\left(m^t\right) = \delta\{N_G\left(m^t\right) + N_B\left(m^t\right)\}.$$

for the sender. As soon as the DM takes an action, both of the players get their respective payoffs. The DM's (expected) payoff when the seeder type is j, which is denoted by $U_{DM}\left(a, j, m^t\right)$, depends on the particular action (accept or reject) taken by the DM, the type of sender, and the communication history after which the DM takes action:

$$U_{DM}\left(A, j, m^t\right) = \mathbb{E}[\theta|j] - C_{DM}\left(m^t\right) \text{ and } U_{DM}\left(R, j, m^t\right) = -C_{DM}\left(m^t\right).$$

When the DM accepts the proposal, her payoff depends on the sender type through the term $\mathbb{E}[\theta|j]$, which should be interpreted that the actual payoff of the decision maker is θ and its expected value is taken.[13] If the DM rejects the proposal, she has

[13]More precisely, the DM's utility depends on the state, action, and message history that is written as $U_{DM}\left(A, \theta, m^t\right) = \theta - C_{DM}\left(m^t\right)$ and $U_{DM}\left(R, \theta, m^t\right) = -C_{DM}\left(m^t\right)$.

an outside option that ensures her payoff of zero, and just pays her communication cost.

The sender's payoff, which is denoted by $U_S(A, m^t)$, depends only on the particular action taken by the DM and the communication history, after which the DM takes the action:

$$U_S\left(A, m^t\right) = V - C_S\left(m^t\right) \text{ and } U_S\left(R, m^t\right) = -C_S\left(m^t\right),$$

where $V \geq \delta$. Hence, the sender's payoff is V, which is the gain from persuading the DM, minus the communication cost if the DM accepts his proposal. It implies that the sooner he can persuade the DM, the higher his payoff is. It is possible that even if he could eventually persuade the DM, the communication cost is larger than the gain of persuasion V. On the other hand, he just pays the communication cost when the DM ends up with rejecting the proposal.

Hence, in our model, the cost of communication, which can be interpreted as a time cost, appears in the players' payoffs in an additively separable form. An alternative setting is one in which players' payoffs are discounted as time goes by. This setting, however, cannot generate the equilibrium that we will characterize; in such a setting the sender does not have an incentive to give up persuasion because his payoff just shrinks and never becomes negative. On the other hand, it is possible to model the DM's payoff in a discounted form and still get the same type of equilibrium, because even in such a setting, she faces the same trade-off between prompt decision making and information collection.

Now, we define the strategies of two players. The sender's (behavior) strategy is a probability measure $\alpha(\cdot, m^t, j)$ over available messages $M(m^t, j)$, parameterized by (m^t, j). It represents the type j sender's strategy after message history m^t, and $\alpha(m, m^t, j)$ represents the probability that he chooses a particular message $m \in \{G, B, S\}$. The strategy of period 1 is denoted by $\alpha(\cdot, \varnothing, j)$, by using a convention of notation $m^0 = \varnothing$. On the other hand, the DM's (behavior) strategy is a probability measures β over $\{A, R, C\}$, parameterized by m^t. Her strategy at period 0 is $\beta(\cdot, \varnothing)$.

We introduce notations and definitions to be used in the subsequent analysis. As the game proceeds, the DM's belief about the sender type evolves. Her belief, which is parametrized by message history m^t, is represented by a vector of function $B_n : M \to [0, 1]$ for $n = 0, 1, 2, .., N$ such that $\sum_{n=0}^{N} B_n(m^t) = 1$, that is, $B_j(m^t)$ is the probability that the DM attaches to the event, the sender type being j, after communication history m^t.

Given a sender's strategy α, we can define the probability that a particular message history is followed, that is

$$\varphi\left(m^t\right) = \sum_{j=0}^{N} f(j)\, \Pi_{\tau=1}^{t} \alpha\left(m_\tau, m^{\tau-1}, j\right).$$

Given the DM as well as the sender's strategy, we can define the set of message history that can be reached with strictly positive probability

$$\Delta = \{m^t \mid \sum_{j=0}^{N} f\left(j\right) \Pi_{s=1}^{t} \alpha\left(m_s, m^{s-1}, j\right) \cdot \beta\left(C, m^{s-1}\right) > 0\}.$$

We simply call elements of Δ, on-equilibrium message history.

In the following analysis, we use the following notations for the ease. The notation (m^t, m) reads "a message history such that m^t is followed by m". In particular, $(m^t, G) \in M^{t+1}$ represents message history m^t followed by G (Also, $(m^t, B) \in M^{t+1}$ should be read similarly). Furthermore, we denote by $G^t \in M^t$ the message history at period t that contains only G.

1.2.2. Equilibrium

Our solution concept is that of perfect Bayesian equilibrium, as is defined in Fudenberg and Tirole (1991, Definition 8.2).[14] This requires that after each history of messages $m^t \in M$, the DM maximizes her expected payoff given her belief about sender's type and their future play of the game, and also the sender maximizes his expected payoff given the DM's strategy.

In order to formally define the equilibrium, we first define the value function of the players. In our game, the decision of each period necessarily depends upon the decisions of the next period, and that in turn depends on the decision of the

[14] Their definition is for finite multistage games. Here, instead, the game has infinite stages and, hence, the definition follows a slight generalization of it.

following period, and so on. The value function we will define makes it possible to summarize all the information about the future play of the game that is necessary for making the current decision.

We start by defining the value function for the DM. In order to do this, let $\varphi(|m^t)$ be a probability distribution function over $\{G, B, S\}$, parameterized by $m^t \in M$, which can be interpreted as the DM's belief about next period's messages she will hear from the sender, should she continue. We say that a function $V_{DM} : M \to \mathbb{R}$ is a value function for the DM given (φ, B) if, for all $m^t \in M$,

$$V_{DM}\left(m^t\right) = \tag{1.1}$$

$$\max\{\max_{a\in\{A,R\}} \sum_{j=0}^{N} B_j\left(m^t\right) U_{DM}(a, j, m^t), \sum_{m\in\{G,B,S\}} \varphi(m|m^t) V_{DM}\left(m_t, m\right)\}$$

and

$$\lim_{t\to\infty} V_{DM}(m^t) = -\infty \text{ for all } \{m_t\}_{t=0}^{\infty}. \tag{1.2}$$

The definition of value function (1.1) says that the DM's value of history m^t is the higher one of the expected payoff when she makes decision immediately after message history m^t, and the expected value for waiting for one more period. The next condition (2) is understood to be the counterpart of "no-Ponzi game condition" in dynamic optimization problems in our model. In a typical formulation of a consumer's dynamic optimization problem, the no-Ponzi game condition ensures that the consumer cannot keep borrowing money over time and accumulating debt and, thereby, makes his utility arbitrary large. Condition (2) is reminiscent of that restriction in our model, which is necessary to pin down the value function for the DM; without it, the uniqueness of the value function is not ensured. Note that (1.2) is the same as requiring $\lim_{t\to\infty} V_{DM}(m^t) = -C_{DM}(m^t)$, because silence is costly for the DM and, hence, $\lim_{t\to\infty} C_{DM}(m^t) \to \infty$ for all sequence of history $\{m^t\}_{t=1}^{\infty}$. We have the following lemma, which states that once we are given the sender's strategy and the DM's belief, the value function for the DM is uniquely determined.

Lemma 1. *Given* (φ, B), V_{DM} *is uniquely determined.*

Similarly, we can define the value function for the sender. Contrary to the value function for the DM, sender's value function should be parameterized by his type. We say that a function $V_S : M \times N \to \mathbb{R}$ is a value function for the sender type j, given the DM's strategy β if

$$\begin{aligned} V_S\left(m^t, j\right) &= \beta\left(A, m^t\right) U_S\left(A, m^t\right) + \beta\left(R, m^t\right) U_S\left(R, m^t\right) \\ &\quad + \beta\left(C, m^t\right) \max_{m\in M(m^t, j)} V_S\left(\left(m^t, m\right), j\right) \end{aligned} \tag{1.3}$$

and

$$\lim_{t\to\infty} V_S\left(m^t, j\right) = -\lim_{t\to\infty} \delta\{N_G\left(m^t\right) + N_B\left(m^t\right)\} \text{ for all } \{m^t\}_{t=0}^{\infty}. \tag{1.4}$$

The max operator in the right hand side of (3) subsumes the fact that the sender behaves optimally at the next period. We also have no-Ponzi game condition as well. Then, we have the same lemma as when we defined the value function of the DM.

Lemma 2. *Given β, V_S is uniquely determined.*

Corollary 1. *Value function V_{DM} satisfies $-C_{DM}\left(m^t\right) \leq V_{DM}\left(m^t\right) \leq 1 - C_{DM}\left(m^t\right)$ for all m^t and value function V_S satisfies $-C_S\left(m^t\right) \leq V_S\left(m^t, j\right) \leq V - C_S\left(m^t\right)$.*

With the above preparations, we can define the equilibrium. We focus on following conditions for a pair of strategies and the DM's belief $(\alpha, \beta, B, \varphi)$.

D1. The optimality of the sender's strategy at every history of messages:

$$\alpha\left(m, m^t, j\right) > 0 \text{ only when } m \in \arg \max_{m \in M(m^t, j)} V_S\left(\left(m^t, a\right), j\right).$$

D2. The optimality of the DM's strategy at every history of messages:

$$\begin{aligned} \beta\left(C, m^t\right) &> 0 \text{ only when } V_{DM}\left(m_t\right) = \mathbb{E}[V_{DM}\left(m^{t+1}\right)|m^t)], \\ \text{and } a &\in \{A, R\},\ \beta\left(a, m^t\right) > 0 \text{ only when } V_{DM}\left(m_t\right) = \mathbb{E}[U(a, j, t)|m^t]. \end{aligned}$$

D3. Bayes' rule for the belief of the DM (B, φ) : For all $m^t \in M$,

$$\varphi\left(m_{t+1}|m^t\right) = \sum_{n=0}^{N} B_n\left(m^t\right)\alpha\left(m_{t+1}, m^t, j\right),$$

and if there is some j such that $\alpha\left(m_t, j, m^{t-1}\right) > 0$ and $B_j\left(m^{t-1}\right) > 0$,

$$B_j\left(m^t\right) = \frac{B_j\left(m^{t-1}\right)\alpha\left(m_t, m^{t-1}, j\right)}{\sum_{n=0}^{N} B_n\left(m^{t-1}\right)\alpha\left(m_t, m^{t-1}, n\right)} \text{ and } B_j\left(m^1\right) = \frac{f\left(j\right)\alpha\left(m_1, \varnothing, j\right)}{\sum_{n=0}^{N} f\left(n\right)\alpha\left(m_1, \varnothing, n\right)}.$$

$$B_j\left(m^t\right) = 0 \text{ for all } j < N_G\left(m^t\right) \text{ and } B_j\left(m^t\right) = 0 \text{ for all } j > N - N_B\left(m^t\right).$$

Our equilibrium is defined by those three conditions.

Definition 1. *A pair $(\alpha, \beta, B, \varphi)$ is a perfect Bayesian equilibrium iff it satisfies D1-D3.*

The first condition D-1 requires that the each time the sender chooses what to show, he chooses the one that maximizes his value. Note that, this must hold not only for message histories that are reached with strictly positive probability (on-equilibrium history), but also the histories that are not supposed to reach with positive probability (off-equilibrium history). D-2 requires the same kind of behavior for the DM. She chooses to continue only when it maximizes her value, in which case her value $V_{DM}\left(m_t\right)$ is equal to $\mathbb{E}[V_{DM}\left(m^{t+1}\right)|m^t)]$, and the same applies for the choices of accept and reject.

Note that D-3 is stronger than simply using Bayes' rule in the usual fashion, since it applies to updating from period t to period $t+1$ when messages history

m^t has probability zero, i.e., $m^t \notin \Delta$. The motivation for this requirement is that if $B_j(m^t)$ represents the DM's beliefs given m^t, and players follow their strategies at $t+1$, the DM should use Bayes' rule to form his belief in period $t+1$.[15]

The final requirement in D-3 simply says that the DM assigns zero probability to the type of sender who has a strictly smaller number of pieces of good (or bad) evidence than already shown. In the terminology of incomplete information game, the set

$$\{(j, m^t) \,|\, j \geq N_G(m^t) \text{ and } N - j \geq N_B(m^t)\} \subset \{0, 1, .., N\} \times M$$

is the information set for the DM after getting message m^t and, hence, the DM has to put all the probability mass in this set.

Note how the two belief functions B and φ play different roles in the DM's decision making. The belief function B, which shows the DM's belief over how good the proposal is, is relevant for choosing whether to accept or reject, if she has to make a decision immediately. On the other hand, the belief function φ, which shows the DM's beliefs about the sender's behavior at the next period, is relevant for choosing whether to decide immediately or to continue.

We conclude this section by showing some immediate results that follow almost directly from the definition of the equilibrium. The first one says that once the sender communicates sufficient number of pieces of good evidence, the DM accepts

[15]For more discussion about the requirements, see Fudenberg and Tirole (1991).

the proposal with certainty, and the sender just remains silent afterwards (thus, such a node should be off-equilibrium). The proof is straightforward and, hence, omitted:

Claim 1. *1. In any equilibrium, for all m^t such that $N_G(m^t) \geq \xi,\ \beta(A, m^t) = 1$.*

2. In any equilibrium, for all m^t such that $N_G(m^t) \geq \xi,\ \alpha(S, m^t, j) = 1$ for all j.

Since the DM, after verifying that the sender's proposal has a sufficient number of good pieces of evidence, already knows that her optimal action is to accept the proposal irrespective of the realizations of the rest of evidence, she does not pay more communication costs and reveals the rest of evidence. On the other hand, knowing that the DM will accept the proposal, the sender does not communicate remaining evidence by incurring the communication cost.

Given an equilibrium, let Ξ be the set of on-equilibrium history of termination with acceptance, that is, $m^\tau \in \Xi$ if and only if $\beta(A, m^\tau) = 1$ and $m^\tau \in \Delta$. From the definition, $\beta(A, m^s) < 1$ for all m^s that is a sub-history of m^τ. The next result shows that the DM accepts the proposal for certain on equilibrium, only after she is shown qualifying evidence.

Proposition 1. *For all $m^\tau = (m^{\tau-1}, m_\tau) \in \Xi$, it must hold that $m_\tau = G$.*

This follows because if there is a message history such that $(m^{\tau-1}, B) \in \Xi$ or $(m^{\tau-1}, S) \in \Xi$, even the lowest type sender among all types who may follow $m^{\tau-1}$ can get accepted at period τ, which implies that there is no screening of a bad sender taking place at period τ. This contradicts the fact that the DM chooses continue after $m^{\tau-1}$.

Remark 1. *One may think that we should impose more restrictions on off-equilibrium belief than when we are working on the usual signaling games. This is because, in our game, the DM's decision to take an action or continue crucially depends on her belief after the current period. In particular, when she decides to take some action and, thereby, terminates the game, that decision must be based on her off-equilibrium behavior of herself and the sender and, moreover, even their off-equilibrium behavior also depends on further off-equilibrium behavior. Therefore, one may want to use the concept of sequential equilibrium (Kreps and Wilson (1982a)), rather than perfect Bayesian equilibrium, just because it imposes more restrictions on off-equilibrium belief of the players. However, in our game it can be shown that those two equilibrium concepts coincide in a fundamental sense. In order to show this, let us define the usage of the term, outcome equivalence. Let $O := M \times \{A, R\}$ be the set of pairs of message history and the DM's action. Then, the outcome of the game is a probability distribution over O. We say that that two different strategies pairs $(\alpha', \beta', \varphi', B')$ and $(\alpha'', \beta'', \varphi'', B'')$ are outcome equivalent if they induce the same outcome. Note that, in such a case we have*

$\beta'(\cdot, m^t) = \beta''(\cdot, m^t)$ on every $m^t \in \Delta'$ and $\alpha'(m, j, m^t) = \alpha''(m \cdot j, m^t)$ for every m if $j \in P(m^t)$ and $m^t \in \Delta'$ such that $\beta'(C, m^t) > 0$, where $\Delta' \subset M$ is the set of nodes that can be reached with strictly positive probability (the set of on-equilibrium history in equilibrium $(\alpha', \beta', \varphi', B')$), and $\Delta'' \subset M$ is defined in a similar manner. It is easy to see that those imply $\Delta' = \Delta''$. We will call PBE $(\alpha, \beta, \varphi, B)$ a sequential equilibrium if there is a sequence of totally mixed strategy $(\alpha^\lambda, \beta^\lambda, \varphi^\lambda, B^\lambda)$, with $\lambda \in \mathbf{N}_+$ such that for each $m^t \in M$ and j, it converges to $(\alpha(m^t, j), \beta(m^t), \varphi(m^t), B(m^t)) \in R^{9+N}$. Now we have the following claim.

Claim 2. *For every PBE, there is an outcome equivalent sequential equilibrium.*

1.3. An Example

This section is devoted to the analysis of the special case in which the number of pieces of evidence is two and every piece of evidence should be good for the expected value of the proposal becomes positive, that is, $N = \xi = 2$. Although this is a special case, it is useful for getting the idea of the construction of the equilibrium and it provides fundamental properties that are shared with more general cases.

The first observation is that the DM's strategy of continuing until two pieces of good evidence is communicated is not supported as an equilibrium. This is because this naive strategy makes the type 1 sender give up persuasion by silence from the beginning, because he knows that two pieces of good evidence are necessary to persuade the DM. However, it makes the DM strictly better to accept immediately after the first piece of good evidence, because it already screened out low type sender (type 0 and type 1). This implies that equilibrium necessarily involves mixing strategies to resolve the tension.

In order to focus on the most interesting case, we impose the following assumptions.

$$\mathbb{E}[\theta|j \geq 1] \geq 0^{16} \text{ and } -\mathbb{E}[\theta|1]f(1) \geq \eta f(2) + \eta_S f(1). \tag{1.5}$$

The second condition says that the cost of communication is low enough, compared to the loss from accepting type 1 sender's proposal. Roughly speaking, the DM is willing to pay the communication cost if she can screen out type one sender when she knows that the sender is either type 1 or 2.

Even in this special case, we have a plethora of equilibria. The next proposition characterizes one of those, where the reason we focus on it is fully discussed in the next section (it is actually the *ex-ante* best equilibrium for the DM). In the statement of the theorem, we omit the description of off-equilibrium behaviors, because it is straightforward to specify those. Remember that the second element of the sender's strategy α is a message history, and the third element represents the type of the sender.

Proposition 2. *A pair of strategies that satisfies the following is an equilibrium.*

[16]This is rewritten as $f(1)\mathbb{E}[\theta|1] + f(2)\mathbb{E}[\theta|2] \geq 0$, which is compatible with $\xi = 2$.

1. Type 2 sender communicates good pieces of evidence in row: $\alpha(G, \varnothing, 2) = \alpha(G, G, 2) = 1$*, and type 0 sender chooses silence at period 1:* $\alpha(S, \varnothing, 0) = 1$.

2. Type 1 sender mixes at period 1:

$$\alpha(G, \varnothing, 1) = c \text{ and } \alpha(S, \varnothing, 1) = 1 - c, \text{ where } c = -\frac{f(2)\eta}{f(1)(\mathbb{E}[\theta|1] + \eta_S)}$$

3. At period 2, the DM accepts if she has been communicated two pieces of good evidence, and rejects otherwise:

$$\beta(A, (G, G)) = 1 \text{ and } \beta\left(R, m^2\right) = 1 \text{ if } m^2 \neq (G, G).$$

4. At period 1, the DM mixes between continuing and acceptance if she has been communicated a piece of good evidence, and rejects otherwise:

$$\beta(A, G) = \delta/V, \ \beta(C, G) = 1 - \delta/V, \text{ and } \beta(R, m_1) = 1 \text{ if } m_1 \neq G.$$

5. At period 0, the DM continues if $W \geq \max\{0, \mathbb{E}[\theta]\}$*, rejects if* $0 > \max\{W, \mathbb{E}[\theta]\}$*, and accepts if* $\mathbb{E}[\theta] > \{W, 0\}$*, where*

$$W = cf(1)(\mathbb{E}[\theta|1] - \eta) + f(2)(\mathbb{E}[\theta|2] - \eta) - \{f(0) + (1 - \alpha) f(1)\}\eta_S.$$

The second period strategies are easy to see. The DM accepts the sender's proposal if the sender communicates the second piece of good evidence again and rejects otherwise, which induces the sender to communicate the last piece of good evidence if he still has it. For the first period strategies, after checking one piece of good evidence, the DM mixes between accepting and continuing. The probability that she accepts is δ/V, which makes the sender type 1 be indifferent between trying persuasion (by communicating good evidence) and giving up by being silent. On the other hand, the probability that type one sender tries persuasion is set in a way that the DM is indifferent between accepting and continuing and thereby screening it out at period 1. Note that if this probability is too low, at period one after checking one piece of good evidence, the DM is sure enough that the sender is type 2 and she strictly prefers to accept, and if it is too high she strictly prefers to continue. The expression c, which is the type one sender's trial probability $\alpha\left(G,\varnothing,1\right)$, follows from the condition[17]

$$-\frac{cf\left(1\right)}{cf\left(1\right)+f\left(2\right)}\mathbb{E}[\theta|1]=\frac{cf\left(1\right)}{cf\left(1\right)+f\left(2\right)}\eta_S+\frac{f\left(2\right)}{cf\left(1\right)+f\left(2\right)}\eta.$$

The left-hand-side, the conditional probability that the sender is type 1 after communicating one piece of good evidence is multiplied by the expected loss, is the benefit of continuing one more time. The right hand side is the expected cost from communicating one more time, given the sender's strategy. Those two must be

[17] Alternatively, we can write it as

$$\frac{f\left(2\right)}{cf\left(1\right)+f\left(2\right)}\mathbb{E}[\theta|2]+\frac{cf\left(1\right)}{cf\left(1\right)+f\left(2\right)}\mathbb{E}[\theta|1]=\frac{f\left(2\right)}{cf\left(1\right)+f\left(2\right)}\left(\mathbb{E}[\theta|2]-\eta\right)+\frac{cf\left(1\right)}{cf\left(1\right)+f\left(2\right)}\eta_S,$$

where the left hand side represents the DM's expected payoff from accepting the sender's proposal after checking a single piece of good evidence, while the right hand side is her expected payoff from continue and screen type one sender out.

equal, because the DM must be indifferent between acceptance and communicating one more time.

At period zero, if the benefit from proceeding to period 1 is higher than the expected payoff from accept or reject without communication, the DM chooses continue. In such a case, we have W, which is characterized in the proposition, becomes $V_{DM}(\varnothing)$.

An important point to note is that at period one, it is optimal for type 2 sender to communicate a piece of good evidence, because he is sure to be able to persuade the DM. This is so even when $2\delta > V$, so that the communication cost he ends up paying is larger than V. This implies that he is expecting "success with regret" to happen with some probability at the beginning of the game, because at period 1 after being required to show one more piece of evidence, his first communication cost is sunk and responding to the DM's request and showing the second good evidence becomes optimal.

The specific equilibrium provided in Proposition 2 has some special characteristics that we focus on in the next section. First, a piece of bad evidence is never communicated on-equilibrium. Second, the acceptance probability after communicating a good evidence is δ/V or 1. Finally, silence meets immediate rejection. Although there are other equilibria that do not satisfy those properties, we will discuss in the next section that the equilibria that have those properties, which more plausible relative to other equilibria.

An important note is that we can also have an equilibrium in which only one piece of good evidence is needed to persuade the DM. If

$$f(1)(\mathbb{E}[\theta|1]-\eta)+f(2)(\mathbb{E}[\theta|2]-\eta)-f(0)\,\eta_S \geq \max\{0,\mathbb{E}[\theta]\},$$

the DM chooses to continue at time zero even if she knows that she is able to screen only type zero sender out. In such an equilibrium, both type one and two senders have success with persuasion by communicating only one piece of good evidence. Obviously, the equilibrium payoff is lower for the DM and higher for the sender, relative to the equilibrium of Proposition 1 (this fact is generalized in subsection 4.3).

In the rest of this section, we investigate more about the equilibrium in our example. First we do some comparative statics with respect to the parameter values η and δ. We start it by looking at the effect of δ.

Proposition 2 shows that the sender's cost δ has no effect on the DM's expected payoff. On the other hand, it has negative effect on the sender's expected payoff $\mathbb{E}_j[V_S(\varnothing,j)]$. To see this, think that we increase δ. It has no effect on type 0 sender's payoff, because it does not participate in the persuasion process. Also, it has no effect on type 1 sender's expected payoff, because the increase of the cost is exactly offset by period 1's acceptance probability. Type 2 sender's expected payoff, however, will be decreased because period $2'$s acceptance probability is still one and, thus, does not fully compensate the burden of the increase in the cost. It is easily seen that the increase in sender's cost decreases the expected time of DM's decision making through the increase in the acceptance probability at period 1.

We next discuss the comparative statics with respect to the DM's cost of communication, η. Proposition 2 demonstrates that it actually has no effect on the equilibrium payoff of the sender, $\mathbb{E}_j[V_S(\varnothing, j)]$, as long as (1.5) is satisfied, because it does not affect the acceptance probability at period one and two.

It is obvious that η has a strictly negative relationship with $V_{DM}(\varnothing)$. An interesting fact is that when η decreases, the DM can enjoy not only direct effect as well as indirect effect of the decrease. It is seen by the following relation:

$$\begin{aligned} \frac{\partial V_{DM}(\varnothing)}{\partial \eta} &= \underbrace{-\alpha(G,\varnothing,1) f(1) - f(2)}_{\text{Direct effect } (-)} \\ &\quad + \underbrace{\frac{\partial \alpha(G,\varnothing,1)}{\partial \eta} f(1)\left(\mathbb{E}[\theta|1] - \eta + \eta_S\right)}_{\text{Indirect effect } (-)}, \\ \text{where } \frac{\partial \alpha(G,\varnothing,1)}{\partial \eta} &= \frac{-f(2)}{f(1)\left(\mathbb{E}[\theta|1] + \eta_S\right)} > 0. \end{aligned}$$

The direct effect is obvious. Since the sender will communicate a piece of good evidence with probability $\alpha(G,\varnothing,1) f(1) + f(2)$ at period one, it becomes the first order effect on the decrease in η. Indirect effect stems from the fact that the DM must be indifferent between communicating and accepting after communicating once. To keep her indifferent after η gets smaller, the probability that type 1 sender tries to persuade should be suppressed so that the gain from screening that type out at period 2 gets smaller.

An important implication is that the DM wants to make commitment if she can write down a contingent plan to follow, rather than playing the original game. In fact, it is easy to see that the following method of commitment, if possible, makes the DM better off, for sure: the DM accepts the offer with probability slightly smaller than δ/V at period one if a piece of good evidence is shown. At period two, she accepts for sure if the second piece of good evidence is shown. She rejects immediately when the sender shows something else. This commitment makes the DM better off because if the sender does not have two good pieces of evidence, he remains silent from the beginning and hence the commitment makes it possible to avoid accepting type one sender's proposal, that happens with some probability in the equilibrium of the original game.

We can also see that the DM can make her better off even if she can make a limited form of commitment. Think of the commitment in the following method: the DM commits to checking two pieces of evidence as long as the sender tries to communicate, and she accepts the proposal if two good pieces of evidence are shown. If, on the other hand, the sender chooses silence, the DM immediately rejects the proposal. To ensure that the sender's incentive compatibility is satisfied, we assume that $V \geq 2\delta$. Then the expected utility for the DM from this limited method of commitment is

$$V_C = f(2)\,\mathbb{E}[\theta|2] - 2f(2)\,\eta - (f(0) + f(1))\,\eta_S.$$

With probability $f(2)$, the sender is type 2 and the DM has to pay the communication cost of 2η. Otherwise, the sender is a bad type (type 0 or 1) and she will pay just a period cost of silence.

The expected payoff from this limited commitment is higher than W, which is the expected payoff from playing the original game. It is computed as

$$V_C = W + \alpha(G, \varnothing, 1)\, f(1)\,\eta, \tag{1.6}$$

Note that $\alpha(G, \varnothing, 1)$ is the probability that the decision maker can prevent type one sender from persuasion by making the commitment, relative to the equilibrium of the game. Given the equilibrium strategy of the sender, the following strategy is optimal for the DM: continue after the first piece of good evidence and accept after the second piece of good evidence, and otherwise reject. Then, the decision maker expects that if the sender is type 1, she communicates a piece of good evidence with probability $\alpha(G, \varnothing, 1)$ and silent with probability $1 - \alpha(G, \varnothing, 1)$. In the former case, she will end up being silent in the next period. Hence, the expected communication cost with sender type 1 in the equilibrium is $\eta_S + \alpha(G, \varnothing, 1)\,\eta$, while it is just η_S when she makes the commitment. Since the expected communication cost with sender type 2 is the same between the equilibrium and the commitment

(2η), the relation (1.6) follows. Succinctly, the commitment makes it possible to avoid checking type one sender's piece of good evidence, by completely discouraging it from persuasion. Whether the DM can make her better off when the condition $V \geq 2\delta$ does not hold is discussed in Section 1.6.

Although this simple example is enough to provide intuition to some of the important results that are valid for more general cases, there are still some questions that cannot be addressed by the simple example. For instance, obviously, in the setting of $N = \xi$, we cannot have an equilibrium in which communicating a piece of *bad* evidence is on-equilibrium. However, such an equilibrium does exist in more general cases of $N > \xi$. To see this, think of the case in which ξ is large and all types of sender chooses B or S at period one. If the sender chooses G, the DM believes that the sender type is exactly 1 (off-equilibrium belief that we have no restriction) and, hence, immediately rejects the proposal. This in turn makes the sender avoiding G. Hence, one possible question is whether such an equilibrium is efficient or not, relative to other equilibria.

1.4. General Analysis

This section is for characterizing the properties of equilibria. In the first subsection, we give some basic properties of all equilibria. In the second subsection, we examine the properties that must be satisfied in an efficient equilibrium. Those are 1. a bad piece of evidence is never communicated. 2. Silence meets immediate rejection. 3. acceptance probability after a good piece of evidence is δ/V or 1. In the third subsection, we characterize the best equilibrium for the DM, which is unique. We show that the best equilibrium must have the longest possible length of communication among Pareto optimal equilibria.

1.4.1. Properties of All Equilibria

As in most signaling games, our model also has a plethora of equilibria. However, it is possible to identify some important properties that all equilibria have to share. The following theorem characterizes the most important properties of the equilibrium in our persuasion game. It says that every time a piece of good or bad evidence is communicated on-equilibrium, the DM must accept the proposal immediately with strictly positive probability. It also characterizes the lower bound of it.

Theorem 1. *In any equilibrium, if* $m^{t+1} = (m^t, G) \in \Delta$ *then* $\beta(A, m^{t+1}) \in [\delta/V, 1]$. *Also, if* $m^{t+1} = (m^t, B) \in \Delta$ *then* $\beta(A, m^{t+1}) = \delta/V$.

This result follows from the fact that communicating a piece of good or bad evidence incurs cost for the sender. If the probability of acceptance is very small right after (m^t, G), for the sender, communicating an evidence does not pay from myopic point of view, which implies that he expects acceptance with high probability in the future. It implies that every on-equilibrium history afterwards reaches a node that the DM accepts with some probability and, hence, acceptance is the best action for the DM at the node. However, it implies that acceptance is optimal in all the contingencies, which implies that the DM should accept the proposal rather than continuing after (m^t, G).[18]

An important implication of the theorem is that in an equilibrium, the DM must not strictly prefer to continue each time she is communicated a piece of good or bad evidence. In constructing the equilibrium, this restriction imposes conditions about how much of bad sender types drop persuasion in the next period and, hence, how much the DM's informational gain is. Note, however, that it is possible that the DM strictly prefers to continue at period 0, at the point where the sender has not yet paid the communication cost.

The next statement is an immediate corollary to Theorem 1, but also provides an important characterization of the equilibrium in our game. It says that silence has essentially no power of persuading the DM.

Theorem 2. *In an equilibrium, if* $m^{t+1} = (m^t, S) \in \Delta$, *then* $\beta\left(A, m^{t+1}\right) = 0$.

From Theorem 1, every essential communication (not silent) meets immediate acceptance with a strictly positive probability. Moreover, if silence, which incurs no cost for the sender, also meets immediate acceptance with a strictly positive

[18] In the alternative setting in which $\eta_S = 0$, the proposition can be rewritten as follows: if $m^{t+1} = (m^t, G) \in \Delta$ then there is a sequence of silence stage $m^{\tau}_{t+1} = (S, .., S)$ such that

$$\beta\left(A, m^{\tau}_{t+1}\right) = \sum_{s=t}^{\tau-1} \beta\left(A, \left(m^{s+1}, m_s\right)\right) \beta\left(C, \left(m^{s+1}, m_s\right)\right)^{s-t} \geq \delta/V,$$

that is, the DM must accept the proposal with probability higher than δ/V before they communicate another evidence.

probability, acceptance is an optimal for all contingencies from the previous period's point of view. However, then for the DM it is strictly better to accept immediately at the previous period, which is a contradiction.

Theorem 1 implies that the value of a message history for the DM after a piece of evidence is shown (not silence) is equal to the expected payoff from accepting the proposal because it is an optimal action, where the expectation is taken with all the information she had gained through the message history. This is stated in the following corollary. Note that this should be the case even after a piece of bad evidence is communicated as long as it is on an on-equilibrium path.

Corollary 2. *In any equilibrium, it holds that*

$$V_{DM}\left(m^t\right) = \sum_{n=0}^{N} B_n\left(m^t\right) U_{DM}(A, n, m^t) \text{ for all } (m_{t-1}, m_t) \in M^{t-1} \times \{G, B\} \in \Delta.$$

1.4.2. Pareto Optimal Equilibria

In this subsection, we demonstrate that an efficient equilibrium is characterized by three properties.[19] Towards this end, first we define the set of Pareto optimal equilibria. Denote by $\mathcal{E}(\eta, \eta_S, \delta)$ the set of all equilibrium for a given pair of parameter values (η, η_S, δ). Also, we denote each value function with superscript

[19]In this section, we ignore cases of some non-generic constellations of parameter values. More precisely, we exclude the cases in which

$$-\frac{f(j)(\mathbb{E}[\theta|j] + \eta_S)}{\sum_{k=j+1}^{N} f(k)} = \eta \text{ for some } j \leq \xi.$$

e when we are mentioning it in a particular equilibrium e. We define the set of Pareto optimal equilibria as follows:

Definition 2. *Given* (η, η_S, δ), *the set of Pareto optimal equilibria* $\mathcal{P}(\eta, \eta_S, \delta) \subset \mathcal{E}(\eta, \eta_S, \delta)$ *is defined as follows: If* $e \in \mathcal{P}(\eta, \eta_S, \delta)$, *there is no* $e' \in \mathcal{E}(\eta, \eta_S, \delta)$ *such that* $V_{DM}^{e'}(\varnothing) \geq V_{DM}^{e}(\varnothing)$ *and* $\mathbb{E}[V_S^{e'}(\varnothing, j)] \geq \mathbb{E}[V_S^{e}(\varnothing, j)]$, *where the expectation is taken with respect to* j[20], *and one of the inequalities is strict.*

While we defined the set of Pareto optimal equilibria in a way that the sender's expected payoff is compared ex-ante, before the state of the world is realized, we can also define it in the interim way, in which the sender's expected payoff is compared after the state of the world is realized, i.e., the condition "$\mathbb{E}[V_S^{e'}(\varnothing, j)] \geq \mathbb{E}[V_S^{e}(\varnothing, j)]$" is replaced by "$V_S^{e'}(\varnothing, j) \geq V_S^{e}(\varnothing, j)$ for all j". However, all the results provided in this section are valid for whichever criteria we choose.

We define an important class of equilibrium that includes the set of Pareto optimal equilibria as a subset. In any equilibrium in the set, silence meets immediate rejection, even a single piece of bad evidence is never communicated, and acceptance probability is minimized among all possible ways of constructing an equilibrium.

Definition 3. *Given* (η, η_S, δ), *the set of benchmark strategy equilibria* $\mathcal{B}(\eta, \eta_S, \delta) \subset \mathcal{E}(\eta, \eta_S, \delta)$ *is defined as follows: If* $e \in \mathcal{B}(\eta, \eta_S, \delta)$,

[20]Hence, $\mathbb{E}[V_S^{e}(\varnothing, j)] = \sum_{j=0}^{N} f(j)\, V_S^{e}(\varnothing, j)$.

1. A bad piece of evidence is never communicated, that is, $(m^t, B) \notin \Delta$ *for all* $m^t \in M$.

2. For all $(m^t, G) \in \Delta$, *it holds that*

$$\beta\left(A, \left(m^t, G\right)\right) \in \{\delta/V, 1\} \text{ and } \beta\left(C, \left(m^t, G\right)\right) = 1 - \beta\left(A, \left(m^t, G\right)\right). \tag{1.7}$$

3. For all $(m^t, S) \in \Delta$, *it holds that* $\beta(R, (m^t, S)) = 1$.

A property of benchmark strategy equilibrium is that once the DM chooses to enter the communication phase (period 1), all sender types higher than the number τ such that $\beta(A, G^\tau) = 1$ keep communicating good evidence until the proposal is accepted (it is the only optimal behavior given the DM's strategy). The DM never rejects the proposal from a high type sender, because he keeps communicating the good pieces of evidence, until the DM accepts eventually. Hence the DM does not make type I error in this sense.

Figure 1.1 describes how the value of the DM evolves over time in a benchmark strategy equilibrium. At period 1, the sender sends either G or S and the DM's value becomes $V_{DM}(G)$ and $-\eta_S$, respectively. Because, in a benchmark strategy equilibrium, only low type sender sends S at period 1, the DM's optimal action is to reject immediately and it results $V_{DM}(S) = -\eta_S$. Once the game reaches the node G, accepting the proposal is an optimal and she is indifferent between doing so and continuing. This means that $V_{DM}(G)$ is the appropriately weighted average of $V_{DM}(G^2)$ and $V_{DM}(G, S)$. The latter is $-\eta_S - \eta$ because again only low type sender sends S at period 2 and hence rejection is the optimal. At the final period where the DM accepts the proposal for sure, say period τ, her value reaches $\mathbb{E}[\theta|j \geq \tau] - \tau\eta$.

It is useful to define the "length" of persuasion for a benchmark strategy equilibrium. Given a benchmark strategy equilibrium $e \in \mathcal{B}(\eta, \eta_S, \delta)$, we call the number λ such that $\beta\left(A, G^{\lambda}\right) = 1$ but $\beta\left(A, G^{t}\right) = \delta/V$ for all $t < \lambda$, as the length of persuasion and denote it by $N_G(e)$. Actually, the length of persuasion is the number of pieces of good evidence to be required to make the DM accept for sure. Note, of course, that the DM may accept the proposal sooner with some probability and hence the terminology should be understood to be an abbreviation of “maximum possible length of persuasion”. Then, an important property of benchmark strategy equilibrium follows directly from the definition.

Claim 3. *1. In a benchmark strategy equilibrium e, for all $j < N_G(e)$, $V_S(\varnothing, j) = 0$. Moreover, for all $j < N_G(e)$ and $j \leq t < N_G(e)$, $V_S(G^t, j) = (t-1)\delta$.*

2. In a benchmark strategy equilibrium e, for all $j \geq N_G(e)$, $V_S(\varnothing, j) > 0$. Moreover, for all $j \geq N_G(e)$ and $t < N_G(e)$, $V_S(G^t, j) > (t-1)\delta$.

In a benchmark strategy equilibrium, after each message history, the sender has only two choices; communicating a piece of good evidence, or being silent.[21]

[21] In a benchmark strategy equilibrium, the acceptance probability after a piece of bad evidence is communicated (off-equilibrium) is set to be small and, thus, silence, which incurs no cost, is better for the sender.

Because silence effectively implies giving up persuasion, the sender's strategy is characterized by a "dropping vector". Formally, type j sender's strategy is characterized by a ξ dimensional vector

$$d_j = (d_j^1, d_j^2, ..., d_j^{\xi}),$$

where d_j^n, which is $\alpha\left(S, G, G^{n-1}\right)$, represents the probability that type j sender drops persuasion by silence at n's trial, i.e., if $d_j^1 = 1$, type j sender drops at period 1 for sure. Obviously, in a benchmark equilibrium e, for all type $j \geq N_G\left(e\right)$, $d_j^n = 0$ for all $n \leq N_G\left(e\right)$, because it never drops out until eventually persuading the DM (this follows from Claim 3). We denote $N \times \xi$ dimensional vector $(d_1, d_2, ..., d_N)$ (collection of all sender types's strategy) by simply d in the subsequent analysis. The next proposition shows the equations that characterize our benchmark strategy equilibrium.

Proposition 3. *Sender's strategy with dropping vector d such that $d_j^1 > 0$ for some j is supported as a benchmark strategy equilibrium if and only if there is κ such that*

$$\begin{aligned}(1.8)\qquad & -\sum_{j\geq t}^{\xi} d_j^{t+1}\Pi_{s=1}^{t}(1-d_j^s)f\left(j\right)\mathbb{E}[\theta|j] \\ = \;& \eta\sum_{j\geq t+1}^{N}\Pi_{s=1}^{t+1}(1-d_j^s)f\left(j\right)+\eta_S\sum_{j\geq t}^{\xi} d_j^{t+1}\Pi_{s=1}^{t}(1-d_j^s)f\left(j\right)\end{aligned}$$

for all $t < \kappa$ *and* $\Pi_{s=1}^{\kappa+1} d_j^s = 0$ *for all* $j \geq \kappa + 1$, *and*

$$\sum_{j}^{N} \left(1 - d_j^1\right) f(j) \left(\mathbb{E}[\theta|j] - \eta\right) - \eta_S \sum_{j}^{N} d_j^1 f(j) \geq 0. \tag{1.9}$$

In equation (1.8), the left-hand-side is the expected gain from screening out low type sender by continuing at period t. From Theorem 1, after message history G^t, the DM's optimal action is acceptance and thus acceptance is the status quo action. By continuing, with the probability

$$\frac{\sum_{j \geq t}^{N} d_j^{t+1} \Pi_{s=1}^{t} (1 - d_j^s) f(j)}{\sum_{j \geq t}^{N} \Pi_{s=1}^{t} (1 - d_j^s) f(j)}, \tag{1.10}$$

she can know that the sender is a low type and change her action to rejection (a high type sender never give up persuasion). On the other hand, this incurs the cost of communication. With probability

$$\frac{\sum_{j \geq t+1}^{N} \Pi_{s=1}^{t+1} (1 - d_j^s) f(j)}{\sum_{j \geq t}^{N} \Pi_{s=1}^{t} (1 - d_j^s) f(j)},$$

the sender is a high type to show next piece of good evidence with whom the DM has to pay the communication cost of η. On the other hand, with probability (1.10), the sender chooses silence and the DM has to pay the communication cost of η_S. Proposition 3 requires that those two values, when adequately weighted, are equal with each other.

Note that (1.8) implies that if the DM accepts for sure at period κ, we must have

$$-\Pi_{s=1}^{\kappa-1} d_{\kappa-1}^{s} f(\kappa-1)\left(\mathbb{E}[\theta|\kappa-1]+\eta_S\right)=\eta\sum_{j\geq\kappa}^{N} f(j), \tag{1.11}$$

because at period κ, only type $\kappa-1$ sender drops persuasion by being silent and $d_{\kappa-1}^{\kappa}=1$. Therefore, if $-\left(f(l)\,\mathbb{E}[\theta|l]+\eta_S\right)<\eta\sum_{k\geq l+1}^{N} f(k)$ for all $l\geq j$, we have no way to have an equilibrium with the maximum length of persuasion longer than $j+1$. The condition (1.9) ensures that after the first piece of good evidence is communicated, the DM's optimal action is A. If this condition is not satisfied, the DM does not accept the proposal, which contradicts Theorem 1.

We have a corollary of Theorem 1 that is used in the subsequent analysis. It determines the value of the DM's value function at the beginning of the game by a simple formula.

Corollary 3. *In a benchmark strategy equilibrium, it holds*

$$V_{DM}(\varnothing)=\max\{0,\mathbb{E}[\theta],\sum_{j=1}^{N}(1-d_j^1)f(j)\left(\mathbb{E}[\theta|1]-\eta\right)-\sum_{j=1}^{N}d_j^1 f(j)\,\eta_S\} \tag{1.12}$$

In the cases of $V_{DM}(\varnothing)=0$ and $V_{DM}(\varnothing)=\mathbb{E}[\theta]$, the DM just rejects and accepts the proposal without requiring a piece of evidence, respectively. When those are not the case, the DM proceeds to period 1 and, hence, her value is determined by the weighted average of payoffs between the case that the sender communicates a piece of good evidence, where her optimal action is acceptance, and the case that he chooses silent, where her optimal action is rejection.

Here we comment on the general procedure to find out an equilibrium. The easiest way to find an equilibrium is to determine the sender's strategy backward. First, we determine the final period at which the DM accepts the proposal for sure, say period κ. Second, let

$$d_j^\tau = 0 \text{ for all } j \geq \kappa \text{ and } \tau \leq \kappa,$$

that is, the sender type higher than κ certainly keeps communicating good pieces of evidence. It must be so in the equilibrium because for the sender type higher than κ showing a piece of good evidence has a strictly higher continuation value, rather than choosing silence and being rejected. Then, we can determine $\Pi_{t=0}^{\kappa-1}\left(1 - d_{\kappa-1}^t\right)$ by (1.11). The rest of the values of d should be chosen in a way that (1.8) as well as $d_j^1 \geq 0$ for all j is satisfied. If there is no such a way of choosing d, we have no equilibrium with communication. Finally, we see if

$$\sum_{j=1}^{N}(1 - d_j^1) f(j) (\mathbb{E}[\theta|1] - \eta) - \sum_{j=1}^{N} d_j^1 f(j) \eta_S \geq \{0, \mathbb{E}[\theta]\}$$

holds. If it does, we can support $\beta(C, \varnothing) = 1$ and, hence, we have a benchmark strategy equilibrium with the sender's dropping vector d.

Remark 2. *It can be shown that when $\xi \leq 4$, for a generic constellation of parameters, the first and the third conditions of a benchmark strategy equilibrium implies that second.*

Now, we state the main result of this subsection. We denote by $\mathcal{B}(\eta, \eta_S, \delta)$ the set of benchmark strategy equilibrium. Then we have the following theorem, which demonstrates that every efficient equilibrium is a benchmark strategy equilibrium.

Theorem 3. *For all (η, η_S, δ), $\mathcal{P}(\eta, \eta_S, \delta) \subseteq \mathcal{B}(\eta, \eta_S, \delta)$.*

Theorem 1, in Section 1.3, demonstrates that even if there is an equilibrium that involves the communication of a bad piece of evidence, the DM must accept it with probability δ/V, before she communicates the next piece of evidence. This means that even from a piece of bad evidence, the DM is actually positively updating the sender type. The message that the set of Pareto equilibria involves no piece of bad evidence says that the way such an equilibrium screens the sender type is not efficient for both players.

The fact that playing an equilibrium with a period of communicating a piece of bad evidence does not benefit the sender can be easily seen. Because communicating a piece of bad evidence only meets with acceptance probability of δ/V, which is just enough to recover the communication cost, playing another equilibrium that skips such a period does not harm the sender (and it is possible to construct such an equilibrium). On the other hand, the fact that it does not benefit the DM is not as straightforward as one may think. To see this, think of an equilibrium such that a piece of bad evidence should be communicated. Then, the sender type N, who has only good pieces of evidence, has to drop at some period. This equilibrium makes it possible to make the right hand side of (1.8), the cost of communication, smaller at each period. Accordingly, this reduces the dropping from the low type sender at each period (the left-hand-side of (1.8)) or equivalently, increases the dropping at period one, which itself benefits the DM. The question is whether this gain outweighs the loss of giving up the best type sender, $f(N)\,\mathbb{E}[\theta|N]$; the answer turns out to be negative (see the Appendix).

It is rather easy to see that silence should meet immediate rejection in an efficient equilibrium. From Theorem 2, the sender cannot be accepted after silence, which means that having such a period does not make him better off, while even silence is costly for the DM. Those imply that given an equilibrium that has silence that does not meet immediate rejection, it is possible to construct another equilibrium that skips such a period, which Pareto dominates the original equilibrium.

To see the reason that the probability of acceptance immediately after a piece of good evidence should be exactly δ/V or 1 in an efficient equilibrium, suppose that communicating a piece of good evidence, for example for the third time, has acceptance probability strictly higher than δ/V, that is, $\beta(A,G) > \delta/V$. Then all the sender types who have more than three pieces of good evidence will communicate them at least three times. However, in such a case, we can make another equilibrium by making the DM accept the proposal with probability one after communicating three pieces of good evidence. This is an equilibrium, since the sender as well as DM's strategy in the original equilibrium remains an optimum. Now the sender is strictly better off because he can persuade the DM sooner, without harming the DM. Hence, the acceptance probability immediately after a piece of good evidence is either maximized or minimized among all possible ways of constructing an equilibrium.

1.4.3. The Best Equilibrium

In this subsection, we characterize the best equilibrium for the DM (we say just the best equilibrium, hereafter), that is, the equilibrium such that $V_{DM}(\varnothing)$, the value of the DM at the initial period, is maximized. Because it is proved that the best equilibrium is unique for all parameter values, it pins down the equilibrium on which we can do comparative statics (Section 1.5). Also, it gives the highest benchmark with which the DM's equilibrium payoff is compared when we examine the commitment problem (Section 1.6).

The result given in the previous subsection already demonstrated that the best equilibrium, which must be a Pareto optimal equilibrium, is one of benchmark strategy equilibrium. Therefore, in this section, we focus our analysis exclusively on the set of benchmark strategy equilibrium. The first property of the best equilibrium for the DM is that it is actually unique.

Proposition 4. *$\mathcal{B}(\eta, \eta_S, \delta)$ has a unique maximizer of $V_{DM}(\varnothing)$.*[22]

An important characteristic of the best equilibrium for the DM is that it maximizes the length of persuasion among the set of Pareto efficient equilibria. Intuitively, increasing the amount of good evidence necessary for persuasion discourages bad senders from trying to persuade.

Theorem 4. *If equilibrium e^* is the best equilibrium for the DM, there is no equilibrium e such that $N_G(e) > N_G(e^*)$.*

Note that there are multiple equilibria even if we focus on the ones that maximize the length of persuasion. Also, note that the theorem does not state that an equilibrium has a higher expected payoff than another equilibrium if the former has longer length of persuasion.[23] It only says that if an equilibrium is the best equilibrium, it must have the maximum length of persuasion.

To see this result in the simplest case, suppose that there are two equilibria, one with the length of persuasion of 1 and the other with the length of persuasion of 2. In the former equilibrium, both type 1 and 2 senders try to persuade, which implies that after checking a single piece of good evidence, the value of the proposal is $\mathbb{E}[\theta | j \geq 1]$ for the DM. On the other hand, in the latter equilibrium, type 1 sender does not try to persuade with probability one, which implies that after checking

[22]We regard two equilibria that are outcome equivalent identical.

[23]This statement holds in the special case of $N = \xi = 2$, where the best equilibrium is unique.

one piece of evidence the value of the proposal is higher than $\mathbb{E}[\theta|j \geq 1]$.[24] Hence, the value of the decision maker at period zero is higher in the latter because it screens out more bad sender (type one sender) by the first piece of evidence.

The procedure of finding the best equilibrium involves 1. find the maximum length of persuasion. 2. given the maximum length of persuasion, find the way that the sender gives up persuasion over time in a way that the low type's trial by showing an evidence at period 1 is suppressed the most. In other words, given the length of equilibrium, we have to connect different periods by equation (1.8) in a way that $\left|\sum_{j}^{\xi-1} d_j^1 f(j)\, \mathbb{E}[\theta|j]\right|$ is maximized.

We can get the basic idea of characterizing the best equilibrium by rewriting the condition (1.8) as

$$-\sum_{j \geq t}^{\xi} d_j^{t+1} \Pi_{s=1}^{t} (1 - d_j^s) f(j) \left(\mathbb{E}[\theta|j] + \eta_S\right) = \eta \sum_{j \geq t+1}^{N} \Pi_{s=1}^{t+1} (1 - d_j^s) f(j) \,. \tag{1.13}$$

As we saw in the previous subsection, we can construct an equilibrium backward. We first determine the last period of persuasion, say κ, and let $d_j^t = 0$ for all $t \leq \kappa$ and $j \geq \kappa$, i.e, the sender type higher than κ never drop persuasion. Then, we choose elements of d backward so that equation (1.13) is satisfied for all period. At each period t, given the value of the right hand side, there are

[24] Under the condition that the latter equilibrium exists, in order to support the former the decision maker has to expect that type 2 sender does not show the second piece of good evidence with probability one at period 2.

multiple ways to assign the probability of dropping, $d_j^{t+1}\Pi_{s=1}^t\left(1-d_j^s\right)$, among different types to make the equality hold. An important observation is that because the absolute value of $\mathbb{E}[\theta|j]$ is decreasing with j as long as $j \leq \xi$, if we decrease $d_j^{t+1}\Pi_{s=1}^t\left(1-d_j^s\right)$ a bit for some j, say by $\Delta d_j^{t+1}\Pi_{s=1}^t\left(1-d_j^s\right)$, we need to increase it for i for more than $\Delta d_j^{t+1}\Pi_{s=1}^t\left(1-d_j^s\right)$ if $i > j$. This re-allocation of the dropping probability leads to higher value of the right hand side at period $t-1$, which leads lower dropping at period 1, i.e., lower $\left|\sum_j^{\xi-1} d_j^1 f(j)\,\mathbb{E}[\theta|j]\right|$. This observation implies that we should use more lower types to tie the consecutive period in the equality (1.13).

Although it is possible to state the general algorithm to construct the best equilibrium that is applicable to general cases, we can introduce an assumption that makes the characterization of the best equilibrium easier. Towards this end, let ζ be the highest $j < \xi$ such that $\mathbb{E}[\theta|j] + \eta_S < 0$. Approximately, ζ is the sender type that the DM does not dare to pay the communication cost to screen it out. Obviously, for any equilibrium e, $N_G(e) < \zeta$.

Think of the function $\Gamma : \{0, 1, .., \zeta\} \to \mathbb{R}$ that is defined as

$$\Gamma(j) = \left|\frac{f(j)}{\sum_{k=j+1}^N f(k)}(\mathbb{E}[\theta|j] + \eta_S)\right|.$$

The assumption we want to impose is function Γ being decreasing. The function $\Gamma(j)$ is made by multiplying the loss from accepting type j sender's proposal with the probability that the sender's type being j relative to the probability that the sender type is strictly higher than j. Approximately, high $\Gamma(j)$ implies that the DM has strong incentive to screen out type j sender, after having already screened out lower types. A sufficient condition for $\Gamma(j)$ to be decreasing is that $|\mathbb{E}[\theta|j]|$ decreases fast enough to compensate for the change in the term $f(j) / \sum_{k=j+1}^{N} f(k)$, which is likely to be increasing. If $f(j) / \sum_{k=j+1}^{N} f(k)$ is increasing with j, the assumption is automatically satisfied.

To see that the assumption makes it easier to find the maximum length of persuasion, see condition (1.11). Under this condition, the maximum length of persuasion is determined by the largest κ such that $\eta < \Gamma(\kappa-1)$. From the condition, we know that $\eta < \Gamma(l)$ for all $l < \kappa$ and this implies that we can find dropping vectors which can take only values less than one, in such a way that (1.8) is satisfied at each period. If the condition is not satisfied, the fact that j is the largest number satisfying $\eta < \Gamma(\kappa-1)$ does not necessarily imply that the maximum length of persuasion is κ. To see this, think of the case that the assumption of Γ decreasing is not satisfied and $\eta < \Gamma(\kappa-1)$ but $\eta > \Gamma(\kappa-2)$. To be an equilibrium with maximum length of persuasion of κ, we must have (1.11) holding in order to support period $\kappa-1$'s behavior of the DM (mixing between accepting and continuing) and we also have

$$-\sum_{j\geq\kappa-2}^{\kappa-1} d_j^{\kappa-1}\Pi_{s=1}^{\kappa-2}\left(1-d_j^s\right)f(j)\left(\mathbb{E}[\theta|j]+\eta_S\right)=\eta\{\Pi_{s=1}^{\kappa-1}\left(1-d_j^s\right)f(\kappa-1)+\sum_{j\geq\kappa}^{N}f(j)\},$$

in order to support the DM's period $\kappa - 2$'s behavior. Here, note that only type $\kappa - 2$ or $\kappa - 1$ sender can drop at period $\kappa - 1$. However, if $\eta > \Gamma(\kappa - 2)$, it may not possible to choose d in such a way that above equality is satisfied, which implies that the maximum length of persuasion should be shorter than $\kappa - 2$.

In sum, under the condition of $\Gamma(j)$ being decreasing, the maximum length of persuasion is determined by j that satisfies

$$\Gamma(j) > \eta > \Gamma(j+1). \tag{1.14}$$

The assumption also makes it easier to find out the optimal dropping vector. As we discussed above, in finding the best equilibrium, we should use more lower type sender's dropping to tie the consecutive period by the equality (1.13). In the equilibrium characterized in the theorem, we use only type j sender's dropping to make period j's equation (1.13). Apparently, type j is the lowest possible type to drop at period t in a benchmark strategy equilibrium. If Γ is decreasing, it is ensured that once we can make the equation (1.13) at period j satisfied by letting only type j sender drops at period j, it is also possible to make the equations hold at previous periods in the same way.

Think of the following procedure to find out a N dimensional vector $c = (c_0, c_1, c_2, .., c_N)$.

· Step 1. Let $c_\xi = c_{\xi+1} = .. = c_N = 1$, and $c_0 = 0$. Find $c_{\xi-1}$ that satisfies

$$\eta = -\frac{c_{\xi-1} f\left(\xi-1\right)}{\sum_{j=\xi}^{N} c_j f\left(j\right)} (\mathbb{E}[\theta|\xi-1] + \eta_S),$$

if we cannot find $c_{\xi-1}$ in a way that $c_{\xi-1} \leq 1$, let $c_{\xi-1} = 1$. Next, find $c_{\xi-2}$ that satisfies

$$\eta = -\frac{c_{\xi-2} f\left(\xi-2\right)}{\sum_{j=\xi-1}^{N} c_j f\left(j\right)} (\mathbb{E}[\theta|\xi-2] + \eta_S).$$

If we cannot have $c_{\xi-2}$ in a way that $c_{\xi-2} \leq 1$, let $c_{\xi-2} = 1$ and rewrite $c_{\xi-1} = 1$, and continue this process until we get c_1. If we get $c_j > 1$ at some period, rewrite $c_j = c_{j+1} = .. = c_N$ and continue. Let the greatest k such that $c_k < 1$ be γ.

· Step 2. Check if $V = \sum_{j=1}^{N} c_j f\left(j\right) (\mathbb{E}[\theta|j] - \eta) - \sum_{j=1}^{N} (1-c_j) f\left(j\right) \eta_S \geq \mathbb{E}[\theta]$ and also $V \geq 0$. If both hold, it is done. If it is not, then let $c_1 = c_2 = .. = c_N = 0$.

The above argument is summarized in the following theorem, which demonstrates that the best equilibrium is found by the procedure when $\Gamma\left(j\right)$ is decreasing.

Theorem 5. *Let $(c_0, c_1, .., c_N)$ be a vector derived from the above procedure. If function $\Gamma\left(j\right)$ is decreasing, there is a unique (in the class of outcome equivalent) DM's utility maximizing equilibrium that is characterized as follows:*

1. $\alpha\left(G, j, \varnothing\right) = c_j,\ \alpha\left(S, j, \varnothing\right) = 1 - c_j$.

2. $\alpha\left(G, j, G^t\right) = 1$ if $t \leq j$ and $t \geq 1$.

3. If $m^t \neq G^t,\ \beta\left(R, m^t\right) = 1$.

4. If $t \leq \gamma,\ \beta(A, G^t) = \delta/V$. *If* $t \geq \gamma + 1,\ \beta(A, G^t) = 1$.

In this equilibrium, each type of sender mixes at period one whether to communicate a good piece of evidence or to give up persuasion by being silent. Once he chooses to communicate a piece of good evidence, he does so until he runs out of evidence. At each period, say t, the DM can screen out exactly type t sender by continuing. The way that type t sender mixes at period one makes the DM's expected benefit from screening out type t and cost of communication equal with each other at period t.[25]

1.5. Comparative Statics

In this section, we examine the effect of changes in the model's cost parameters (η, η_S, δ) on the equilibrium. In order to do this, hereafter, we focus solely on the best equilibrium for the decision maker and, hence, from the analysis of Section 1.6, we focus on the best benchmark strategy equilibrium, where value functions are denoted with superscript "*". We have the next theorem, whose proof is easy and thus omitted:

Theorem 6. *1. Fix* (η_S, δ). *Suppose that the prior of the proposal is bad, i.e.,* $\mathbb{E}[\theta] < 0$. *Then* $\mathbb{E}_j[V_S^*(\varnothing, j)]$ *is a step function of* η *and there is a threshold value of* η *under which it is increasing, and above which it is zero.*

[25] The equilibrium characterized in the theorem has a dropping vector such that $d_j^1 = 1 - c_j$, $d_j^k = 0$ for all $k \in \{2, .., j\}$, and $d_j^{j+1} = 1$.

2. Fix (η_S, δ). Suppose that the prior of the proposal is good, i.e., $\mathbb{E}[\theta] > 0$. *Then* $\mathbb{E}_j[V_S^*(\varnothing, j)]$ *is a step function of* η *and there is a threshold value of* η *under which it is increasing, and above which it is* V.

If the prior of the proposal is bad, the DM whose communication cost η is very high, does not talk with the sender and just rejects the proposal. For the sender, this is the worst case because he has no chance of persuading her. The best DM for the sender is the DM whose communication cost is low enough to communicate, but not too low to be willing to communicate for a long time. The DM may communicate as long as $\Gamma(j)$ defined in the previous section exceeds η and thus the maximum length of persuasion is decreasing with η. In this case of the prior of the proposal being bad, the expected payoff of the sender is non-monotonic. Note that the relation between $\mathbb{E}[V_S(\varnothing, j)]$ and η is a step function whose values depends on the length of persuasion.

Figure 1.2 describes the relation when $\Gamma(j)$ is decreasing, in which the length of persuasion is determined by (1.14). Because $\Gamma(j)$ is decreasing, as we gradually increase η from zero, the length of persuasion decreases one by one and, thereby, increases the expected payoff of the sender.

On the other hand, if the proposal is ex-ante good and the DM has a very high communication cost, the DM does not require evidence from the sender and just rubber-stamps the proposal. For the sender, this is the best possible case in which his expected payoff is maximized. Therefore, in such a case of ex-ante good proposal, the expected payoff of the sender becomes monotonic, which is again a step function.

The effect of change in η on the DM's expected payoff is divided into direct and indirect effects, as we have seen in the example of Section 1.3. We have

$$\begin{aligned} V_{DM}^{*}(\varnothing) &= \sum_{j=1}^{N_G(e)-1} \alpha(G,\varnothing,j) f(j)(\mathbb{E}[\theta|1]-\eta) + \sum_{j=N_G(e)}^{N} f(j)(\mathbb{E}[\theta|1]-\eta) \\ & -\eta_S \sum_{j=1}^{N_G(e)-1} \{1-\alpha(G,\varnothing,j)\} f(j), \end{aligned}$$

since the sender will communicate a piece of good evidence or being silent at period 1. Hence we have $\frac{\partial V_{DM}^{*}(\varnothing)}{\partial \eta} =$

$$\underbrace{-\sum_{j=1}^{N_G(e^*)-1} \alpha(G,\varnothing,j) f(j) - \sum_{j=N_G(e^*)}^{N} f(j)}_{\text{Direct effect }(-)}$$

$$+\underbrace{\sum_{j=1}^{N_G(e^*)-1} \frac{\partial \alpha(G,\varnothing,j)}{\partial \eta} f(j)(\mathbb{E}[\theta|1]-\eta+\eta_S)}_{\text{Indirect effect }(-)}.$$

We can see that the direct effect is negative, and we are able to show that the indirect effect is also negative (see Appendix). Interpretation of the direct effect is straightforward: it just reduces the cost of communication at period one. The indirect effect comes from the later periods. The reduction in η makes it possible to make the DM indifferent between acceptance and continue at each period with a small benefit from screening and, hence, with a high dropping of bad senders at period 1.

Another important implication is that, assuming that the best equilibrium always holds, the probability that the decision maker makes a wrong decision, either choosing A when the sender type is less than ξ (type II error) or choosing R when the sender type is bigger than ξ (type I error), monotonically converges to zero as η converges to zero. Thus, by denoting the probability by $\digamma(\eta, \eta_S)$, the next theorem follows, where its proof is omitted:

Theorem 7. *Fix δ. Then,*

$$\lim_{\eta \downarrow 0} \sup_{\eta_S \leq \eta} \digamma(\eta, \eta_S) \to 0.$$

To see the theorem, first note that the probability that the decision maker makes the wrong decision of accepting the bad proposal, type II error, is smaller than $\sum_{j=1}^{N_G(e^*)-1} \alpha(G, \varnothing, j) f(j)$. It is easily seen by (1.8) that the absolute value of it is decreasing. On the other hand, in the best equilibrium, type I error never happens when $\mathbb{E}[\theta] > 0$. Even when $\mathbb{E}[\theta] \leq 0$, type I error never happens as long as the DM chooses C at period 0, which is the case when η is sufficiently small.

We next consider comparative statics with respect to the sender's communication costs. It is easy to see from the construction of equilibrium that the DM's expected payoff is invariant with the sender' cost of communication. On the other hand, the sender's expected payoff can be naturally shown to be decreasing with his cost of communication.

Theorem 8. *Fix* (η, η_S). *Then* $V_{DM}^*(\varnothing)$ *is constant with respect to* δ *and* $\mathbb{E}[V_S^*(\varnothing, j)]$ *is strictly decreasing with* δ.

The reason for $\mathbb{E}[V_S^*(\varnothing, j)]$ being strictly decreasing with δ is easy to see. From Claim 3, the low type sender's expected payoff is 0, irrespective of his communication cost δ, which comes from the fact that acceptance probability will adjust in an equilibrium. However, the acceptance probability at period $N_G(e^*)$, which is 1, cannot adjust with the change in δ, which implies that an increase in δ decreases the high type sender's expected payoff. These also tell that a decrease in δ lengthens the expected time before acceptance.

1.6. Commitment

In this section, we examine whether the DM can be better off by making a commitment if she can write down a contingent plan to follow. If we think of the DM as an organization which is frequently making decisions based on the advice of concerned parties, this question is particularly important for designing the rule used to handle this advice. We consider two different forms of commitment.

Optimal Stochastic Commitment

The answer to the problem, however, is easy if the DM is allowed to make commitment in a very sophisticated way. In fact, it is easy to see that the following method of commitment, if possible, makes the DM better off for sure: the DM accepts the offer with probability δ/V (or actually, slightly lower than) each time a piece of good evidence is shown, until enough good evidence is shown at which point she accepts the offer for sure. It follows that if the sender does not have enough good pieces of evidence to show, he remains silent from the beginning. Once the DM knows that, she has an incentive to accept as soon as possible. The probability δ/V is the largest probability of acceptance that can make the screening of sender types possible. Furthermore, it can be shown that this is the optimal commitment that the DM can make.

Theorem 9. *The optimal commitment takes the following form: the DM accepts the proposal with probability δ/V each time the sender communicates a piece of good evidence until κ pieces of good evidence are communicated, where κ is the number characterized by*

$$\begin{aligned}\kappa &= \arg\max_{k} \sum_{j\geq k} f(j)\,\mathbb{E}[\theta|j] \\ &\quad - \sum_{j\geq k} f(j)\,[\sum_{n\geq 1}^{k} n\eta\left(1-\frac{\delta}{V}\right)^{n-1}\frac{\delta}{V} + k\eta\left(1-\frac{\delta}{V}\right)^{\kappa-1}] - \eta_S \sum_{j<k} f(j)\,.\end{aligned}$$

An important point is that the stochastic commitment is a Pareto improvement from the best equilibrium. This follows because in the best equilibrium, the low type sender's expected payoff is zero, and it can be shown that the length of communication is shorter in the commitment case than in the best equilibrium, which makes the sender better-off. This tells us that the persuasion game involves an inevitable waste of time or energy, which can be mitigated by the commitment.

Optimal Limited (Non-Stochastic) Commitment

Although stochastic commitment attains a desirable outcome compared to the best equilibrium, making stochastic commitment possibly difficult because the agent cannot verify the DM's behavior and the DM cannot prove that she is actually following the committed plan. In order to consider such a case, in particular, we think of the following form of commitment that is easier to make: she decides to listen to the sender for a predetermined length of time, say τ, as long as good pieces of evidence are shown. If she is shown τ pieces of good evidence in a row, she accepts the offer, while she rejects the offer as soon as she is shown other evidence or silence. If the DM makes such a commitment, it is optimal for the sender types lower than τ to remain silent at period 1 and get rejected, because they know that they cannot persuade the DM. We call this type of commitment "limited commitment," hereafter.

The optimization problem the DM has to solve when she makes the limited commitment is as follows:

$$\begin{aligned}\max_k \Upsilon_{DM}(k) &= \max_k \{\sum_{j\geq\kappa} f(j)(\mathbb{E}[\theta|j]-\kappa\eta)-\eta_S\sum_{j<\kappa} f(j)\}, \\ \text{subject to } \kappa\delta &\geq V.\end{aligned}$$

With probability $\sum_{j\geq\kappa} f(j)$, the sender is of high type and the DM has to pay the communication cost of $\kappa\eta$. Otherwise, the sender is low type and the DM pays just a period cost of silence. We have a participation constraint for the sender, $\kappa\xi \geq V$. Unless this condition is satisfied, the sender does not try to persuade the DM by paying the communication cost.

We have the following result, which shows that the DM is better off by making a limited commitment if the sender's persuasion gain V is high enough relative to his communication cost.

Theorem 10. *Suppose that $V \geq \delta N_G(e^*)$,[26] where $N_G(e^*)$ is the length of persuasion of the best equilibrium. If $\beta(C,\varnothing)=1$ in the best equilibrium, the DM prefers to make limited commitment, i.e.,* $\max_k \Upsilon_{DM}(k) > V^*_{DM}(\varnothing)$.

This result follows from the same reason as we discussed in Section 1.4. In the best equilibrium, an optimal strategy of the decision maker, given the sender's

[26]Note that $N_G(e^*)$ is an endogenous variable. Another sufficient condition that uses only exogenous variable is $V \geq \delta\xi$, which is stronger because $\xi \geq N_G(e^*)$.

strategy, is to require further pieces of evidence until enough good evidence is communicated, and otherwise reject. This prevents bad senders from communicating good evidence and then giving up, which causes the DM to incur communication costs.

Contrary to the stochastic commitment, the limited commitment is not a Pareto improvement from the best equilibrium. This follows because in the best equilibrium, the high type sender has a chance of succeeding with persuasion quickly, while he has to communicate for a certain amount of time in a limited commitment case. Because the low type sender's expected payoff is zero, both in the best equilibrium and with limited commitment, it can happen that the sender is worse-off in a limited commitment case than in the best equilibrium.[27]

The above result, however, can only be guaranteed if $V \geq \delta N_G(e^*)$, i.e., the sender is willing to pay the persuasion cost in order to induce his preferred action from the DM even if it takes $N_G(e^*)$ periods to communicate with certainty. Once this condition is violated, it is possible to have a situation where the DM prefers to play the persuasion game instead of making a limited commitment. An example is shown in the following claim.

Claim 4. *Suppose that the parameter values of the model are as follows:* $N = \xi = 2$. $2\delta > V$, $\mathbb{E}[\theta | j \geq 1] > 0$, *and the DM's communication cost* η *is small to*

[27] It can be shown that the length of persuasion is shorter in the commitment case than in the best equilibrium.

the extent that (1.5) is satisfied. Then, the DM prefers to play the best equilibrium than the limited commitment, i.e., $V^*_{DM}(\varnothing) > \max_k \Upsilon_{DM}(k)$.

The proof is easy. In the setting above, the best limited commitment is to require only one good piece of evidence. If she requires two pieces, no sender type tries to persuade her. Hence, she can only require at most one piece of good evidence in the limited commitment, which gives her the same expected payoff, as playing the game with the equilibrium of $N_G(e) = 1$. Then, the claim follows from Theorem 4, which states that the equilibrium that attains the highest expected payoff for the DM has the longest length of persuasion.

More generally, in the best equilibrium, we may have $\delta N_G(e^*) > V$, which means that the sender communicates for too long and pays more persuasion cost than what he can get (V) if the decision maker postpones the decision the most. This makes it possible for the DM to extract more information from the sender, relative to the case of limited commitment where the sender is perfectly knowledgeable about the outcome of the persuasion and, hence, never pay the communication cost excessively.

1.7. Conclusion

In this study, we developed a model that describes the dynamic process of persuasion. We show that the equilibrium necessarily involves probabilistic behavior from both parties. We characterized the set of Pareto efficient equilibria and the best equilibrium for the decision maker.

Although we provided an entrepreneur-venture capitalist relation as a primary example, there are a lot of real world examples that fit out model. Glazer and Rubinstein (2004) provide a number of nice examples of persuasion through hard evidence.[28] Those include, for example, the case in which a worker wishes to be hired by an employer for a certain position. The worker tells the employer about his previous experience and the employer wishes to hire the worker if his ability is above a certain level.

It may be interesting to extend the model in a way that parameters η and δ, which represent players' costs of communication, have non-degenerate distributions and also are private information. Then, we will obtain more complicated strategic interactions because the fact that the game did not terminate until a particular period conveys some information about the players' types. This gives our game an additional flavor of Fudenberg and Tirole's (1986) war of attrition model.

Obviously, this is just a first step for a deeper understanding of the process of persuasion. There are a lot of questions that cannot be addressed in this study. These include interesting questions such as 1. In which order should pieces of evidence be released when each piece of evidence has a different value? 2. If the sender is allowed to show multiple pieces of evidence at a time, how does this

[28] They work on a setting that the DM is restricted to checking only one piece of evidence. In this sense, they think of the case where players face a very tight constraint in communication relative to our model.

change the nature of persuasion? 3. Can we render a reasonable explanation for why sometimes a persuader reveals unfavorable information? Those questions are left up to future research.

1.8. Appendix: Proofs

In the following, we use the following notations.

T : The set of terminal message histories that can be reached with strictly positive probabilities, i.e., $T = \{m^t \in \Delta \ |\beta (C, m^t) = 0\}$.

$P(m^t)$: The set of types of sender that follows message history m^t with strictly positive probabilities, i.e., if $j \in P(m^t)$ then $\Pi_{s=1}^{t} \alpha (m_s, m^{s-1}, j) > 0$.

$\succ$: Incomplete order on M defined as $m^s \succ m^t$ if and only if $m^s = (m^t, m_{t+1}^s)$ for some m_{t+1}^s such that $s \geq t+1$, i.e., m^s is a continuation from m^t.

$\alpha(m^\tau, j)$: The probability that type j sender follows communication history m^τ, i.e., $\alpha(m^\tau, j) = \Pi_{s=1}^{\tau} \alpha (m_s, m^{s-1}, j)$.

$\alpha\left(m_{t-1}^\tau, j\right)$: The probability that type j sender follows message history m_{t-1}^τ from period $t-1$, i.e., $\alpha\left(m_{t-1}^\tau, j\right) = \Pi_{s=t}^{\tau} \alpha (m_s, m^{s-1}, j)$.

Proofs of Lemma 1 and 2: We first prove the uniqueness of V_{DM}. Let (α, φ, B) be given. Suppose that we have two value functions V_{DM} and V'_{DM} that satisfy conditions (1.1) and (1.2). Let

$$W = \{m^t | V_{DM}\left(m^t\right) \neq V'_{DM}\left(m^t\right)\},$$

which is the set of message histories such that two value functions take different values. To get a contradiction, suppose that $W \neq \varnothing$. Then there must be some $m^\tau \in W$ such that

$$\begin{aligned} V_{DM}\left(m^\tau\right) &> \max_{a\in\{A,R\}} \mathbb{E}[U_{DM}(a,j,m^\tau)|m^\tau] \\ \text{and } V'_{DM}\left(m^\tau\right) &= \max_{a\in\{A,R\}} \mathbb{E}[U_{DM}(a,j,m^\tau)|m^\tau], \end{aligned} \tag{1.15}$$

or

$$\begin{aligned} V_{DM}\left(m^\tau\right) &= \max_{a\in\{A,R\}} \mathbb{E}[U_{DM}(a,j,m^\tau)|m^\tau] \\ \text{and } V'_{DM}\left(m^\tau\right) &> \max_{a\in\{A,R\}} \mathbb{E}[U_{DM}(a,j,m^\tau)|m^\tau], \end{aligned} \tag{1.16}$$

where we denote $\sum_{n=0}^{N} B_n\left(m^t\right) U_{DM}(a,j,m^t)$ by $\mathbb{E}[U_{DM}(a,j,m^t)|m^t]$.

To see this, note that if neither holds,

$$V_{DM}\left(m^\tau\right) > \max_{a\in\{A,R\}} \mathbb{E}[U_{DM}(a,j,m^\tau)|m^\tau]$$

and

$$V'_{DM}\left(m^\tau\right) > \max_{a\in\{A,R\}} \mathbb{E}[U_{DM}(a,j,m^\tau)|m^\tau]\ for all m^\tau \in W.$$

These, respectively, imply that

$$\begin{aligned} V_{DM}\left(m^{\tau}\right) &= \mathbb{E}[V_{DM}\left(m^{\tau}, m_{\tau+1}\right)|m^{\tau})] = \sum_{m_{t+1}} \varphi\left(m_{\tau+1}|m^{\tau}\right) V_{DM}\left(m_{\tau}, m^{\tau}\right) \\ \text{and } V'_{DM}\left(m^{\tau}\right) &= \mathbb{E}[V'_{DM}\left(m^{\tau}, m_{\tau+1}\right)|m^{\tau})] = \sum_{m_{t+1}} \varphi\left(m_{\tau+1}|m^{\tau}\right) V'_{DM}\left(m_{\tau}, m^{\tau}\right). \end{aligned}$$

Since $m^{\tau} \in W$, we let $V_{DM}\left(m^{\tau}\right) > V'_{DM}\left(m^{\tau}\right)$, without loss of generality. Then above relations imply that there is some $m_{\tau+1}$ such that

$$V_{DM}\left(m^{\tau}, m_{\tau+1}\right) > V'_{DM}\left(m^{\tau}, m_{\tau+1}\right) > \max_{a \in \{A,R\}} \mathbb{E}[U_{DM}(a, j, \left(m^{\tau}, m_{\tau+1}\right)|m^{\tau}, m_{\tau+1}].$$

By continuing the same argument, we have a sequence $\{m_s\}_{s=t}^{\infty}$ such that

$$V_{DM}\left(m^{s}\right) \leq V_{DM}\left(m^{s}, m_{s+1}\right) \text{ for all } s \geq \tau.$$

From (1.2), we have to have $\lim_{s\to\infty} V'_{DM}\left(m^{s}, m_{s+1}\right) = \lim_{s\to\infty} -\eta\{N_G\left(m^{s}\right) + N_B\left(m^{s}\right)\} = -\infty$, but this contradicts $V_{DM}\left(m^{\tau}\right) > \max_{a \in \{A,R\}} \mathbb{E}[U_{DM}(a, j, m^{\tau})|m^{\tau}]$.

Without loss of generality, let (1.15) holds. For every history m^t that can be reached from m^{τ} with a strictly positive probability, we can find the smallest $s \leq t$ such that $V_{DM}\left(m^{s}\right) = \max_{a \in \{A,R\}} \mathbb{E}[U_{DM}(a, j, m^{s})|m^{s}]$. To see this, note that if it is not, we have a sequence $\{m_s\}_{s=\tau}^{\infty}$ such that $V_{DM}\left(m^{s}\right) \leq V_{DM}\left(m^{s}, m_{s+1}\right)$ for all $s \geq \tau$ and thus $V_{DM}\left(m^{s}\right) \leq \lim_{s\to\infty} V_{DM}\left(m^{s}\right) = -\infty$, which contradicts $V_{DM}\left(m^{\tau}\right) > \max_{a \in \{A,R\}} \mathbb{E}[U_{DM}(a, j, m^{\tau})|m^{\tau}]$. Let the set of such histories Λ, and probability measure on Λ generated by φ be ω. Then we have $V_{DM}\left(m^{\tau}\right) =$

$\int V_{DM}(s)\,d\omega(s)$. For each $s \in \Lambda$, we must have

$$V'_{DM}(m^s) \geq \max_{a \in \{A,R\}} \mathbb{E}[U_{DM}(a, j, m^s)|m^s] = V_{DM}(m^s),$$

and, hence, we have $\int V'_{DM}(s)\,d\omega(s) \geq \int V_{DM}(s)\,d\omega(s)$. However, it must hold that $V'_{DM}(m^\tau) \geq \int V'_{DM}(s)\,d\omega(s)$ from the definition of the value function. This implies $V'_{DM}(m^\tau) \geq V_{DM}(m^\tau)$, which is a contradiction. Hence the uniqueness of the value function follows.

Next, we prove the uniqueness of V_S. To get a contradiction, suppose that we have two different value functions, and let $V_S(m^t, j) > V'_S(m^t, j)$ for some j and m^t. Make the sequence $\{m^s\}_{s=t}^{\infty}$ by

$$\begin{aligned} V_S\left(m^t, x\right) &= \beta\left(A, m^t\right)\left(V - C_S\left(m^t\right)\right) \\ &\quad -\beta\left(R, m^t\right) C_S\left(m^t\right) + \beta\left(C, m^t\right) V_S\left(\left(m^{t+1}, a\right), j\right). \end{aligned}$$

Then we have

$$\begin{aligned} &\lim \Pi_{t=0}^{\infty} \beta\left(C, m^t\right) V_S\left(m^{t+1}\right) - \lim \Pi_{t=0}^{\infty} \beta\left(C, m^t\right) V'_S\left(m^{t+1}, j\right) \\ > \; & V_S\left(m^t, j\right) - V'_S\left(m^t, j\right) > 0, \end{aligned}$$

which contradicts (1.2).

Next, we will prove the existence of V_{DM}. To shorten the notation, denote by $g(m^t)$ the highest value of expected utility of the DM when she decides whether to accept or reject, i.e., $g(m^t) = \max_{a\in\{A,R\}} \mathbb{E}[U(a,j,m^t)|m^t]$. Take an arbitrary m^t and fix it. Make the sequence of real numbers $V_0(m^t), V_1(m^t), ...$ as follows. Let

$$V_0\left(m^t\right) = g\left(m^t\right),\ V_1\left(m^t\right) = \max\{g\left(m^t\right), \sum_{m^{t+1}} \varphi\left(m^{t+1}|m^t\right) g\left(m^{t+1}\right),$$

and $V_2(m^t) =$

$$\max\{g\left(m^t\right), \sum_{m_{t+1}} \max\{\varphi\left(m^{t+1}|m^t\right) g\left(m^{t+1}\right), \sum_{m_{t+2}} \{\varphi\left(m^{t+2}|m^{t+1}\right) g\left(m^{t+2}\right)\},$$

and so on. That is, $V_k(m^t)$ is constructed by $V_{k-1}(m^t)$ by replacing terms

$$\varphi\left(m^{t+k-1}|m^{t+k-2}\right) g\left(m^{t+k-1}\right)$$

with

$$\max\{\varphi\left(m^{t+k-1}|m^{t+k-2}\right) g\left(m^{t+k-1}\right), \sum_{m_{t+k-1}} \{\varphi\left(m^{k+t}|m^{k+t-1}\right) g\left(m^{k+t}\right)\}.$$

Obviously, $V_n(m^t)$ is an increasing sequence with each satisfies $V_n(m^t) \leq 1 - \eta\{N_G(m^t) + N_B(m^t)\}$. Hence it converges to some value $V_\infty(m^t) \leq 1 - \eta\{N_G(m^t) + N_B(m^t)\}$. Let this value be $V_{DM}(m^t)$, and do this for all elements in H. It is a routine work to verify that those satisfy the condition for being the value function.

To prove the existence of $V_S(m^t, j)$, pick a pair (m^t, j) and fix it. Again, define the sequence $V_0(m^t, j)$, $V_1(m^t, j)$, .. as follows:

$$V_0\left(m^t, j\right) = \beta\left(A, m^t\right)\left\{V - C_S\left(m^t\right)\right\} - \beta\left(R, m^t\right) C_S\left(m^t\right) - \beta\left(C, m^t\right) C_S\left(m^t\right),$$

$$\begin{aligned} V_1\left(m^t, j\right) &= \beta\left(A, m^t\right)\left(V - C_S\left(m^t\right)\right) - \beta\left(R, m^t\right) C_S\left(m^t\right) \\ &\quad -\beta\left(C, m^t\right)[\max_{a \in M(m^t, j)}\{\beta\left(A, \left(m^t, a\right)\right)\left(V - C_S\left(m^t, a\right)\right) \\ &\quad -\beta\left(R, \left(m^t, a\right)\right) C_S\left(m^t, a\right) - \beta\left(C, m^t\right) C_S\left(m^t, a\right)\}], \end{aligned}$$

and so on. That is, $V_{k+1}(m^t, j)$ is constructed by using $V_k(m^t)$ by replacing terms $\beta\left(C, m^{t+k}\right) C\left(m^{t+k}\right)$ with

$$\begin{aligned} &\beta\left(C, m^{t+k}\right) \max_{a \in M\left(m^{t+k}, j\right)}[\beta\left(A, \left(m^{t+k}, a\right)\right)\left(V - C_S\left(m^{t+k}, a\right)\right) \\ &-\beta\left(R, \left(m^{t+k}, a\right)\right) C_S\left(m^{t+k}, a\right)) - \beta\left(C, \left(m^{t+k}, a\right)\right) C_S\left(m^{t+k}, a\right)]. \end{aligned}$$

Then obviously, the sequence $V_n(m^t, j)$ is an increasing sequence with each satisfies

$$-C_S\left(m^t\right) \leq V_n\left(m^t, j\right) \leq V - C_S\left(m^t\right) \leq V.$$

Hence it converges to some value $V_\infty(m^t)$. Let this value be $V_S(m^t, j)$, and do this for all elements in M. It is a routine work to verify that these satisfy the condition for being the value function. *Q.E.D.*

Proof of Claim 2: Take an arbitrary PBE and let it $\left(\widehat{\alpha}, \widehat{\beta}, \widehat{\varphi}, \widehat{B}\right)$. We will construct a sequence of totally mixed strategy $(\alpha^\lambda, \beta^\lambda, \varphi^\lambda, B^\lambda)$, with $\lambda \in \boldsymbol{N}_+$ (and $\lambda \geq 2$) that converges to $\left(\widehat{\alpha}, \widehat{\beta}, \widehat{\varphi}, \widehat{B}\right)$ as $\lambda \to \infty$. More precisely, we show that there is a sequence $\left(\alpha^\lambda, \beta^\lambda, \varphi^\lambda, B^\lambda\right)$ such that for each $m^t \in M$ and j, it converges to $\left(\widehat{\alpha}\left(m^t, j\right), \widehat{\beta}\left(m^t\right), \widehat{\varphi}\left(m^t\right), \widehat{B}\left(m^t\right)\right) \in R^{9+N}$ in a way such that $\beta^\lambda\left(m^t\right) \geqslant 0$ as a vector in R^3, $\alpha^\lambda\left(S, m^t, j\right) > 0$, $\alpha^\lambda\left(G, m^t, j\right) > 0$ if $j > N_G\left(m^t\right)$, $\alpha^\lambda\left(B, m^t, j\right) > 0$ if $N - j > N_B\left(m^t\right)$, and $\varphi^\lambda\left(m^t\right)$ and $B^\lambda\left(m^t\right)$ are induced by Bayes rule.

In order to do this, we first specify the sequence of sender's first period strategy as follows. We will choose, for each $m_1 \in M$, sufficiently large number $\Psi\left(m_1\right)$, and a function $\varepsilon^\lambda\left(\cdot, \cdot, \varnothing\right) : M \times N \to [0, 1]$. Let the following conditions are satisfied: first, if $j \notin M\left(m_1, j\right)$, $\varepsilon^\lambda\left(m_1, \varnothing, j\right) = 0$. Next think of m_1 such that $\sum_{n=0}^{N} \alpha\left(m_1, \varnothing, j\right) = 0$ (hence m_1 is an off-equilibrium message). If $N_G\left(m_1\right) < \xi$, let it satisfies the followings: for j such that $m_1 \in M\left(\varnothing, j\right)$ and $j < \xi$, we have

$$\frac{f\left(j\right) \varepsilon^\lambda\left(m_1, \varnothing, j\right)}{\sum f\left(n\right) \varepsilon^\lambda\left(m_1, \varnothing, n\right)} = \frac{1}{\left|\left\{n | m_1 \in M\left(m_1, n\right) \text{ and } n < \xi\right\}\right|} \left(\frac{\lambda - 1}{\lambda}\right),$$

and for j such that $m_1 \in M\left(\varnothing, j\right)$ and $j \geq \xi$,

$$\frac{f\left(j\right) \varepsilon^\lambda\left(m_1, \varnothing, j\right)}{\sum\limits_{m_1 \in M\left(m_1, n\right)} f\left(n\right) \varepsilon^\lambda\left(m_1, \varnothing, n\right)} = \frac{1}{\left|\left\{n | m_1 \in M\left(m_1, n\right) \text{ and } n \geq \xi\right\}\right|} \frac{1}{\lambda}.$$

If $N_G\left(m_1\right) \geq \xi$ (and thus $\xi = 1$ and $m_1 = G$), let it satisfies

$$\varepsilon^\lambda\left(m_1, \varnothing, j\right) = 1 \text{ for } j \geq 1 \text{ and } \varepsilon^\lambda\left(m_1, \varnothing, x\right) = 0 \text{ for } j < 0.$$

On the other hand, for each m_1 such that $\sum_{j=0}^{N} \alpha(m_1, \varnothing, j) > 0$ (hence m_1 is an on-equilibrium message), let it satisfies

$$\varepsilon^{\lambda}(m_1, \varnothing, j) = \frac{1}{\Psi(m_1)\lambda} \text{ if } \alpha(m_1, \varnothing, j) = 0,\ \varepsilon^{\lambda}(m_1, \varnothing, j) = 0 \text{ if } \alpha(m_1, \varnothing, j) > 0.$$

Finally, let ε^{λ} be a sufficiently small number. More precisely, let it satisfy

$$\sum_{m_1 \in M(x, \varnothing)} \varepsilon^{\lambda}(m_1, \varnothing, x) < \min_{a \in M(x,\varnothing),\ \alpha(m_1, \varnothing, x) > 0} \alpha(m_1, \varnothing, x).$$

By using $\varepsilon^{\lambda}(m_1, \varnothing, j)$, we will construct $\alpha^{\lambda}(m_1, \varnothing, j)$ as follows. For m_1 such that $\alpha(m_1, \varnothing, j) = 0$, let $\alpha^{\lambda}(m_1, \varnothing, j) = \varepsilon^{\lambda}(m_1, \varnothing, j)$ (hence $\alpha(m_1, \varnothing, j) = 0$ if $m_1 \notin M(j, \varnothing)$) and for m_1 such that $\alpha(m_1, \varnothing, j) > 0$, let

$$\alpha^{\lambda}(m_1, \varnothing, j) = \alpha(m_1, \varnothing, j) - \frac{1}{|\{m_1 | \alpha(m_1, \varnothing, j) > 0\}|} \sum_{m_1 \in M} \varepsilon^{\lambda}(m_1, \varnothing, j). \tag{1.17}$$

Note that $\sum_{m_1 \in M(j, \varnothing)} \widehat{\alpha}^{\lambda}(m_1, \varnothing, j) = 1$, and it constitutes the totally mixed first period strategy for the sender.

In order to construct a sequence of the first period's strategy for the DM, let $\varkappa(m^t) = |\{a \in \{A, R, C\}\ |\beta(a, m^t) > 0\}|$, which is the number of actions that DM takes with a strictly positive probability after message history m^t (and hence less than three). Then for $m_t \in \Delta$, let it be

$$\begin{aligned}
\beta^{\lambda}(a, m_1) &= \widehat{\beta}(a, m_1) - \frac{1}{\varkappa(m_1)\lambda} \text{ if } \widehat{\beta}(a, m_1) > 0, \\
\text{and } \beta^{\lambda}(a, m_1) &= \frac{1}{\varkappa(m_1)\lambda} \text{ if } \widehat{\beta}(a, m_1) = 0.
\end{aligned}$$

And for $m_1 \notin \Delta$, let it be

$$\begin{aligned}
\beta^{\lambda}(Q, m_1) &= \frac{2\lambda - 1}{2\lambda} \text{ and } \beta^{\lambda}(A, m_1) = \beta^{\lambda}(C, m_1) = \frac{1}{2\lambda} \text{ if } N_G\left(m^t\right) < \xi, \\
\text{and } \beta^{\lambda}(A, m_1) &= \frac{2\lambda - 1}{2\lambda} \text{ and } \beta^{\lambda}(Q, m_1) = \beta^{\lambda}(C, m_1) = \frac{1}{2\lambda} \text{ if } N_G\left(m^t\right) \geq \xi.
\end{aligned}$$

$$B_j^{\lambda}(m_1) = \frac{f(j)\,\alpha^{\lambda}(m_1, \varnothing, j)}{\sum_{n=0}^{N} f(n)\,\alpha^{\lambda}(m_1, \varnothing, j)} \textit{ and } \varphi^{\lambda}(m_1|\varnothing) = \sum_j f(j)\,\alpha^{\lambda}(m_1, \varnothing, j).$$

The idea is to make DM's strategy put a strictly high probability of rejection (acceptance) after every off-equilibrium messages with low (high) number of good evidence in the original equilibrium.

Next we define the totally mixed strategy for period 2. Fix $m_1 \in H_1$. For each m_2, choose sufficiently large number $\Psi(m_1, m_2)$, and function $\varepsilon^{\lambda}(\cdot, \cdot, m_1) : M \times N \to [0, 1]$. Let the following conditions are satisfied: If $j \notin M(m_1, j)$, $\varepsilon^{\lambda}(m_1, j, \varnothing)$ is zero. Think of the case in which $\displaystyle\sum_{n=0}^{N} \alpha(m_1, n, \varnothing) = 0$. If $N_G(m_1, m_2) <$

ξ, let it satisfy the following: for j such that $m_2 \in M(m_1, j)$ and $j < \xi$, we have

(1.18)
$$\begin{aligned} & \frac{B_j^\lambda (m_1) \varepsilon^\lambda (m_2, m_1, j)}{\sum B_n^\lambda (m_1) \varepsilon^\lambda (m_2, m_1, j)} \\ = & \frac{1}{|\{n | m_1 \in M((m_1, m_2), n) \text{ and } n < \xi\}|} \left(\frac{\lambda - 1}{\lambda}\right), \end{aligned}$$

and if $j \geq \xi$ and $m_2 \in M(m_1, j)$, let it satisfy

(1.19)
$$\frac{B_j^\lambda (m_1) \varepsilon^\lambda (m_2, m_1, j)}{\sum_{m_1 \in M(m_1, n)} B_j^\lambda (m_1) \varepsilon^\lambda (m_2, m_1, j)} = \frac{1}{|\{n | m_1 \in M((m_1, m_2), n) \text{ and } n \geq \xi\}|} \frac{1}{\lambda}.$$

On the other hand, for each m_2 such that $\sum_{j=0}^{N} \alpha(m_2, m_1, j) > 0$,

$$\begin{aligned} \varepsilon^\lambda (m_2, m_1, j) &= \frac{1}{\Psi(m_1, m_2) \lambda} \text{ if } \alpha(m_2, m_1, j) = 0, \\ \text{and } \varepsilon^\lambda (m_2, m_1, j) &= 0 \text{ if } \alpha(m_2, m_1, j) > 0, \end{aligned}$$

and moreover,

$$\sum_{a \in M(j, (m_1, m_2))} \varepsilon^\lambda (m_2, m_1, j) < \min_{a \in M(j, (m_1, m_2)),\ \alpha(m_2, j, m_1) > 0} \alpha(m_2, m_1, j).$$

By using $\varepsilon^\lambda (m_2, m_1, j)$, we construct $\alpha^\lambda (m_2, m_1, j)$ as follows. For m_2 such that $\alpha(m_2, m_1, j) = 0$, let $\alpha^\lambda (m_2, m_1, j) = \varepsilon^\lambda (m_2, m_1, j)$. For m_2 such that

$\alpha(m_2, m_1, j) > 0$, let

$$\text{(1.20)} \qquad \begin{aligned} & \alpha^\lambda(m_2, m_1, j) \\ = \; & \alpha(m_2, m_1, j) - \frac{1}{|\{m_1 | \alpha(m_2, m_1, j) > 0\}|} \sum_j \varepsilon^\lambda(m_2, m_1, j). \end{aligned}$$

Note that $\sum_a \widehat{\alpha}^\lambda(m_2, m_1, j) = 1$, and it constitutes the totally mixed first period strategy for the sender. Let

$$\begin{aligned} B_j^\lambda(m_1, m_2) &= \frac{B_j^\lambda(m_1)\, \alpha^\lambda(m_2, m_1, j)}{\sum_{n=0}^N B_j^\lambda(m_1)\, \alpha^\lambda(m_2, m_1, j)} \\ \text{and } \varphi(m_2 | m_1) &= \frac{\sum_{n=0}^N B_n(m_1)\, \alpha(m_2, m_1, j)}{\sum_{n=0}^N B_n(m_1)\, \alpha(m_2, m_1, j)}. \end{aligned}$$

For the second period's strategy for the DM, for $m^2 = (m_1, m_2) \in \Delta$, let it be

$$\begin{aligned} \beta^\lambda\left(a, m^2\right) &= \widehat{\beta}\left(a, m^2\right) - \frac{1}{\varkappa(m^2)\,\lambda} \text{ if } \widehat{\beta}\left(a, m^2\right) > 0, \\ \text{and } \beta^\lambda\left(a, m^2\right) &= \frac{1}{\varkappa(m^2)\,\lambda} \text{ if } \widehat{\beta}\left(a, m^2\right) = 0. \end{aligned}$$

And for $m_1 \notin \Delta$, let it be

$$\begin{aligned} \beta^\lambda\left(Q, m^2\right) &= \frac{2\lambda - 1}{2\lambda} \text{ and } \beta^\lambda\left(A, m^2\right) = \beta^\lambda\left(C, m^2\right) = \frac{1}{2\lambda} \text{ if } N_G\left(m^2\right) < \xi, \\ \text{and } \beta^\lambda\left(A, m^2\right) &= \frac{2\lambda - 1}{2\lambda} \text{ and } \beta^\lambda\left(Q, m^2\right) = \beta^\lambda\left(C, m^2\right) = \frac{1}{2\lambda} \text{ if } N_G\left(m^2\right) \geq \xi. \end{aligned}$$

Strategies after period 3 are constructed inductively, and we eventually get the sequence $\left(\alpha^\lambda, \beta^\lambda, B^\lambda, \varphi^\lambda\right)$. Using almost the same procedure, we can construct

$\left(\alpha^{\lambda+1}, \beta^{\lambda+1}, B^{\lambda}, \varphi^{\lambda+1}\right)$. In doing this, replace λ with $\lambda+1$, but use the same $\Psi(m)$ for all $m \in M$. Also, choose $\varepsilon^{\lambda+1}(m_t, j, h)$ in a way that $\max \varepsilon^{\lambda+1}(m_t, j, h) < \frac{1}{2}\varepsilon^{\lambda}(m_t, j, h)$, from which we can get $\lim_{\lambda\to\infty} \varepsilon^{\lambda+1}(m_t, j, m) = 0$ for all j, a, and $m \in M$. It is possible since the left hand side in (1.18) and (1.19) are homogeneous of degree 0 with respect to $\varepsilon^{\lambda}(a, j, h)$. Let

$$(\alpha, \beta, \varphi, B) := \lim_{\lambda\to\infty}(\alpha^{\lambda}, \beta^{\lambda}, \varphi^{\lambda}, B^{\lambda}).$$

It is easy to see that $\left(\widehat{\alpha}, \widehat{\beta}, \widehat{\varphi}, \widehat{B}\right) = (\alpha, \beta, \varphi, B)$ on every $h \in \Delta$, and $\Delta = \Delta'$. To see that $\left(\widehat{\alpha}, \widehat{\beta}, \widehat{\varphi}, \widehat{B}\right)$ is an equilibrium, note that for any $m^t \notin \Delta$, $\widehat{V}_{DM}(m^t) \leq V_{DM}(m^t)$ and hence for all m^t such that $m^t \in \Delta$ and $\widehat{\beta}(C, m^t) = 0$, $\widehat{\beta}(C, m^t) = 0$ is also optimal. This shows the optimality of DM. It is easy to see the optimality of the sender's strategy. *Q.E.D.*

Proof of Proposition 1: Suppose that $m^{\tau} = (m^{\tau-1}, S) \in \Xi$. Because being silent incurs no cost but can persuade the DM, it must hold that $\alpha(S, m^{\tau-1}, j) = 1$ for all j, which implies that

$$\varphi\left(S|m^{\tau-1}\right) = 1 \text{ and } V_{DM}\left(m^{\tau-1}\right) > \mathbb{E}[V_{DM}\left(m^{\tau-1}, m_{\tau}\right)|m^{\tau-1}].$$

These imply $\beta(A, m^{\tau-1}) = 1$. However, this contradicts $(m^{\tau-1}, S) \in \Xi$. That $m^{\tau} = (m^{\tau-1}, B) \notin \Xi$ follows from Theorem 1. *Q.E.D.*

Proof of Theorem 1: To prove the theorem, suppose that we have an equilibrium in which $\beta(A, (m^t, G)) = 0$ for some $(m^t, G) \in \Delta$. Obviously, it must hold that $\beta(C,(m^t,G))>0$. Think about period $t+2$, after message history (m^t,G). Then there must be some message m'_{t+2} such that $\beta\left(A,(m^t,G,m'_{t+2})\right)=0$ and $\left(m^t,G,m'_{t+2}\right)\in\Delta$, since otherwise, we have

$$V_{DM}\left(m^t,G,m_{t+2}\right)=\sum_{j=0}^{N}B_j\left(m^t\right)U_{DM}(A,j,m^t)$$

for all $(m^t,G,m_{t+2})\in\Delta$, which implies

$$V_{DM}\left(m^t,G\right)\geq\sum_{j=0}^{N}B_j\left(m^t,G\right)U_{DM}(A,j,\left(m^t,G\right))$$

$$\begin{aligned}
&> \sum_{m_{t+2}\in M}\varphi\left(m_{t+2}|\left(m^t,G\right)\right)\sum_{j=0}^{N}B_j\left(m^t,G,m_{t+2}\right)U_{DM}(A,j,\left(m^t,G,m_{t+2}\right))\\
&= \sum_{m_{t+2}\in M}\varphi\left(m_{t+2}|\left(m^t,G\right)\right)V_{DM}\left(m^t,G,m_{t+2}\right),
\end{aligned}$$

which contradicts $\beta(A,(m^t,G))=0$. On the other hand, $\beta\left(R,(m^t,G,m'_{t+2})\right)$ has to be smaller than one, since if so sender types from $P\left(A,(m^t,G,m'_{t+2})\right)$ should have chosen S at period $t+1$ after m^t, which contradicts $\left(m^t,G,m'_{t+2}\right)\in\Delta$. This implies that $\beta\left(C,(m^t,G,m'_{t+2})\right)>0$ and, hence,

$$V_{DM}\left(m^t,G,m'_{t+2}\right)=\sum_{m_{t+3}\in M}\varphi\left(m_{t+3}|m^t,G,m'_{t+2}\right)V_{DM}\left(m^t,G,m'_{t+2},m_{t+3}\right).$$

These imply that at period $t+2$, after all on-equilibrium messages, either A is optimal or C is optimal, where in the latter case, the probability of acceptance is zero. Take such a message and let it m_{t+2}, and denote (m^t, G, m_{t+2}) by m^{t+2}. At period $t+3$, there is no $\left(m^{t+2}, m'_{t+3}\right) \in \Delta$ such that $\beta\left(R, m^t, \left(m^{t+2}, m'_{t+3}\right)\right) = 1$ and $\left(m^{t+2}, m'_{t+3}\right) \in \Delta$, since otherwise, senders from $P\left(m^{t+2}, m'_{t+3}\right)$ should have chosen S at period $t+1$ after m^t, which contradicts $\left(m^{t+2}, m'_{t+3}\right) \in \Delta$. It in turn, implies that $\beta\left(A, m^t, \left(m^{t+2}, m'_{t+3}\right)\right) > 0$ or $\beta\left(C, m^t, \left(m^{t+2}, m'_{t+3}\right)\right) > 0$. In sum we have

$$V_{DM}\left(m^{t+2}, m'_{t+3}\right) = \sum_{j=0}^{N} B_j\left(m^{t+2}, m'_{t+3}\right) U_{DM}(A, j, \left(m^{t+2}, m'_{t+3}\right))$$

or

$$V_{DM}\left(m^{t+2}, m'_{t+3}\right) = \sum_{m_{t+4} \in M} \varphi\left(m_{t+4} | m^{t+2}, m'_{t+3}\right) V_{DM}(m^{t+2}, m'_{t+3}, m_{t+4}).$$

Repeat the same reasoning, we can see that for every on-equilibrium history m^τ that is a continuation from (m^t, G), it must hold

$$V_{DM}\left(m^\tau\right) = \max\{\sum_{j=0}^{N} B_j\left(m^\tau\right) U_{DM}(A, j, m^t), \sum_{m_{\tau+1} \in M} \varphi\left(m_{\tau+1} | m^\tau\right) V_{DM}(m^\tau, m_{\tau+1})\}.$$

Then, it is easy to see that

$$V_{DM}\left(m^t, G\right) > \sum_{m_{t+2} \in M} \varphi\left(m_{t+2} | m^t, G\right) V_{DM}(m^t, G, m_{t+1}),$$

which contradicts $\beta\left(C, (m^t, G)\right) > 0$. We can apply the same proof to show that $(m^t, B) \in \Delta$ implies $\beta\left(A, (m^t, B)\right) > 0$.

To see that $(m^t, G) \in \Delta$ implies $\beta(A, (m^t, G)) \geq \delta/V$, suppose that there is some $(m^t, G) \in \Delta$ such that $\beta(A, (m^t, G)) < \delta/V$. If $\beta(C, (m^t, G)) = 0$, then the sender should choose S after m^t rather than G, which contradicts $(m^t, G) \in \Delta$. Hence we have $\beta(C, (m^t, G)) > 0$. Using the same type of argument above, we can see that there must be some type of sender in $P(m^t, G)$ such that he is never accepted after period $t+2$ (he sends S after (m^t, G)). For a such type of sender, it is strictly better to send S after m^t, which gives him at least $-\delta\{N_G(m^t) + N_B(m^t)\}$, rather than G, which gives him only $\beta(A, (m^t, G)) \cdot V - \delta\{N_G(m^t) + N_B(m^t)\} - \delta < -\delta\{N_G(m^t) + N_B(m^t)\}$.

Next, we will show that $(m^t, B) \in \Delta$ implies $\beta(A, (m^t, B)) = \delta/V$. We can apply the same proof as above to show that $(m^t, B) \in \Delta$ implies $\beta(A, (m^t, B)) \geq \delta/V$. Hence suppose that $(m^t, B) \in \Delta$ and $\beta(A, (m^t, B)) > \delta/V$. These and $m^t \in \Delta$ imply that $(m^t, S) \in \Delta$ and $\beta(A, (m^t, S)) = 0$ and, thus, there must be some $j \in P(m^t, S)$ that follows a message history such that he is never accepted after period $t+1$. However rather than that, he can send B after m^t and get the strictly higher expected payoff of

$$\beta\left(A, \left(m^t, B\right)\right) \cdot V - \delta\{N_G\left(m^t\right) + N_B\left(m^t\right)\} - \delta > -\delta\{N_G\left(m^t\right) + N_B\left(m^t\right)\},$$

unless he has no more pieces of bad evidence. If he has no more pieces of bad evidence, it contradicts the optimality of the DM's behavior of not accepting him. Thus, we must have $\beta(A, (m^t, B)) \leq \delta/V$. *Q.E.D.*

Proof of Theorem 2: In order to get a contradiction, suppose that there is some $(m^t, S) \in \Delta$ such that $\beta(A,(m^t,m)) > 0$. If $(m^t, G) \in \Delta$ (or $(m^t, B) \in \Delta$), from Theorem 1 it must hold that $\beta(A,(m^t,m)) > 0$ (or $\beta(A,(m^t,m)) > 0$). In such a case, choosing A at period $t+2$ becomes optimal after all on-equilibrium message at $t+2$ and, hence, contradicts $\beta(C, m^t) > 0$. Thus we must have $(m^t, G) \notin \Delta$ and $(m^t, B) \notin \Delta$. However if it is the case, $\varphi(S|m^t) = 1$ and the DM does not expect to update belief at period $t+2$, which implies $\beta(C,(m^t,m)) = 1$ and, hence, $\beta(A,(m^t,m)) = 0$. *Q.E.D.*

Proof of Proposition 4: Only if direction: Suppose that we have a benchmark strategy equilibrium e with the sender's strategy α, and let $\kappa = N_G(e)$. Obviously, $\beta(A, G^{\kappa+1}) = 1$ and $\beta(A, G^t) = \eta/V$ for $t \leq \kappa$. Because for $t \leq \kappa$, A as well as C are optimal for the DM, from $D1$, it must hold that

$$\sum_{j=0}^{N} B_j\left(G^t\right) U_{DM}(A,j,G^t) = \sum_{m \in M} \varphi(m|m^t) V_{DM}\left(m_t, m\right) \tag{1.21}$$

$$= \varphi(G|G^t) V_{DM}\left(G^{t+1}\right) + \varphi(S|G^t) V_{DM}\left(G^t, m\right).$$

In the benchmark strategy equilibrium, for all $(m,S) \in \Delta$, $V_{DM}(G^t,m) = -C_{DM}(m^t,m) = -\eta t - \eta_S$, because $\beta(R,(G^t,m)) = 1$. Moreover, we have

$$B_j(G^\tau) = \frac{\Pi_{s=1}^{\tau}\left(1-d_j^s\right)f(j)}{\sum_{j\geq\tau}\Pi_{s=1}^{\tau}\left(1-d_j^s\right)f(j)} \tag{1.22}$$

$$\text{and } \varphi(G|G^\tau) = \frac{\sum_{j\geq\tau+1}\left(1-d_j^{\tau+1}\right)\Pi_{s=1}^{\tau}\left(1-d_j^s\right)f(j)}{\sum_{j\geq\tau}\Pi_{s=1}^{\tau}\left(1-d_j^s\right)f(j)}. \tag{1.23}$$

Then by substituting those into (1.21), we can see that (1.8) must hold. Since the optimal action for the DM after $m^1 = G$ is A, (1.9) must hold as well.

If direction: let $\beta(A,G^{\kappa+1}) = 1$ and $\beta(A,G^t) = \eta/V$ for $t \leq \kappa$, $\alpha(B,m^t,j) = 0$ for all $m^t \in H$ and B and φ satisfy (1.22), as well as

$$\begin{aligned} B_N\left(m^t\right) &= 1 \text{ for all } m^t \text{ such that } m^t \neq G^{\kappa+1} \text{ and } N_G\left(m^t\right) \geq \xi, \\ B_{N_G(m^t)}\left(m^t\right) &= 1 \text{ for all } m^t \text{ such that } m^t \neq G^{\kappa+1} \text{ and } N_G\left(m^t\right) < \xi, \\ \varphi(S|m^t) &= 1 \text{ for all } m^t \text{ such that } m^t \neq G^t \text{ for some } t \leq \kappa, \end{aligned} \tag{1.24}$$

where (1.24) corresponds to off-equilibrium beliefs. It is easily seen that D3 is satisfied.

Let the value function for the sender as follows. For sender type $j \leq \kappa+1$,

$$\begin{aligned} V_S\left(m^t,j\right) &= (t-1)\,\delta_G \text{ for } m^t = G^t,\ t < \kappa+1. \\ V_S\left(m^t,j\right) &= -\delta\{N_G\left(m^t\right) + N_B\left(m^t\right)\} \text{ otherwise.} \end{aligned}$$

and for sender type $j > \kappa + 1$,

$$\begin{aligned} V_S\left(m^t, j\right) &= V - \delta\{N_G\left(m^t\right) + N_B\left(m^t\right)\} \text{ for } m^t \text{ such that } N_G\left(m^t\right) \geq \xi, \\ V_S\left(G^{\kappa+1}, j\right) &= V - \delta\left(\kappa+1\right), \end{aligned}$$

$$V_S\left(G^t, j\right) = \sum_{j=0}^{\kappa-t+1} \left(\frac{\delta}{V}\right)\left(1-\frac{\delta}{V}\right)^j \left(V - \delta\left(t+j\right)\right)$$

$$+\left(1-\frac{\delta}{V}\right)^{\kappa-t+1} \left(V - \delta\left(\kappa+1\right)\right) \text{ for } t \leq \kappa+1,$$

and

$$V_S\left(m^t, j\right) = -\delta\{N_G\left(m^t\right) + N_B\left(m^t\right)\} \text{ otherwise.}$$

It is straightforward to verify that V_S satisfies condition for being a value function, given β.

Take a sender with $j \leq \kappa$ number of pieces of good evidence. From above, it follows that

$$V_S\left(\left(m^t, G\right), j\right) = \delta - \delta t = -\delta\left(t-1\right) = V_S\left(\left(m^{k-1}, S\right), j\right).$$

Then he is indifferent between sending G and any S after $m^t = G^t$, with $t \leq j$, which shows that $\alpha\left(\cdot,\cdot,j\right)$ satisfies $D2$ for $j \leq \kappa$. For sender with $j > \kappa$ number of good aspects, on the other hand, we have

$$V_S\left(\left(G^t, G\right), j\right) > V_S\left(\left(G^t, S\right), j\right) = -\delta t > V_S\left(\left(G^t, B\right), j\right) = -\delta t - \delta,$$

for all $t \leq \kappa$, which shows that $\alpha(\cdot, \cdot, j)$ satisfies $D2$ for $j > \kappa$.

For the DM, let the value function be

$$\begin{aligned} V_{DM}\left(m^t\right) &= \sum_{j=0} B_j\left(G^t\right) \mathbb{E}[\theta|j] - \eta t \text{ for all } G^t, \\ V_{DM}\left(m^t\right) &= \eta\{N_G\left(m^t\right) + N_B\left(m^t\right)\} \\ \text{for all } m^t \text{ such that } m^t &\neq G^t \text{ and } N_G\left(m^t\right) < \xi, \end{aligned}$$

and

$$\begin{aligned} V_{DM}\left(m^t\right) &= E[\theta|N] - \eta\{N_G\left(m^t\right) + N_B\left(m^t\right)\} \\ \text{for all } \left(m^t, m\right) \text{ such that } m^t &\neq G^t \text{ and } N_G\left(m^t\right) \geq \xi. \end{aligned}$$

It is straightforward to show that V_{DM} is a value function.

Think about the decision at the first period. The expected payoff from not continuing is $\max\{E[\theta], 0\}$. On the other hand, the expected payoff from entering period 1 and making decision at period 1 is given by (1.9). Hence from D1, $\beta(C, \varnothing) = 1$ is optimal. Think about the decision after $m^t = G^t$ and $t \leq \kappa$. From the description of sender's strategy, the DM never receives B by choosing C and thus we can calculate

$$V_{DM}\left(G^t\right) = \mathbb{E}[U_{DM}(A,j,t)|G^t] - t\eta = \sum_{m\in\{G,B,S\}} \varphi(m|G^t)V_{DM}\left(G^t,m\right),$$

and hence $\beta\left(A,G^t\right)$ satisfies $D1$. It is easy to see that $D1$ is satisfied for the other cases as well, because in such cases we have

$$V_{DM}\left(m^t\right) > \sum_{m\in\{G,B,S\}} \varphi(m|G^t)V_{DM}\left(G^t,m\right)$$

and

$$\beta\left(R,m^t\right) = 1 \text{ if } N_G\left(m^t\right) < \xi \text{ and } \beta\left(A,m^t\right) = 1 \text{ if } N_G\left(m^t\right) \geq \xi.$$

Q.E.D.

In order to prove Theorem 3, we first prove several lemmata. In the following, we fix an equilibrium (α,β,φ,B).

Lemma 3. *If* $(m^t,G)\in\Delta$ *and* $\beta\left(A,(m^t,G)\right)<1$, *there must be some* $j\in P\left((m^t,G)\right)$ *such that*

$$\max_{m\in M((m^t,G),j)} V_S\left(\left(m^t,G,m\right),j\right) = -\delta\{N_G\left(m^t,G\right)+N_B\left(m^t,G\right)\}.$$

Proof. It is seen from the proof of Proposition 1. □

Lemma 4. *For all* $n^s\in\Xi$ *and* $k^\tau\in\Xi$, *it holds that* $N_G\left(n^s\right)=N_G\left(k^\tau\right).$

Proof. Take $n^{\tau} = (n_1, .., n_s) \in \Xi$ and $k^t = (k_1, .., k_{\tau}) \in \Xi$ and suppose that $N_G(n^s) > N_G(k^{\tau})$. Let n^t be a longest sub-history of n^{τ} and k^{τ} such that $n^t = k^t$. Note that this can be $\varnothing$ if $n_1 \neq k_1$. We can show that $N_G(n^s) - 1 \in P(n^{s-1})$. To see this, note that from $N_G(n^s) > N_G(k^{\tau})$, there must be some $n^{s'} \prec n^s$ such that $N_G(n^{s'}) = N_G(n^s) - 1$, which means that the DM knows that the sender has at least $N_G(n^s) - 1$ number of good evidence after history $n^{s'}$. Since $\beta(C, n^{s-1}) > 0$, the DM has to expect some type from $P(n^{s-1})$ chooses S, which can be only type $N_G(n^s) - 1$ sender. This implies that $V_S((n^t, n_{t+1}), N_G(n^s) - 1) \geq V_S((n^t, k_{t+1}), N_G(n^s) - 1)$. However, since the type $N_G(n^s) - 1$ sender can follow the path k^s instead and $\beta(A, k^s) = 1$, this implies that

$$\begin{aligned} & \sum_{l=t}^{\tau-1} \left(1 - \beta\left(A, n^l\right)\right)^{t'-t} \beta\left(A, n^l\right)^l U_S\left(A, n^{t+1}\right) \\ > & \sum_{l=t}^{\tau-1} \left(1 - \beta\left(A, k^l\right)\right)^{t'-t} \beta\left(A, k^l\right)^l U_S\left(A, k^{t+1}\right), \end{aligned}$$

where each term is the expected payoff for the sender summed up before period $\tau - 1$. In other words, the path n^s has a higher probability of acceptance than k^s does in the early period of the path. However, this implies that

$$V_S\left((n^t, n_{t+1}), N_G(k^{\tau-1}) - 1\right) > V_S\left(\left(n^t, k_{t+1}\right), N_G(k^{\tau}) - 1\right)$$

and, hence, $N_G(k^{\tau}) - 1 \notin P(k^{\tau-1})$. However this is a contradiction because we can show that $N_G(k^{\tau}) - 1 \in P(k^{\tau-1})$, from the same reasoning used above. $\square$

Take an equilibrium $(\alpha, \beta, \varphi, B)$. Let $\Delta_1 \in \Delta$ be a subset of Δ such that each $m \in \Delta_1$ contains only one G as its final element, that is, m takes the form of $(G), (m_1, G), (m_1, m_2, G)$. Define Δ_j in a similar manner, that is, each $m \in \Delta_1$ contains j number of G and also its final element is G. We can show the following lemma.

Lemma 5. *Take an equilibrium. Then for all $m \in \Delta_j$ and $m' \in \Delta_j$, $\beta(A, m) = \beta(A, m')$ and $\beta(C, m) = \beta(C, m')$ for all j.*

Proof. We first show that for all $m \in \Delta_1$ and $m' \in \Delta_1$, $\beta(A, m) = \beta(A, m')$. To see this, suppose that there is some pair $m \in \Delta_1$ and $m' \in \Delta_1$ such that $\beta(A, m) > \beta(A, m')$. Then obviously, $1 \notin P(m')$. Also because $\beta(A, m') < 1$, from Lemma 3 there must be some $j \in P(m')$ such that $V_S((m', m), j) = -\delta\{N_G(m') + N_B(m')\}$. However, then we have $V_S(m, j) > V_S(m', j)$, which contradicts $j \in P(m')$. Then we can prove that $\beta(A, m) = \beta(A, m')$ for $m \in \Delta_j$ and $m' \in \Delta_j$ inductively. We can also prove that $\beta(C, m) = \beta(C, m')$ for all $m \in \Delta_j$ and $m' \in \Delta_j$ in a similar manner. □

Lemma 6. *For all $m \in \Delta$ and $m' \in \Delta$ such that $\beta(A, m) = \beta(A, m') = 1$, $N_B(m) = N_B(m')$.*

Proof. Suppose that $m \in \Delta$ and $m' \in \Delta$, $\beta(A, m) = \beta(A, m') = 1$ but $N_B(m) > N_B(m')$. Then it is immediately seen that from Lemma 5, for all $j \in P(m)$, we have $V_S(m'_1, j) > V_S(m_1, j)$, where $m'_1 \in \Delta_1$, $m'_1 \prec m'$, $m_1 \in \Delta_1$, $m_1 \prec m$ (note that history m' reaches the node of sure acceptance faster than m does), which contradicts $m \in \Delta$. □

This lemma allows us to define the function that gives the number of pieces of good and bad evidence to persuade the DM for a given equilibrium, which we denote by $N_G(e)$ and $N_B(e)$.

Proof of Theorem 3 is divided into three parts.

Lemma 7. *For all* $e \in \mathcal{P}(\eta, \eta_S, \delta)$, $\beta(A, (m^t, G)) = 1$ *or* $\beta(A, (m^t, G)) = \delta/V$ *for all* $(m^t, G) \in \Delta$.

Proof. Suppose that the set of message history $M^+ = \{m^t | (m^\tau, G) \in \Delta$ such that $\beta(A, (m^\tau, G)) \in (\delta/V, 1)$ is non-empty. Also, let M^{++} be the set of the smallest elements of M^+ with respect to the order of $\prec$, that is, if $m^s \in M^{++}$ there is no $m^t \in M^+$ such that $m^t \prec m^s$. Note that from Lemma 6, $N_G(m^s) = N_G\left(m^{s'}\right) = \gamma$ for some γ for all $m^s \in M^{++}$ and $m^{s'} \in M^{++}$. Also, for all $m^\tau \in \Xi$ we can find $m^t \in M^{++}$ such that $m^t \prec m^\tau$.

In the equilibrium e, for every sender type j such that $N - \gamma > j \geq \gamma$, there is a $m_1 \in M(\varnothing, j)$ such that $V_s(m_1, j) > 0$, that is, every sender types from $\{N_G(M^{++}), .., N - N_B(M^{++}) - 1\}$ can get strictly positive payoffs by following a message path from M^{++} (note that from proposition 1, each time a sender communicates an aspect, he is accepted with a probability of as high as η/V, which is enough to recover the cost of communication one time).

In this case, we can construct equilibrium $\widehat{e} = \left(\widehat{\alpha}, \widehat{\beta}, \widehat{B}, \widehat{\varphi}\right)$ in a way that $\left(\widehat{\alpha}, \widehat{\beta}, \widehat{B}, \widehat{\varphi}\right) = (\alpha, \beta, B, \varphi)$ except $\widehat{\beta}\,(A, m^s) = 1$ for all $m^s \in M^{++}$. To see that this is an equilibrium, first note that

$$\begin{aligned} V_S^e(m, m^s, x) &= V_S^{\widehat{e}}(m, m^s, x) \text{ for all } m^s \preceq m^\tau \text{ such that } m^\tau \notin \Xi^{\widehat{e}} = M^{++} \\ V_S^e(m, m^s, x) &= V_S^{\widehat{e}}(m, m^s, x) \text{ for all } m^s \preceq m^\tau \text{ such that } m^\tau \in \Xi^{\widehat{e}} = M^{++}, \end{aligned}$$

and $V_S^{\widehat{e}}(m, m^s, j) = V_S^{\widehat{e}}(m, m^s, j)$ for all m^s, m_{s+1}, and $m_{s+1}^{'}$ such that there exist m^τ and $m^{\tau'}$ such that $(m^s, m_{s+1}) \preceq m^\tau \in \Xi^{\widehat{e}}$ and $\left(m^s, m_{s+1}^{\prime}\right) \preceq m^{\tau'} \in \Xi^{\widehat{e}}$. Those imply that $\widehat{\alpha}$ satisfies D-2 given $\widehat{\beta}$, from the fact that α satisfies D-2 given β. On the other hand, we have $V_{DM}^{\widehat{e}}\,(m^t) \leq V_{DM}^{e}\,(m^t)$ for all m^t such that there is $m^\tau \in \Xi^{\widehat{e}}$ and $m^t \prec m^\tau$. Also, $V_{DM}^{\widehat{e}}\,(m^t) = V_{DM}^{e}\,(m^t)$ if there is not such m^τ, and hence $\widehat{\beta}$ satisfies D-1 given $\left(\widehat{\varphi}, \widehat{B}\right)$, from the fact that β satisfies D-1 given (φ, B). Hence $\widehat{e}$ is an equilibrium. We can also see that $\widehat{e}$ Pareto dominates e since it has strictly higher payoff for sender types $\{N_G\,(M^{++})\,, .., N - N_B\,(M^{++}) - 1\}$ while giving exactly the same payoff for other types of sender and the DM. $\square$

Lemma 8. *For all* $e \in \mathcal{P}\,(\eta, \eta_S, \delta)\,,$ *it holds that* $\beta\,(R, (m^t, S)) = 1$ *for all* $(m^t, S) \in \Delta.$

Proof. From Lemma 7, we assume that $\beta\,(A, (m^t, G)) \in \{1, \delta/V\}$ for all $(m^t, G) \in \Delta$ holds in any equilibrium we consider.

Take $e \in \mathcal{E}(\eta, \eta_S, \delta)$ such that $\beta(R, (m^t, S)) < 1$ for some $(m^t, S) \in \Delta$. Let Ψ be a set of such m^t. In order to get a contradiction, suppose that Ψ is not empty. From Lemma 4, $N_G(m^\tau) = N_G\left(m^{\tau'}\right) = N_G(e)$ and $N_B(m^\tau) = N_B\left(m^{\tau'}\right) = N_B(e)$ for all $m^\tau \in \Xi^e$ and $m^{\tau'} \in \Xi^e$. From the equilibrium condition, for all $m^{t+1} = (m^t, m_{t+1}) \in \Delta$ such that $m_{t+1} \in \{G, B\}$, we have

$$
\begin{aligned}
(1.25) \qquad & -\sum_{j\geq 1}^{N}\{1 - \sum_{m\in\{G,B\}} \alpha\left(m, m^{t+1}, j\right)\}\alpha\left(m^{t+1}, j\right) f(j)\left(\mathbb{E}[\theta|j] + \eta_S\right) \\
\geq \; & \eta \sum_{n\geq 1}^{N} \sum_{m\in\{G,B\}} \alpha\left(m, m^{t+1}, j\right) f(j),
\end{aligned}
$$

which follows from $\sum_{j=0}^{N} B_j(m^t)\, U_{DM}(A, j, m^t) = \sum_{m\in M} \varphi(m|m^t) V_{DM}(m_t, m)$ and Proposition 1. Note that this holds with equality when $\beta(R, (m^{t+1}, S)) = 1$ and strict inequality otherwise, and from the maintained assumption we have (1.25) with strictly inequality for at least one $m^t \in \Delta$. Let $\Gamma(e)$ be a set of message histories such that their last element is G or B and rest of elements are S, that is, it is the set of message history that an piece of evidence is communicated for the first time. Then, we can construct following equilibrium $\widehat{e}$: there is only one acceptance history $m^\tau = (B, .., B, G, .., G)$, that is, $\widehat{\beta}(R, m^t) = 1$ for all $m^t \npreceq m^\tau$, $\widehat{\beta}(A, m^\tau) = 1$. The sender's strategy $\widehat{\alpha}$ satisfies (1.8) with inequality for all $m^t \prec m^\tau$. Then, it is possible to choose $\widehat{\alpha}$ in such a way that

$\widehat{\alpha}(B,\varnothing,j) \leq \sum_{m\in\Gamma(e)} \alpha(m,j)$ for all $j \leq N_G(m^\tau)$ and at least one strictly inequality. However this implies that

$$\begin{aligned} V_{DM}^{\widehat{e}}(\varnothing) &= \sum_{n\geq 1}^{N} \widehat{\alpha}(B,\varnothing,j)(\mathbb{E}[\theta|j]-\eta) - \eta_S \sum_{n\geq 1}^{N} (1-\widehat{\alpha}(B,\varnothing,j)) \\ &> \sum_{n\geq 1}^{N} \sum_{m\in\Gamma(e)} \alpha(m,j)(\mathbb{E}[\theta|j]-\eta) - \eta_S \sum_{n\geq 1}^{N} (1-\sum_{m\in\Gamma(e)} \alpha(m,j)) \\ &\geq V_{DM}^{e}(\varnothing). \end{aligned}$$

□

Lemma 9. *If there is an equilibrium e such that $(m^t, B) \in \Delta$ for some m^t, then there is a benchmark strategy equilibrium e' that gives the DM a strictly higher payoff.*

Proof. Take an equilibrium $e = (\alpha, \beta, B, \varphi)$ that involves communicating a piece of bad aspect one time. From Lemma 6, all the equilibrium acceptance path involves at least one time communication of a piece of bad evidence, which implies that type N sender has to drop before reaching "acceptance for sure" nodes. Without loss of generality, assume that $m^\tau = (m_1, .., m^\tau) \in \Xi$ contains no S, that is, $m_t \in \{G, B\}$ and thus $N_G(e) + 1 = \tau$. Moreover, from Lemma 7, we can assume that for all $(m^t, G) \in \Delta$, $\beta(A, (m^t, G)) \in \{\eta/V, 1\}$. Then, $V_{DM}(m_1) =$

$$\sum_{j\geq 1}^{N} \alpha(m_1,\varnothing,j) f(j)(\mathbb{E}[\theta|j]+\eta) - \eta_S \sum_{j\geq 1}^{N} \{1-\alpha(m_1,\varnothing,j)\} f(j)$$

$$= \eta \sum_{t=1}^{2} \sum_{j\geq 1}^{N} \Pi_{s=1}^{t} \alpha\left(m_s, m^{s-1}, j\right) f(j) + \sum_{j\geq 1}^{N} \Pi_{s=1}^{k} \alpha\left(m_s, m^{s-1}, j\right) f(j)\, \mathbb{E}[\theta|j]$$

$$= \sum_{j\geq \kappa}^{N-1} f(j)\, \mathbb{E}[\theta|j] - \eta \sum_{t=1}^{\tau} \sum_{j\geq 1}^{N} \Pi_{s=1}^{\tau} \alpha\left(m_s, m^{s-1}, j\right) f(j),$$

where we used the relation

$$-\sum_{j\geq 1}^{N} \{1 - \alpha\left(m_{t+1}, m^{t}, j\right)\} \Pi_{s=1}^{t} \alpha\left(m_s, m^{s-1}, j\right) f(j) \left(\mathbb{E}[\theta|j] + \eta_S\right) \tag{1.26}$$

$$= \eta \sum_{j\geq 1}^{N} \Pi_{s=1}^{t+1} \alpha\left(m_s, m^{s-1}, j\right) f(j), \tag{1.27}$$

for all m_{t+1}. Since type N sender has to drop at some period and $V_{DM}(m_1) \geq 0$, this implies that

$$-\frac{\sum_{j\geq\kappa}^{N-1} f(j) \left(\mathbb{E}[\theta|j] + \eta_S\right)}{\sum_{j\geq\kappa}^{N-1} f(j)} \geq \eta \frac{\sum_{t=1}^{\tau} \sum_{j\geq 1}^{N} \Pi_{s=1}^{t} \alpha\left(m_s, m^{s-1}, j\right) f(j)}{\sum_{j\geq\kappa}^{N-1} f(j)} \tag{1.28}$$

We will first show that there exists a benchmark strategy equilibrium such that it gives a strictly higher payoff to the DM than equilibrium e does. We moreover assume that $m_1 = B$.

Take a benchmark strategy equilibrium $e' = (\alpha', \beta', \varphi', B')$ that maximizes $N_G(e')$. First, suppose that we have $N_G(e') \geq N_G(e) - 1$. Define

$$\begin{aligned}
\Delta\alpha(t) &= \Pi_{s=1}^{t-1}\alpha'\left(m_s, m^{s-1}, j\right)\left(1-\alpha'\left(m_t, m^{t-1}, j\right)\right)f(j) \\
&\quad -\Pi_{s=1}^{t-1}\alpha\left(m_s, m^{s-1}, j\right)\left(1-\alpha'\left(m_t, m^{t-1}, j\right)\right)f(j) \\
\text{and } \Delta E(j) &= \Pi_{s=1}^{t-1}\alpha'\left(m_s, m^{s-1}, j\right)\left(1-\alpha'\left(m_t, m^{t-1}, j\right)\right)\mathbb{E}[\theta|j] \\
&\quad -\Pi_{s=1}^{t-1}\alpha\left(m_s, m^{s-1}, j\right)\left(1-\alpha'\left(m_t, m^{t-1}, j\right)\right)\mathbb{E}[\theta|j]
\end{aligned}$$

If $N_G(e') = N_G(e) - 1$, because we cannot construct an equilibrium in a way that k pieces of evidence are communicated, we must have $-f(\kappa-1)\{\mathbb{E}[\theta|\kappa-1] + \eta_S\} < \eta\sum_{j\geq\kappa}^{N} f(j)$. To see this, note that If $-f(N_G(e))\{\mathbb{E}[\theta|N_G(e)] + \eta_S) \geq \eta\sum_{j\geq N_G(e)+1}^{N}$, by the fact that we can construct an equilibrium e' with $N_G(e') = N_G(e)$ implies that we can construct another equilibrium e' with $N_G(e') = N_G(e)$ by setting

$$\alpha'\left(G, G^{k-1}, N_G(e)\right) = -\eta\sum_{j\geq N_G(e)+1}^{N} /\{\mathbb{E}[\theta|N_G(e)] + \eta\},$$

and $\alpha' \leq \alpha(G, G^{n-1}, j)$ for all $n,\ j$, which is a contradiction. Then because we have

$$\Pi_{s=1}^{\tau}\alpha\left(m_s, m^{s-1}, j\right)f(N_G(e)-1)\{\mathbb{E}[\theta|N_G(e)-1] + \eta_S\} = \eta\sum_{j\geq\kappa}^{N-1} f(j)$$

from the construction of equilibrium e, we have

$$\Delta E(\tau) < \eta f(N) \text{ and } \Delta\alpha(\tau) < \frac{1}{E(N_G(e)-1)}\eta f(N) \tag{1.29}$$

On the other hand, if $N_G(e') \geq \kappa$, (1.29) follows immediately. Those in turn imply that

$$\Delta E(\tau-1) < \eta f(N) + \Delta\alpha(\tau)\eta \text{ and } \Delta\alpha(\tau-1) < \frac{1}{E(N_G(e)-1)}\Delta\mathbb{E}(\tau-1).$$

By continuing this, we will get

$$\Delta E(t) < \eta f(N) + \sum_{s=t+1}^{\tau}\Delta\alpha(s) \text{ for all } j > 1.$$

Next we show that

$$\Delta E(t) < \eta f(N)\frac{\sum_{j\geq 1}^{N}\Pi_{s=t}^{\tau}\alpha(m_s, m^{s-1}, j) f(j)}{\sum_{j\geq\kappa}^{N-1} f(j)}. \tag{1.30}$$

To see this, note that

$$\begin{aligned}
&\frac{\sum_{j\geq 1}^{N}\Pi_{s=1}^{t}\alpha(m_s, m^{s-1}, j) f(j)}{\sum_{j\geq\kappa}^{N-1} f(j)} \\
= &\frac{\sum_{s=t}^{\tau}\sum_{j\geq 1}^{N}\Pi_{n=t}^{s}(1-\alpha(m_{s+1}, m^s, j))\alpha(m_n, m^{n-1}, j) f(j) + \sum_{j\geq\kappa}^{N-1} f(j)}{\sum_{j\geq\kappa}^{N-1} f(j)} \\
= &\frac{\Gamma(t+1) + ... + \Gamma(\tau-1) + \Gamma(\tau)}{\sum_{j\geq\kappa}^{N-1} f(j)},
\end{aligned}$$

where we defined $\Gamma(r) = \sum_{s=1}^{r}\sum_{j\geq 1}^{N}\Pi_{s=1}^{r}(1-\alpha(m_r, m^{r-1}, j))\alpha(m_{s-1}, m^s, j) f(j)$.

On the other hand we can show that

$$
\begin{aligned}
\Delta E(\tau) &< \eta f(N) = \eta f(N) \frac{\sum_{j\geq\kappa}^{N-1} f(j)}{\sum_{j\geq\kappa}^{N-1} f(j)} = \eta f(N)\Gamma(\tau), \\
\Delta E(\tau-1) &< \eta f(N) + \Delta\alpha(\tau)\eta = \eta f(N)\Gamma(\tau) + \eta f(N) \frac{\eta}{\mathbb{E}(\kappa-1)} \\
&\leq \eta f(N)\Gamma(\tau) + \eta f(N) \frac{\Pi_{s=t}^{\tau-1}\alpha(m_s, m^{s-1}, \kappa-1) f(\kappa-1)}{\sum_{j\geq\kappa}^{N-1} f(j)} \\
&= \eta f(N)\Gamma(\tau) + \eta f(N)\Gamma(\tau-1),
\end{aligned}
$$

and more generally, $\Delta E(r) < \eta f(N) \sum_{j\geq r}^{\tau} \Gamma(\tau)$, which implies

$$
\Delta E(r) < \frac{\sum_{j\geq 1}^{N} \Pi_{s=1}^{r+1}\alpha(m_s, m^{s-1}, j) f(j)}{\sum_{j\geq\kappa}^{N-1} f(j)} \text{ for all } r \in \{1, .., \kappa-1\}
$$

Obliviously, we have

$$
0 \leq V_{DM}^{e}(\varnothing) - V_{DM}^{e'}(\varnothing) < \sum_{j=1}^{\kappa-1} \Delta E(\tau) - \mathbb{E}(\theta|N)
$$

and thus (1.30) implies

$$
\mathbb{E}(\theta|N) > \frac{\sum_{t=1}^{\tau} \sum_{j\geq 1}^{N} \Pi_{s=1}^{t}\alpha(m_s, m^{s-1}, j) f(j)}{\sum_{j\geq\kappa}^{N-1} f(j)},
$$

which contradicts (1.28) since $\mathbb{E}[\theta|j]$ is increasing with j.

Next, think of the case in which we have a benchmark strategy equilibrium $e' = (\alpha', \beta', \varphi', B')$ such that $N_G(e') = \kappa - 2$ but not $\kappa - 1$. The fact that we cannot construct an equilibrium in a way that $k-1$ aspects are communicated implies

$$-f(\kappa-1)\,\mathbb{E}[\theta|\kappa-1]-f(\kappa-2)\,\mathbb{E}[\theta|\kappa-2]<2\eta\sum_{j\geq\kappa}^{N}f(j)+\eta^{2}\frac{\sum_{j\geq\kappa}^{N}f(j)}{\mathbb{E}[\theta|\kappa-1]}.$$

Then by using the same argument we can get a contradiction. Other cases can be treated similarly. □

Theorem 3 follows immediately from Lemma 7, 8, and 9. *Q.E.D.*

Proof of Proposition 4: Fix model's parameter values (η,η_S,δ). In a benchmark strategy equilibrium, the sender's strategy α can be seen as an element of $[0,1]^{2(\xi-1)}$ (each represents the probability of communicating G), since $\alpha(G,G^t,j)=1$ for all $j\geq\xi$ and $t\leq\xi$. Let $\mathcal{E}\subset[0,1]^{2(\xi-1)}$ be the set of the sender's strategy that is supported as an equilibrium, and let $\mathcal{E}_\lambda\subset\mathcal{E}$ for $\lambda\in\{1,...,\xi\}$ be a subset of the sender's strategy that is supported by an equilibrium e such that $N_G(e)=\lambda$. We first show that set $\mathcal{E}_\lambda$ is closed in the usual sense of Euclidean topology. However this is easy because if we take a sequence $\{\alpha^n\}_{n=1}^{\infty}$ from $\mathcal{E}$ that converges to α, it holds that

$$\begin{aligned}&-\sum_{j\geq x}^{N}\{1-\alpha^n\left(G,G^t,j\right)\}\Pi_{s=1}^{t}\alpha^n\left(G,G^{s-1},j\right)f(j)\left(\mathbb{E}[\theta|j]+\eta_S\right)\\ =\;&\eta\sum_{j\geq x}^{N}\Pi_{s=1}^{t+1}\alpha^n\left(G,G^{s-1},j\right)f(j),\end{aligned}$$

for all n and $t\leq\lambda$, which implies that the same condition holds for α from $\lim_{n\to\infty}\alpha^n\to\alpha$. Hence $\alpha\in\mathcal{E}_\lambda$ and $\mathcal{E}_\lambda$ is closed. Then $V_{DM}(\varnothing)$, calculated as

$\sum_{j\geq 1}^{N}\alpha\left(G,\varnothing,j\right)f\left(j\right)\left(\mathbb{E}[\theta|j]-\eta\right)$, is a continuous function on $\mathcal{E}_{\lambda}$, which is closed and bounded, and hence has a maximum point in $\mathcal{E}_{\lambda}$. Then $V_{DM}\left(\varnothing\right)$ has the maximum on $\mathcal{E}$, which is a finite union of $\mathcal{E}_{\lambda}$.

We next prove the uniqueness. Towards this end, suppose that we have two different equilibria e and e' such that $V_{DM}^{e}\left(e\right)=V_{DM}^{e}\left(e\right)$ and those maximize the DM's expected payoff (best equilibria). From Theorem 4, which will be proved below, $N_{G}\left(e\right)=N_{G}\left(e'\right)=\lambda$ and $\alpha\left(G,G^{t},j\right)=\alpha'\left(G,G^{t},j\right)=1$ for all $t\leq N_{G}\left(e\right)$ and $j\geq N_{G}\left(e\right)$. Moreover, since (1.11) must hold at period $\lambda-1$, we must have $\alpha\left(G^{\lambda-1},\lambda-1\right)=\alpha'\left(G^{\lambda-1},\lambda-1\right)$. Let $h<\lambda-1$ be the biggest t such that $\alpha\left(G^{t},j\right)\neq\alpha'\left(G^{t},j\right)$ for some $j\leq h$, and let l be the biggest $l\leq h$ such that $\alpha\left(G^{h},j\right)\neq\alpha'\left(G^{h},j\right)$. Without loss of generality, let $1\geq\alpha\left(G^{h},j\right)>\alpha'\left(G^{h},j\right)$. Since we have (1.8) for period h, there must be some q such that $\alpha\left(G^{h},q\right)<\alpha'\left(G^{h},q\right)\leq 1$. Then we can find a pair of strictly positive numbers $\varepsilon<\alpha'\left(G^{h},q\right)-\alpha\left(G^{h},q\right)$, $\delta<\alpha\left(G^{h},l\right)-\alpha'\left(G^{h},l\right)$, and ϵ such that

$$
\begin{aligned}
& \{1-\alpha\left(G,G^{h-1},q\right)\}\Pi_{s=1}^{h-2}\alpha^{n}\left(G,G^{s-1},q\right)f\left(q\right)\{\mathbb{E}[\theta|q]+\eta_{S}\} \\
& +\{1-\alpha\left(G,G^{h-1},l\right)\}\Pi_{s=1}^{h-2}\alpha^{n}\left(G,G^{s-1},l\right)f\left(l\right)\{\mathbb{E}[\theta|l]+\eta_{S}\} \\
= & \{1-\alpha\left(G,G^{h-1},q\right)-\varepsilon\}\left(\alpha\left(G^{h-1},q\right)-\epsilon\right)f\left(q\right)\{\mathbb{E}[\theta|q]+\eta_{S}\} \\
& +\{1-\alpha\left(G,G^{h-1},l\right)+\delta\}\left(\alpha\left(G^{h-1},l\right)-\epsilon\right)f\left(l\right)\{\mathbb{E}[\theta|l]+\eta_{S}\},
\end{aligned}
$$

because $\mathbb{E}[\theta|j]$ is strictly increasing with j. Then think of the following strategy after after period $h-1:$ $\widehat{\alpha}\left(G^{t},j\right)=\alpha\left(G^{t},j\right)$ for all $j\neq l,q$, and

$$\begin{aligned}
\widehat{\alpha}\left(G,G^{h-1},q\right) &= \alpha\left(G,G^{h-1},q\right)+\varepsilon \\
\text{and } \widehat{\alpha}\left(G,G^{h-1},q\right) &= \alpha\left(G,G^{h-1},q\right)-\delta; \\
\Pi_{s=1}^{h-2}\widehat{\alpha}\left(G,G^{s-1},q\right) &= \alpha\left(G^{h-1},q\right)-\epsilon \\
\text{and } \Pi_{s=1}^{h-2}\widehat{\alpha}\left(G,G^{s-1},l\right) &= \alpha\left(G,G^{h-1},l\right)-\epsilon.
\end{aligned}$$

Then obviously, $\sum_{j}\Pi_{s=1}^{h-2}\widehat{\alpha}\left(G,G^{s-1},j\right)f\left(j\right)<\sum_{j}\Pi_{s=1}^{h-2}\alpha\left(G,G^{s-1},j\right)f\left(j\right),$ and equilibrium condition (1.8) is satisfied after period $h-1$. Then, by taking the same steps as in the proof of Theorem 3, we can construct an equilibrium that can support such strategy after $h-1$ in a way such that $\widehat{\alpha}\left(G,\varnothing,j\right)f\left(j\right)\leq\alpha\left(G,\varnothing,j\right)$ for all j holds with at least one strict inequality. Obviously, such an equilibrium attains a strictly higher $V_{DM}\left(\varnothing\right)$ than e does, and it is a contradiction. *Q.E.D.*

Proof of Theorem 4: To get a contradiction, suppose that equilibrium e is the best equilibrium for the DM, but there is another equilibrium $\widehat{e}$ such that $N_{G}\left(\widehat{e}\right)>N_{G}\left(e\right)$. Obviously, $N_{G}\left(\widehat{e}\right)\leq\xi$. Then from the fact that e being an equilibrium, we have (1.8) for all $t<N_{G}\left(e\right)$ and $\Pi_{s=1}^{t}\alpha\left(G,G^{s-1},j\right)=1$ for all

$j \geq N_G(e)$. On the other hand, since $N_G(\widehat{e}) > N_G(e)$, we have

$$\begin{aligned} & -\sum_{j\geq 1}^{N} \{1 - \widehat{\alpha}\left(G, G^t, j\right)\} \Pi_{s=N_G(e)}^{t} \widehat{\alpha}\left(G, G^{s-1}, j\right) f(j) \left(\mathbb{E}[\theta|j] + \eta_S\right) \\ = \;& \eta \sum_{j\geq 1}^{N} \Pi_{s=N_G(e)}^{t+1} \widehat{\alpha}\left(G, G^{s-1}, j\right) f(j) > 0, \end{aligned}$$

for all $t < N_G(\widehat{e})$ and

$$\Pi_{s=1}^{t} \widehat{\alpha}\left(G, G^{s-1}, j\right) = 1 \text{ for all } j \geq N_G(\widehat{e}) \text{ and } t \leq N_G(\widehat{e}).$$

Then from the assumption 1, it holds that $\Pi_{s=1}^{N_G(e')} \widehat{\alpha}\left(G, G^{s-1}, N_G(e') - 1\right) < 1$. Hence

(1.31)
$$-\eta \sum_{j\geq N_G(e)}^{N} \Pi_{s=1}^{N_G(e)} \alpha\left(G, G^{s-1}, j\right) f(j) - \eta \sum_{j\geq N_G(e)}^{N} \Pi_{s=1}^{N_G(e)} \widehat{\alpha}\left(G, G^{s-1}, j\right) f(j) > 0.$$

Then we can construct a benchmark strategy equilibrium $e' = (\alpha', \beta', \varphi', B')$ in the following way.

$$\alpha'(,,j) = \widehat{\alpha}(,,j) \text{ for all } j \geq N_G(\widehat{e}) - 1, \; \beta' = \widehat{\beta},$$

(1.8) for all $t < N_G(\widehat{e})$, and

$$\alpha'\left(G, G^t, j\right) \leq \alpha\left(G, G^t, j\right) \text{ for all } j < N_G(\widehat{e}) \text{ and } t < N_G(\widehat{e}),$$

which is possible from (1.31). Then it follows that

$$\begin{aligned}
V_{DM}^{e'}(\varnothing) &= \sum_{j\neq N_G(e')-1}^{N} \alpha'(G,\varnothing,j)\, f(j)\,(\mathbb{E}[\theta|j]-\eta) \\
&\quad +\alpha'(G,\varnothing,j)\, f(N_G(e')-1)\,(\mathbb{E}[\theta|N_G(e')-1]-\eta) \\
&\quad -\eta_S[\sum_{j\neq N_G(e')-1} \{1-\alpha'(G,\varnothing,j)\} f(j) \\
&\quad +\{1-\alpha'(G,\varnothing,N_G(e')-1)\} f(N_G(e')-1)] \\
&> \sum_{j\neq N_G(e')-1}^{N} \alpha(G,\varnothing,j)\, f(j)\,(\mathbb{E}[\theta|j]-\eta) \\
&\quad +\alpha(G,\varnothing,j)\, f(N_G(e')-1)\,(\mathbb{E}[\theta|N_G(e')-1]-\eta) \\
&\quad -\eta_S[\sum_{j\neq N_G(e')-1}^{N} \{1-\alpha(G,\varnothing,j)\} f(j) \\
&\quad +\{1-\alpha(G,\varnothing,N_G(e')-1)\} f(N_G(e')-1)] \\
&= V_{DM}^{e}(\varnothing),
\end{aligned}$$

which contradicts e being the best equilibrium for the DM. *Q.E.D.*

Proof of Theorem 5: That the strategies constructed by the procedure is an equilibrium follows from Theorem 4. Let λ be the largest j such that $c_j < 1$ in the above process. Then we have $\Gamma(j) < \eta$ for all $j > \lambda$. Suppose that $\widehat{e} = (\widehat{\alpha}, \widehat{\beta}, \widehat{\varphi}, \widehat{B})$ is the best equilibrium. From Lemma 9 and Theorem 4, it must hold that $\Pi_{s=1}^{\lambda}\widehat{\alpha}(G, G^{s-1}, \lambda) = c_\lambda$. Suppose that $\Pi_{s=1}^{\lambda-1}\widehat{\alpha}(G, G^{s-1}, \lambda-1) \neq c_{\lambda-1}$. because if $\Pi_{s=1}^{\lambda-1}\widehat{\alpha}(G, G^{s-1}, \lambda-1) > c_{\lambda-1}$ the (1.8) cannot be satisfied at period $\gamma - 2$, it must hold that $\Pi_{s=1}^{\lambda-1}\widehat{\alpha}(G, G^{s-1}, \lambda-1) < c_{\lambda-1}$. From the choice of $c_{\lambda-1}$, we have to have $1-\widehat{\alpha}\left(G,G^{\lambda-2},\lambda\right)>0$ and, hence,

$$
\begin{aligned}
&(1.32) - \{1-\widehat{\alpha}\left(G,G^{\lambda-2},\lambda-1\right)\}\Pi_{s=1}^{\lambda-2}\widehat{\alpha}\left(G,G^{s-1},\lambda-1\right)f(\lambda-1)\left(\mathbb{E}[\theta|\lambda-1]+\eta_S\right) \\
&\quad -\{1-\widehat{\alpha}\left(G,G^{\lambda-2},\lambda\right)\}\Pi_{s=1}^{\lambda-2}\widehat{\alpha}\left(G,G^{s-1},\lambda\right)f(\lambda)\left(\mathbb{E}[\theta|\lambda]+\eta_S\right) \\
&= \eta\sum_{j\geq\lambda-1}^{N}\Pi_{s=1}^{\lambda-1}\widehat{\alpha}\left(G,G^{s-1},j\right)f(j) \\
&= -c_{\lambda-1}f(\lambda-1)\left(\mathbb{E}[\theta|\lambda-1]+\eta_S\right)-c_{\lambda}f(\lambda)\left(\mathbb{E}[\theta|\lambda]+\eta_S\right).
\end{aligned}
$$

Because $|\mathbb{E}[\theta|\lambda-1]|>|\mathbb{E}[\theta|\lambda]|$, (1.32) implies that

$$
\begin{aligned}
&\eta\sum_{j\geq\lambda-1}^{N}\Pi_{s=1}^{\lambda-1}\widehat{\alpha}\left(G,G^{s-1},j\right)f(j) \\
&= \eta\{\Pi_{s=1}^{\lambda-1}\widehat{\alpha}\left(G,G^{s-1},\lambda-1\right)f(\lambda-1)+\Pi_{s=1}^{\lambda-1}\widehat{\alpha}\left(G,G^{s-1},\lambda\right)f(\lambda)+\sum_{j\geq\lambda+1}^{N}f(j)\} \\
&> \eta\{c_{\lambda-1}f(\lambda-1)+c_{\lambda}f(\lambda)+\sum_{j\geq\lambda+1}^{N}f(j)\}.
\end{aligned}
$$

This implies in the equilibrium condition (1.8) at period $\lambda-3$, the right hand side is strictly higher in equilibrium $\widehat{e}$ than in the equilibrium generated by the procedure. However, then it is possible to construct another equilibrium $e'=(\alpha',\beta',\varphi',B')$ by letting $\Pi_{s=1}^{\lambda-1}\alpha'(G, G^{s-1}, \lambda-1) = c_{\lambda-1}$, $\Pi_{s=1}^{\lambda}\alpha'(G, G^{s-1}, \lambda) = c_{\lambda}$, and

$$\begin{aligned} & \eta \sum_{j\geq\lambda-1}^{N} \Pi_{s=1}^{t}\alpha'\left(G, G^{s-1}, j\right) f(j) \\ < \quad & \eta \sum_{j\geq\lambda-1}^{N} \Pi_{s=1}^{t}\widehat{\alpha}\left(G, G^{s-1}, j\right) f(j) \text{ for all } t \in \{2, .., \lambda-2\}. \end{aligned}$$

Obviously, we have

$$\begin{aligned} V_{DM}^{e'}(\varnothing) &= \sum_{j}^{N} \Pi_{s=1}^{t}\alpha'(G, \varnothing, j) f(j) (\mathbb{E}[\theta|j] - \eta) \\ &\quad -\eta_S \sum_{j}^{N} \Pi_{s=1}^{t}\{1 - \alpha'(G, \varnothing, j)\} f(j) \end{aligned}$$

$$\begin{aligned} &> \sum_{j}^{N} \Pi_{s=1}^{t}\widehat{\alpha}(G, \varnothing, j) f(j) (\mathbb{E}[\theta|j] - \eta) - \eta_S \sum_{j}^{N} \Pi_{s=1}^{t}\{1 - \widehat{\alpha}(G, \varnothing, j)\} f(j) \\ &= V_{DM}^{\widehat{e}}(\varnothing), \end{aligned}$$

which contradicts $\widehat{e}$ being the best equilibrium. Then, $\Pi_{s=1}^{\lambda-1}\widehat{\alpha}(G, G^{s-1}, \lambda-1) = c_{\lambda-1}$ follows. Following the same procedure, we eventually get $\Pi_{s=1}^{j}\widehat{\alpha}(G, G^{s-1}, j) = c_j$ for all j, which shows that our procedure generates the best equilibrium. *Q.E.D.*

Proof of Theorem 8: The first statement is straightforward. In the benchmark strategy equilibrium e, we have $V_S(\varnothing, j) = 0$ for $j < N_G(e)$ and

$$\begin{aligned}
V_S(\varnothing, j) &= \sum_{s \in \{1,..,N_G(e)-1\}} (V - s\delta)(1-\delta/V)^{s-1}\delta/V \\
&\quad + (V - N_G(e)\delta)(1-\delta/V)^{N_G(e)-1} \\
&= \sum_{s \in \{1,..,N_G(e)-1\}} (V - s\delta)(1-\delta/V)^{s-1}\delta/V - (1-\delta/V)^{N_G(e)-1} N_G(e) \\
&\quad + (1-\delta/V)^{N_G(e)-1} N_G(e) + (V - N_G(e)\delta)(1-\delta/V)^{N_G(e)-1} \\
&= V_S(\varnothing, N_G(e)-1) + (1-\delta/V)^{N_G(e)-1} N_G(e) \\
&\quad + (V - N_G(e)\delta)(1-\delta/V)^{N_G(e)-1} \\
&= (1-\delta/V)^{N_G(e)-1} N_G(e) + (V - N_G(e)\delta)(1-\delta/V)^{N_G(e)-1},
\end{aligned}$$

for $j \geq N_G(e)$, where we used the fact $V_S(\varnothing, N_G(e)-1) = 0$. Now obviously $\frac{\partial V_S(\varnothing,j)}{\partial\delta} < 0$, which implies the result. *Q.E.D.*

Proof of $\sum_{j=1}^{N_G(e)-1} \frac{\partial\alpha(G,\varnothing,j)}{\partial\eta} f(j)(\mathbb{E}[\theta|1] - \eta) \leq 0$.

In order to get a contradiction, suppose that $\sum_{j=1}^{N_G(e)-1} \frac{\partial\alpha(G,\varnothing,j)}{\partial\eta}_{|\eta=\eta'} f(j)(\mathbb{E}[\theta|1] - \eta') > 0$ for some η'. Then, we have $\eta'' > \eta'$ for η'' sufficiently close to η' and $\sum_{j=1}^{N_G(e)-1} \alpha'(G,\varnothing,j) f(j)(\mathbb{E}[\theta|1] - \eta') > \sum_{j=1}^{N_G(e)-1} \alpha''(G,\varnothing,j) f(j)(\mathbb{E}[\theta|1] - \eta')$, where α' and α'' correspond the sender' strategy in the best equilibrium for the DM when $\eta = \eta'$ and $\eta = \eta''$, respectively.

From the fact that α' is supported as an equilibrium when $\eta = \eta'$ implies that (1.8) holds for all $t \leq N_G(e')$. Since $\eta' > \eta''$, this implies that

$$\widehat{\alpha}\left(G^{N_G(e')-1}, N_G(e') - 1\right) f\left(N_G(e') - 1\right) \mathbb{E}[\theta | N_G(e') - 1] = \eta'' \sum_{n \geq N_G(e')}^{N} f(n)$$

for some $\widehat{\alpha}\left(G^{N_G(e')-1}, N_G(e') - 1\right) < \alpha'\left(G^{N_G(e')-1}, N_G(e') - 1\right)$, where we used the notation $\alpha(m^\tau, j) = \Pi_{s=1}^{\tau} \alpha(m_s, m^{s-1}, j)$. The fact that (1.8) holds for $t = N_G(e') - 1$ implies that

$$\begin{aligned} &\sum_{j \geq N_G(e')-2}^{N_G(e')-1} \{1 - \widehat{\alpha}\left(G^{N_G(e')-1}, j\right)\} \widehat{\alpha}\left(G^{N_G(e')-2}, j\right) f(j)\, \mathbb{E}[\theta | j] \\ = \;& \eta''[\widehat{\alpha}\left(G^{N_G(e')-1}, N_G(e') - 1\right) f\left(N_G(e') - 1\right) + \sum_{n \geq N_G(e')}^{N} f(n)], \end{aligned}$$

for some

$$\widehat{\alpha}\left(G^{N_G(e')-2}, N_G(e') - 2\right) \leq \alpha'\left(G^{N_G(e')-2}, N_G(e') - 2\right)$$

and

$$\widehat{\alpha}\left(G^{N_G(e')-2}, N_G(e') - 1\right) \leq \alpha'\left(G^{N_G(e')-2}, N_G(e') - 1\right),$$

with at least one with strictly inequality. By continuing this, we will eventually get $\widehat{\alpha}(G, j) \leq \alpha'(G, j)$ for all j with at least one strict inequality. This implies that a sender's strategy $\widehat{\alpha}$ can be supported as an equilibrium when $\eta = \eta''$. However, it contradicts the fact that α'' is the best equilibrium, since

$$\begin{aligned}
&\sum_{j=1}^{N_G(e'')-1} \widehat{\alpha}\left(G,\varnothing,j\right) f\left(j\right)\left(\mathbb{E}[\theta|1]-\eta''\right) - \sum_{j=N_G(e'')}^{N} f\left(j\right)\left(\mathbb{E}[\theta|1]-\eta''\right)\\
>&\sum_{j=1}^{N_G(e'')-1} \alpha'\left(G,\varnothing,j\right) f\left(j\right)\left(\mathbb{E}[\theta|1]-\eta''\right) - \sum_{j=N_G(e'')}^{N} f\left(j\right)\left(\mathbb{E}[\theta|1]-\eta''\right)\\
>&\sum_{j=1}^{N_G(e'')-1} \alpha''\left(G,\varnothing,j\right) f\left(j\right)\left(\mathbb{E}[\theta|1]-\eta''\right) - \sum_{j=N_G(e'')}^{N} f\left(j\right)\left(\mathbb{E}[\theta|1]-\eta''\right).
\end{aligned}$$

Q.E.D.

Proof of Theorem 9: A probabilistic commitment is characterized by a $\xi - 1$ dimensional vector $\sigma = (\sigma_1, \sigma_2, .., \sigma_{\xi-1})$, where σ_j is the probability that the DM accepts the proposal after requiring j pieces of good evidence. We will prove that for any σ the the probabilistic commitment given in the theorem attains higher expected payoff for the DM. Towards this end, pick a commitment σ and fix it. Also, denote by $\pi(\sigma)$ be the DM's expected payoff associated with commitment σ, and $k(\sigma)$ be the threshold type of sender above which he is eventually accepted by the DM. It is without loss of generality to assume the followings:

$$(1.33)\qquad (1-\sigma_l)\sum_{j\geq l+1}^{N} f\left(j\right)\eta < \sum_{j\geq l}^{k(\sigma)-1} f\left(j\right)\mathbb{E}[\theta|j] \text{ for all } l \leq k\left(\sigma\right)-1,$$

because otherwise, another commitment $\sigma' = (\sigma_1, \sigma_2, ..\sigma_{l-1}, 1, 1, 1)$ attains higher expected payoff for the DM.

First suppose that $\sigma_1 > \delta/V$. Then, every sender $j \geq 1$ communicates a piece of good evidence at period 1. Then we have $\pi(\sigma) =$

$$\sigma_1\{\sum_{j\geq 1}^{N} f(j)(\mathbb{E}[\theta|j]-\eta) - \eta_S f(0)\} + (1-\sigma_1)\sigma_2\Psi_2 + (1-\sigma_1)(1-\sigma_2)\sigma_3\Psi_3$$

$$+\cdot+(1-\sigma_1)(1-\sigma_2)\cdot\left(1-\sigma_{k(\sigma)-1}\right)\{\sum_{j\geq k(\sigma)}^{N} f(j)(\mathbb{E}[\theta|j]-(k(\sigma)-1)\eta)\},$$

where Ψ_j is the expected payoff for the DM when she accepts at period j.

Think of the commitment $\sigma' = (\eta/V, \sigma_2, .., \sigma_\xi)$. We have $\pi(\sigma) \geq$

$$\delta/V\{\sum_{j\geq k(\sigma)}^{N} f(j)(\mathbb{E}[\theta|j]-\eta) - \eta_S f(0)\} + (1-\delta/V)\sigma_2\Psi_2$$

$$+\cdot+(1-\delta/V)(1-\sigma_2)\cdot\left(1-\sigma_{k(\sigma)-1}\right)\{\sum_{j\geq k(\sigma)}^{N} f(j)(\mathbb{E}[\theta|j]-(k(\sigma)-1)\eta)$$

$$-\eta_S\sum_{j<k(\sigma)}^{N} f(j)\},$$

which is strictly higher than $\pi(\sigma)$, because of (1.33). This implies that for all commitment σ such that $\sigma_1 > \delta/V$, there is a commitment σ' such that $\sigma_1' = \delta/V$ and attains higher expected payoff for the DM. Applying the same reasoning inductively, we can prove that for all commitment σ such that $\sigma_j > \delta/V$ for some j, there is a commitment σ' such that $\sigma_j' = \delta/V$ for all j and attains higher expected payoff for the DM.

Next, suppose that $\sigma_{k(\sigma)-1} < \delta/V$. We have

$$\begin{aligned}\pi(\sigma) &= \sigma_1\Psi_1 + (1-\sigma_1)\sigma_2\Psi_2 + \cdot + \\ &\quad (1-\sigma_1)\cdot\cdot\left(1-\sigma_{k(\sigma)-2}\right)\sigma_{k(\sigma)-1}\{\sum_{j\geq k(\sigma)}^{N} f(j)\left(\mathbb{E}[\theta|j] - (k(\sigma)-1)\eta\right)\} \\ &\quad (1-\sigma_1)\cdot\cdot\left(1-\sigma_{k(\sigma)-1}\right)\{\sum_{j\geq k(\sigma)}^{N} f(j)\left(\mathbb{E}[\theta|j] - k(\sigma)\eta\right)\}.\end{aligned}$$

Think of the commitment $\sigma' = (\sigma_1, ., \sigma_{k(\sigma)-2}, \delta/V, 1, ., \sigma_\xi)$. Now we have

$$\begin{aligned}\pi(\sigma') &= \sigma_1\Psi_1 + (1-\sigma_1)\sigma_2\Psi_2 + \cdot + \\ &= (1-\sigma_1)\cdot\cdot\left(1-\sigma_{k(\sigma)-2}\right)\frac{\delta}{V}\{\sum_{j\geq k(\sigma)}^{N} f(j)\left(\mathbb{E}[\theta|j] - (k(\sigma)-1)\eta\right)\} \\ &= (1-\sigma_1)(1-\sigma_2)\cdot\left(1-\frac{\delta}{V}\right)\{\sum_{j\geq k(\sigma)}^{N} f(j)\left(\mathbb{E}[\theta|j] - k(\sigma)\eta\right)\},\end{aligned}$$

which is strictly higher than $\pi(\sigma)$. Applying the same reasoning inductively backward, we can prove that for all commitment σ such that $\sigma_j < \delta/V$ for some j, there is a commitment σ' such that $\sigma'_j = \delta/V$ for all j and attains higher expected payoff for the DM.

Proof of Theorem 10: We denote the solution to the commitment problem by $r(\eta, \eta_S, \delta)$, and let $\kappa(\eta, \eta_S, \delta)$ be the length of persuasion of the best equilibrium for DM, i.e.,

$$\beta\left(A, G^{\kappa}\right)=1 \text{ and } \beta\left(A, G^{t}\right)=\delta / V \text{ for all } t \leq \kappa-1.$$

Since the result is trivially true when $\kappa\left(\eta, \eta_{S}, \delta\right)=0$, think of the case in which $\kappa\left(\eta, \eta_{S}, \delta\right) \geq 1$, i.e.,

$$V_{DM}^{*}(\varnothing)=\sum_{n=1}^{N} f(n)\, \alpha(G, n, \varnothing)(\mathbb{E}[\theta \mid n]-\eta)-\eta_{S} \sum_{n=1}^{N} f(n)(1-\alpha(G, n, \varnothing)).$$

Let $r=r\left(\eta, \eta_{S}, \delta\right)$ and $\kappa=\kappa\left(\eta, \eta_{S}, \delta\right)$.

Because $\alpha(G, \varnothing, j)=1$ for all $j \geq \kappa$, we have

$$\begin{aligned}
& \Upsilon_{DM}(\kappa)-V_{DM}^{*}(\varnothing) \\
= & \sum_{j \geq \kappa} f(j)(\mathbb{E}[\theta \mid j]-\kappa \eta)-\eta_{S} \sum_{j<\kappa} f(j) \\
& -\sum_{j=1}^{N} f(j)\, \alpha(G, \varnothing, j)(\mathbb{E}[\theta \mid j]-\eta)+\eta_{S} \sum_{j=1}^{N} f(j)\{1-\alpha(G, \varnothing, j)\} \\
= & -\sum_{j \geq \kappa} f(j)\, \kappa \eta+\sum_{j=1}^{N} f(j)\, \alpha(G, \varnothing, j)\, \eta \\
& -\sum_{j \geq 1}^{\kappa-1} f(n)\, \alpha(G, \varnothing, j)\, \mathbb{E}[\theta \mid n]-\eta_{S} \sum_{j=1}^{k-1} f(j)\, \alpha(G, \varnothing, j),
\end{aligned}$$

where we used $\alpha(G,\varnothing,j)=1$ for all $j\geq\kappa$. Then it follows that

$$\begin{aligned}
& \Upsilon_{DM}(\kappa)-V_{DM}(\varnothing) \\
= \; & -\kappa\eta\sum_{j\geq\kappa} f(j)+\eta\sum_{j=1}^{N} f(j)\,\alpha(G,j,\varnothing)-\eta_S\sum_{j=1}^{k-1} f(j)\,\alpha(G,\varnothing,j). \\
& -\sum_{j\geq 1}^{\kappa-1} f(n)\,\alpha(S,j,G)\,\alpha(G,j,\varnothing)\,\mathbb{E}[\theta|j]-\sum_{j\geq 1}^{\kappa-1} f(n)\,\alpha(G,j,G)\,\alpha(G,j,\varnothing)\,\mathbb{E}[\theta|j] \\
= \; & -\eta\kappa\sum_{j\geq\kappa} f(j)+\eta\sum_{j=1}^{N} f(j)\,\alpha(G,\varnothing,j)-\eta_S\sum_{j=1}^{k-1} f(j)\,\alpha(G,\varnothing,j) \\
& +\eta\sum_{j\geq 1}^{N} f(j)\,\alpha(G,G,j)\,\alpha(G,\varnothing,j)+\eta_S\sum_{j\geq 1}^{\kappa-1} f(j)\,\alpha(S,j,G)\,\alpha(G,j,\varnothing) \\
& -\sum_{j\geq 1}^{\kappa-1} f(j)\,\alpha(G,G,j)\,\alpha(G,\varnothing,j)\,\mathbb{E}[\theta|j] \\
= \; & \cdot\cdot=-\eta\kappa\sum_{j\geq\kappa} f(j)+\eta\sum_{j\geq 1}^{N}\sum_{t\geq 1}^{\kappa}\Pi_{s=0}^{t-1}\alpha(G,G^s,j)\,f(j) \\
= \; & \eta\sum_{j\geq 1}^{\kappa-1}\sum_{t\geq 1}^{\kappa}\Pi_{s=0}^{t-1}\alpha(G,G^s,j)\,f(j)\geq 0,
\end{aligned}$$

where the last inequality is strict when $k\geq 1$. Note that we used the conditions of benchmark equilibrium repeatedly, i.e.,

$$\begin{aligned}
& -\sum_{j\geq 1}\alpha\left(S,G^{n-1},j\right)\Pi_{j=0}^{n-1}\alpha\left(G,G^j,j\right)\left(\mathbb{E}[\theta|j]+\eta_S\right) \\
= \; & \eta\sum_{j\geq 1}^{\kappa-1}\alpha\left(G,G^{n-1},j\right)\Pi_{j=0}^{n-1}\alpha\left(G,G^j,j\right),
\end{aligned}$$

for all $n < \kappa$. *Q.E.D.*

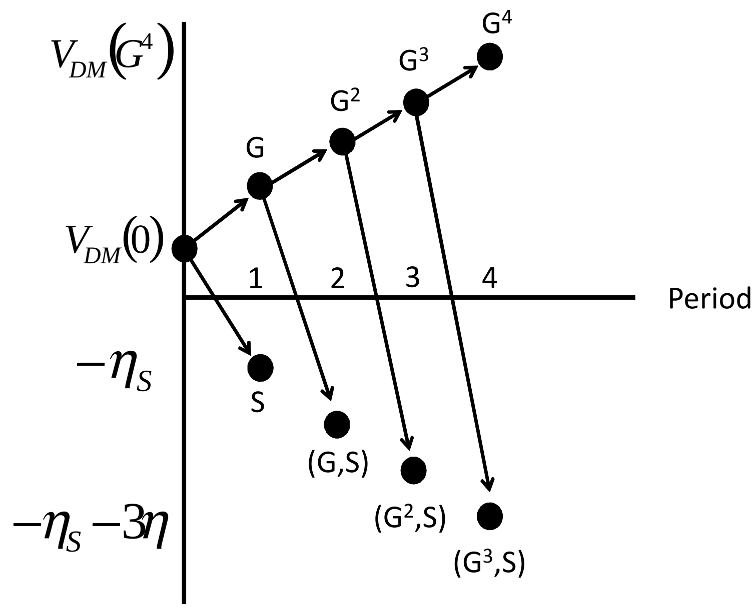

Figure 1.1. DM's value on a benchmarket strategy equilibrium

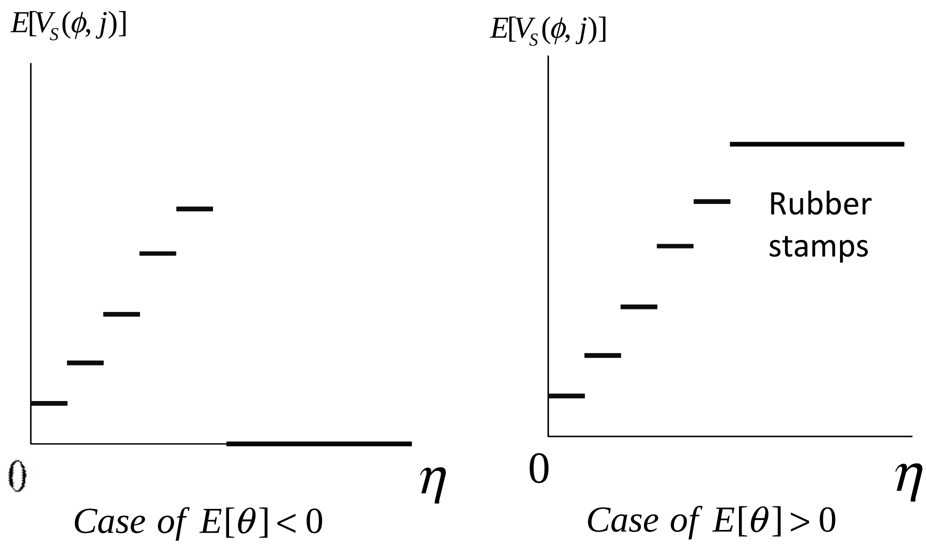

Figure 1.2. Comparative statics on the sender's expected payoff

CHAPTER 2

Signaling Competence in Elections

2.1. Introduction

At a Republican presidential debate held in Iowa in summer 2011, amid the economic turmoil, after the debate over raising the debt ceiling is settled, all candidates rejected the idea of a deficit-reduction plan that included one dollar of tax increase for every \$10 of spending cuts. Apparently, they have been building the reputation of being extreme, putting more weight on their traditional small government doctrine over the textbook prescription of fiscal expansion in a downturn. In the previous 2008 US presidential campaign, the presumptive Republican nominee, Senator John McCain picked Governor Sarah Palin, whose political position was far more conservative than his own, to be his running mate and, thereby shifted public perception about his political position. Although those behaviors completely contradict the celebrated median voter theorem, the casual observation that being extreme increases public appeal seems to be a very general phenomena. This chapter tries to explain this casual observation.

Towards this end, this chapter incorporates a dimension of "competence" of politicians, which is their private information. Our perception of competence follows the point made by Stokes (1963). According to Stokes, the role of politicians includes identifying the electorate's concerns and trying to convey the message that their policy proposals effectively address those concerns. This perception defines a dimension of politicians' competence in our model: the ability to discover the most effective policies, where effectiveness depends on the circumstances. In an election, the electorate tries to choose a competent person as their leader, under the presumption that only a portion of political candidates possess competence. This makes it necessary for electoral candidates to behave in a way that projects the appearance of competence.

Our results show that proposing an extreme platform serves as a signal of competence and gives candidates a strictly higher winning probability in comparison to a median platform. This result stems from the fact that an extreme platform, which can be very bad depending on the circumstance, is a gamble for incompetent candidates.[1] This gives the extreme platform a signaling effect of competence. In an equilibrium, extreme platforms have to have strictly higher winning probabilities than the median platform in order to make incompetent candidates bet on the gamble and, thereby, prevent perfect separation of the incompetent from the competent.

We model electoral competition by adding a state space to a standard Hotelling-Downs one-dimensional policy location game. The bliss policy of the median voter depends on the realized state of the world and is probabilistic in nature. In our

[1]This effect is first found by Majumdar and Mukand (2004), although they are not dealing with an electoral competition game.

model the probability distribution over states is such that the best policy, from an ex-ante point-of-view, is the median policy. Competent candidates can observe the state of the world at the time of the election and, hence, they are aware of the ex-post best policy, while incompetent candidates are not. This means that our model is a signaling game, in which competent candidates' strategies can be state-dependent.

The equilibrium behavior of incompetent candidates depends on the extent to which electoral candidates are policy-motivated and how uncertain the states are. When candidates are sufficiently policy-motivated, the motive to signal competence is dominated by the policy motive, and they choose the median platform, that is, the ex-ante best platform. However, when they are strongly motivated to win the office and the states are uncertain, they may choose an extreme platform to pretend to be competent and, thereby, increase their winning probability. In contrast, competent candidates always choose an extreme platform in an extreme state and how they behave in the median state depends on how uncertain the states are.

We also provide some comparative static analysis. Perhaps surprisingly, we show that an increase in the degree of office motivation does not affect the voter's expected payoff, once it exceeds a certain level. This follows because a high motivation to obtain office increases the probability that the median platform will win and, thereby, mitigates the candidates' signaling competence motivation. This makes candidates' equilibrium strategy invariant with respect to their degree of office motivation, which makes the voter's expected payoff invariant also. However, when their office motivation is low enough, below a threshold value, then we may have an equilibrium such that all the candidates just propose the best policy for the voter according to their information, which attains the highest level of payoff among all the equilibria.

Because our basic model setting is based on that of Kartik and McAfee (2007), it is important to discuss the differences between our study and theirs in some detail. In their model, a fraction of candidates have "character", which is unobserved by voters like "competence" in our model. Our model differs from their model in not treating competence as an attribute that voters intrinsically prefer. While in their model whether a candidate has a character or not (indicator function) enters directly in the voter's payoff function, in our model the voter cares only about what policy is implemented, contingent on the state of the world. Hence, the fact that the voter prefers competent candidates is an equilibrium phenomenon and, hence, is a result of the model.

One more important difference is that in our model each platform has a different winning probability. In Kartik and McAfee (2007), as in most models of electoral competition, each platform (on-equilibrium) has the same winning probability: otherwise, candidates concentrate on the platform that has the highest winning probability particularly, when they care only about winning the election, which is the setting most often used. Our study is novel in constructing a model that supports different winning probabilities for different platforms, which is accomplished by using the hybrid of office motivation and policy motivation in candidates' payoffs.

Finally, the technology of equilibrium construction is very different. In their model, the voter's payoff from choosing a candidate is independent from the choice made by another candidate and, hence, the voter simply compares the two payoffs and votes for the preferred one. In our model, the voter's (expected) payoff from voting for a candidate depends on both candidates' choices because types of them are correlated, which gives some information about the payoff relevant state of the world. In this sense, two candidates are more strategically interacted with each other, which requires different technology for solving the model. In constructing a signaling game of electoral competition, this chapter is also related to Banks (1990), Callander (2008) and Callander and Wilkie (2005).

This study is different from the most traditional electoral competition model (Downs (1957) and Davis, Hinich, and Ordeshook (1970)) in which the candidates are concerned solely with winning the election, where they are endowed with complete knowledge about what the election results will be, given any particular choice of platforms by the candidates. In this study, we assume that candidates are also policy motivated in the usual sense the term is used (Calvert (1985) and Wittman (1977)), which is crucial for deriving the main insights.

In developing a model with uncertainty about the policy-relevant state of the world and where politicians receive private signals about the true state, this study is related to a number of other studies. Majumdar and Mukand (2004) develop a dynamic model of policy choice in which politicians, who are either high ability or low ability, care about both public welfare and future electoral prospects. They show the possibility of inefficient persistence of a previously enacted policy, since changing policies signals low ability of candidates.

Schultz (1996) analyzes an election model in which ideologically biased candidates are informed about the policy-relevant state. He finds that the degree to which candidates reveal their private information depends on the degree their policy preferences are biased. Martinelli (2001) analyzes the case in which voters also have private information about the policy-relevant state and shows that equilibrium does not result in policy convergence. Heidhues and Lagerlof (2003) show that pandering happens, that is, the candidates have a strong incentive to bias their platform choice toward the electorate's prior belief.[2] On the contrary, in a recent paper, Kartik, Squintani, and Tinn (2012) show that candidates have an incentive to exaggerate their private information, i.e., to anti-pander. This study also shows the possibility of anti-pandering, which comes from their motive to signal their competence.

In focusing on the vertical difference between candidates, rather than on the horizontal difference (policy preference), this chapter is also related to Aragones and Palfrey (2001). They study an electoral model in which one candidate enjoys an advantage in the sense that when his opponent candidate chooses the same

[2]See also Jensen (2009), Laslier and Straeten (2004), and Loertscher (2012), for more studies on how candidates convey their policy relevant private information through their platform choices.

platform, the voter votes for him. Our perception of competence differs from theirs because the voter does not necessarily prefer a competent over an incompetent candidate as long as they choose the same policy; rather, the preference over competence is generated endogenously by the fact that competent candidates tend to choose appropriate policies. In Caselli and Morelli (2004), competent candidates have a high ability to implement policy, but they tend to have better outside options in comparison to being politicians. Their focus is on the relation between the demand and supply of electoral candidates, while our focus is on the outcomes of elections.

This chapter is organized as follows. Section 2.2 introduces the basic structure of the model. In Section 2.3, we characterize some important equilibria of our game. We also give some discussion about refinement issue. Section 2.4 presents some concluding remarks. Formal proofs are given in Appendix.

2.2. The Model

The basic element of the model is a standard Hotelling-Downs unidimensional policy location game augmented with an uncertain state of the world. There is a one-dimensional policy space $X = \{-1, 0, 1\}$, and a set of states of the world, $\Theta = \{-1, 0, 1\}$[3]. There is a probability mass function over states $f : \Theta \to [0, 1]$, that satisfies $f(0) = m \in (0, 1)$ and $f(-1) = f(1) = (1 - m)/2$, where m is seen to represent the degree of uncertainty. Typical elements of X and Θ are denoted by

[3] We will discuss the choice of this discrete model setting in the concluding section.

$x \in X$ and $\theta \in \Theta$, respectively. There is one voter who has a policy preference defined on the product of X and Θ, or we can interpret it as a median voter which has a strictly positive mass.[4] Specifically, we assume that the voter's utility $u(x, \theta)$ takes the quadratic loss form, $u(x, \theta) = -(x - \theta)^2$, which implies that the voter wants the implemented policy and the state of the world to be as close as possible.[5] It follows immediately that policy 0 maximizes the voter's expected utility $\mathbb{E}_\theta[u(x, \theta)]$, where the expectation is taken with respect to random variable θ. Hereafter, we call platform 0 the median, and 1 and -1 extreme.

We now introduce *competence* of candidates as follows. There are two candidates, A and B. Competence of candidates is a binary variable: candidate $i \in \{A, B\}$ either possesses competence ($c^i = C$) or does not ($c^i = I$). This is private information and it is drawn independently from a Bernoulli distribution with $\Pr(c^i = C) = c > 0$. Before choosing his platform, a competent candidate receives a perfect signal about the state of the world, θ, whereas an incompetent candidate does not receive any signal.[6]

[4] Alternatively, we can think that a voter is characterized by her preference parameter, b, and her preference over policy is $-(x - \theta - b)^2$. The median value of preference parameter is 0, which has a strictly positive mass.

[5] One way to understand this setting is to interpret x as a level of a government's fiscal spending, and θ as a state of its economy. The people's preference over the level of fiscal spending swings with the state of the economy.

[6] Majumdar and Mukand (2004) define a high ability politician and a low ability politician in a similar way.

Our election game proceeds as follows. First, Nature chooses the types of the two candidates, determining whether they are competent or not, and this becomes private information. Second, Nature chooses a state from Θ according to pmf f, and only competent candidates observe it. Then, the two candidates simultaneously choose their respective platforms. Since a competent candidate observes the state of the world, his platform choice can be state-dependent, while an incompetent candidate's platform choice is not state-dependent. After observing the two candidates' platforms, the voter votes sincerely to maximize her expected utility, without knowing the true state of the world or receiving any signal by herself. After the election, the two candidates receive a payoff from policy of $u(x, \theta)$, which is the same as the voter's utility. To focus on the vertical difference (competence) between the candidates, we assume that each candidate has the same policy preference. In addition to the payoff from policy, the winning candidate obtains the office rent, $k \in [0, \infty)$, which is the degree of motivation to win the office, where the case of $k = 0$ corresponds to the pure policy motivation case. The value of k is common knowledge and common across candidates, which is dealt with as a parameter of the model.[7]

Since incompetent candidates do not observe the state of the world, their strategies are state-independent. Allowing the possibility of mixed strategies, a strategy for incompetent candidate $i \in \{A, B\}$ is represented by a probability mass

[7]In a benchmark Hotelling-Downs models of electoral competition, candidates are purely office motivated, i.e., $k = 0$.

function $g^i : X \to [0,1]$.[8] In contrast, since competent candidates observe the state, their strategies can be state-dependent. A strategy for competent candidate $i \in \{A, B\}$ when the state of the world is θ is represented by a probability mass function $g^i_\theta : X \to [0,1]$.

Since the voter does not observe the state, she has to decide which candidate to vote for based only on candidates' chosen platforms. Her voting strategy is described by a voting function $v : X \times X \to [0,1]$, where its value represents the probability of voting for Candidate A.

With this preparation, given Candidate B's strategy and the voter's voting strategy, Candidate A's expected payoff from choosing platform x^A when he is competent, observing the state of θ is written as

$$c \sum_{x^B \in X} \{v(x^A, x^B)(k + u\left(x^A, \theta\right)) + \left(1 - v(x^A, x^B)\right)u\left(x^B, \theta\right)\} g^B_\theta\left(x^B\right) \tag{2.1}$$

$$+ (1-c) \sum_{x^B \in X} \{v(x^A, x^B)(k + u\left(x^A, \theta\right)) + \left(1 - v(x^A, x^B)\right)u\left(x^B, \theta\right)\} g^B\left(x^B\right),$$

[8]As in Kartik and McAfee (2007), we take the interpretation of mixed strategies according to Bayesian view of opponents' conjectures, originating in Harsanyi (1973). That is, a candidate's mixed strategy need not represent him literally randomizing over platforms; instead, it represents the uncertainty that the other candidate and the electorate have about his pure strategy choice.

which we denote by $U^A\left(x^A,\theta\right)$. On the other hand, Candidate A's expected payoff when he is incompetent is

$$\mathbb{E}_\theta[U^A\left(x^A,\theta\right)] = \sum_{\theta\in\Theta} U^A\left(x^A,\theta\right) f\left(\theta\right). \tag{2.2}$$

Candidate B's expected payoff can be described in an analogous way.

As this is a signaling game, voter beliefs about the state of the world are critical. Let $\varphi\left(\cdot,x^A,x^B\right):\Theta\to[0,1]$ be the posterior probability mass over the states given Candidate A's and Candidate B's platforms x^A and x^B, respectively. Given the posterior belief over the states, φ, the voter votes for Candidate A if

$$\left|x^A - \sum_{\theta\in\Theta}\theta\varphi\left(\theta,x^A,x^B\right)\right| < \left|x^B - \sum_{\theta\in\Theta}\theta\varphi\left(\theta,x^A,x^B\right)\right|, \tag{2.3}$$

and votes for Candidate B if the opposite inequality holds.[9] Any voting rule is optimal when those two terms in (2.3) are equal.

Our solution concept is that of perfect Bayesian equilibrium (Fudenberg and Tirole 1991). This requires that the platform distributions, g^A, g^B, g_θ^A and g_θ^B, maximize the expected payoff for each candidate given voter belief φ, and that they are consistent with Bayes' rule. Therefore, competent candidates maximize (2.1), whereas, incompetent candidates maximize (2.2). To simplify the analysis, we focus on the anonymous equilibrium, where $g^A = g^B$, $g_\theta^A = g_\theta^B$ for all θ, $v\left(x,y\right) =$

[9] This follows from the quadratic loss form of the voter's utility function.

$1-v(y,x)$ for all $x,y\in X$.[10] Hence, two candidates choose the same strategy and the voting rule treats two candidates equally. This allows us to drop the superscript on candidate strategies.

Those imply that in an equilibrium, the following must hold

$$\varphi\left(\theta,x^A,x^B\right)=\frac{f(\theta)\,\Delta\left(x^A,x^B,\theta\right)}{\sum_\theta f(\theta)\,\Delta\left(x^A,x^B,\theta\right)}, \tag{2.4}$$

where

$$\begin{aligned}\Delta\left(x^A,x^B:\theta\right) &= c^2 g_\theta(x^A)g_\theta(x^B)\\ &\quad +c\,(1-c)\left(g_\theta(x^A)g(x^B)+g(x^A)g_\theta(x^B)\right)+(1-c)^2\,g(x^A)g(x^B),\end{aligned}$$

which represents the probability that a particular pair of platforms (x^A,x^B) is chosen, conditional on the realization of the state. We have no restrictions on $\varphi\left(\theta,x^A,x^B\right)$ when $\left(x^A,x^B\right)$ is never chosen in an equilibrium (no restrictions on off-equilibrium belief).

Finally, given equilibrium (g,g_θ,v,φ), which is a pair of candidates' strategies, the voter's voting strategy and the voter's updated belief over states, the (ex-ante) winning probability of platform x before observing the opponent's choice, which is

[10] Note that we are not assuming $g_\theta(-1)=g_\theta(1)$ nor $g(-1)=g(1)$.

denoted by $W(x)$, is

$$W(x) = c \cdot \mathbb{E}_{\theta}[\sum_{y \in X} g_{\theta}(y) v(x,y)] + (1-c) \sum_{y \in X} g(y) v(x,y).$$

This follows because with probability c, the opponent is competent, using state contingent strategy $g_{\theta}(y)$ while with probability $1-c$ the opponent is incompetent using non-state contingent strategy $g(y)$.

In the following, in order to state our results, let $T = \{I\} \cup C \times \Theta$ be the type space of candidates with an element $t = (C, \theta)$, where $I \in T$ corresponds to the incompetent type, and $(C, \theta) \in T$ corresponds to the type of competent candidate observing the state θ.

2.3. Signaling Competence

In this section, we examine strategies that constitute a perfect Bayesian equilibrium behavior. As in usual signaling games, our model has multiple equilibria where some of them are supported only by using unreasonable off-equilibrium beliefs. From this reason, except for Theorem 1, we put emphasis on the equilibria that have no off-equilibrium platform choice, that is, from ex-ante point of view each of the three platforms may be chosen with a strictly positive probability. The possibilities of other types of equilibria and their plausibility are examined in the end of the section, and we provide more formal argument when we discuss refinement issues in the Appendix.

Before characterizing the equilibria that have no off-equilibrium platform choice, we start by giving the equilibrium in which the median platform is never chosen, but its existence is ensured for all parameter specifications. In this sense, we have complete polarization in the equilibrium.

Theorem 11. *For all parameter values, there is a perfect Bayesian equilibrium such that*

1. *Competent candidates completely polarize, that is,*

$$g_{-1}(-1) = g_1(1) = 1, \text{ and } g_0(0) = 0.$$

2. *Incompetent candidates completely polarize, that is,*

$$g(-1) = g(1) = \frac{1}{2}.$$

3. *The voting strategy is such* $W(-1) > W(0)$ *and* $W(1) > W(0)$.

In this equilibrium, it does not need to be $g_0(-1) = g_0(1)$, because the expected value of the state for the voter after observing candidates' platform choices is invariant with those values, that is, $\sum_{\theta\in\Theta} \theta\varphi(\theta, -1, 0) = \sum_{\theta\in\Theta} \theta\varphi(\theta, 0, 1) = 1/2$ for all allocations of $g_0(-1)$ and $g_0(1)$. The equilibrium is supported by attaching a low belief about competence, to those who deviate and propose the median.

Although the existence of the equilibrium in Theorem 11 is ensured for all parameter specifications, we have a good reason to doubt the plausibility of it for some cases. The equilibrium is supported because the winning probability of the median platform is minimized by making the voter's off-equilibrium beliefs extreme, after observing one candidate chose the median and another chose an extreme, (we are free to attach any belief, from the definition of perfect Bayesian equilibrium). A simple intuition, however, tells us that when the median state is very likely, or candidates are very likely to be incompetent, the equilibrium hardly seems to be supported. For example, if the voter uses the simple updating rule such that after a candidate deviates to the median, she updates her belief based only on the platform choice made by the non-deviating candidate by using Bayesian rule, it can be shown that we need $c \geq m/(1-m)$ to make the voter want to vote for a non-deviating candidate.

In the following, we characterize different types of equilibrium in order of the degree of polarization. We start from an equilibrium in which no candidates polarize and proceed to ones with more and more polarization. Generally speaking, the degree of polarization depends on how uncertain the state is and how much political candidates are motivated to obtain the office. However, all the equilibria characterized in this section share the following common facts, whose proofs follow from Theorem 12 through 16:

Proposition 5. *In an equilibrium such that all the three platforms may be chosen, that is, in an equilibrium such that* $\max\{g(x), \sum_{\theta} g_{\theta}(x)\} > 0$ *for all* x*, the following facts are satisfied:*

Fact 1: *The (ex-ante) winning probability from proposing an extreme platform (-1 or 1) is strictly higher than that from proposing the median platform (0), that is,* $W(-1) > W(0)$ *and* $W(1) > W(0)$.

Fact 2: *The probability that a candidate who chooses an extreme platform is competent is higher than* c*, that is,* $\frac{c\sum_{\theta} g_{\theta}(-1)}{c\sum_{\theta} g_{\theta}(-1)+(1-c)g(-1)} > c$ *and* $\frac{c\sum_{\theta} g_{\theta}(1)}{c\sum_{\theta} g_{\theta}(1)+(1-c)g(1)} > c$.

Fact 3: *Competent candidates who observe an extreme state choose the same platform as policy as the state, that is* $g_{-1}(-1) = g_1(1) = 1$.

Fact 4: *Competent candidates who observe the median state choose the median platform (weakly) less often than incompetent candidates do, that is,* $g_0(0) \leq g(0)$.

There is a subtlety in Fact 2. The mathematical expressions given there say that conditional on observing a candidate choosing an extreme platform, but without using information that may be taken from another candidate's choice, the probability of the candidate being competent is higher than c. In our model, the voter observes both candidates' choices simultaneously and because candidates' choices are correlated, one candidate's choice should tell the voter something about another candidate's type. This implies that the voter does not update the candidate's type by using simple Bayesian rule given in the statement of Fact 2. We can also state Fact 2 in the alternative way, that is, for all combination of candidates' platform choices, the probability of a candidate choosing an extreme platform being competent is higher than c. Although Fact 2 is stated in the former form, both of the statements are correct in our model.

Fact 3 says that competent candidates observing an extreme state simply choose the (ex-post) best platform that matches the true state. Because such a platform actually maximizes the voter's utility given the state, we do not call such behavior of candidates "polarization". Here, it might be useful to explicitly define the usage of the term polarization in our study. We say that a candidate is polarizing when he is choosing a platform -1 or 1, when the median platform is most preferred according to his policy preference and available information. This says that the word is used only for the behavior of incompetent candidates and competent candidates observing the true state of the median. Furthermore, we say that a candidate is completely polarizing if he never chooses the median and mildly polarizing when he puts some positive probability (but less than one) on choosing the median, when the median is the most preferred according to his policy preference.

Fact 4 comes from the fact that candidates are risk averse. No that for a type $(C, 0)$ candidate, choosing an extreme platform is not very risky compared to an incompetent candidate because he already knew that the state is the median. This implies that when type $(C, 0)$ candidates are moderately polarizing or not polarizing, incompetent candidates are not polarizing either.

One important corollary of Proposition 5 is that the winning probability of competent candidates is strictly higher than that of incompetent candidates. It follows because the winning probability from proposing an extreme platform is strictly higher than that of the median platform and competent candidates are more likely to choose the extreme platform.

Corollary 4. *In an equilibrium such that every platform may be chosen, that is, in an equilibrium such that* $\max\{g(x), \sum_\theta g(x)\} > 0$ *for all* x*, competent candidates have a strictly higher winning probability than incompetent candidates.*

2.3.1. Low Motivation for Office

We start from the case in which candidates are sufficiently policy motivated (low k). The following theorem demonstrates that there is a perfect Bayesian equilibrium in which candidates always choose the best policy for them, that is, the same policy as the state for competent candidates and median policy for incompetent candidates (all the proofs can be found in the Appendix).

Theorem 12. *There is a perfect Bayesian equilibrium that satisfies the following if and only if* $k \leq 2$.

1. Competent candidates choose the best platform for each state, that is, $g_0(0) = g_{-1}(-1) = g_1(1) = 1$.

2. Incompetent candidates always choose median platform, that is, $g(0) = 1$.

3. The voter votes for an extreme candidate over a moderate one, that is, $v(-1,0)=1,\ v(1,0)=1.$

In the equilibrium, the voter always votes for a candidate who chooses an extreme platform over a moderate one because choosing an extreme platform is a perfect signal of competence, and thereby, conveys the perfect information about the state being the extreme. An incompetent candidate chooses the median platform, which is the ex-ante best platform, and wins the election only when his opponent is incompetent or when the median platform is actually the ex-post best platform, in which case the election results in a tie with the competent opponent. Although choosing an extreme platform gives incompetent candidates a strictly higher winning probability than choosing the median platform, they do not do so because such a policy is likely to turn out to be wrong. We can easily see that Facts 1 through 4 hold, and in particular, Fact 4 holds with $W(-1)=W(1)>W(0)$. Note that, it is not necessarily the case that we have $v(-1,1)=1/2$, as long as $W(-1)=W(1)$ is ensured (we can adjust by $v(-1,0)$ and $v(1,0)$, accordingly).

An important implication of Theorem 12 is the validity of the median voter theorem. In our model, the appropriate statement of the median voter theorem is that each type of candidate chooses the most desirable platform based on his knowledge. One interesting insight to note is that, contrary to the literature, the incompetent candidate's reason for choosing the median platform arises from policy motivation, whereas, in the usual Hotelling-Downs model of electoral competition, candidates' motivation for choosing the median platform arises from the motivation to win the office.

As the above reasoning clearly shows, the equilibrium in Theorem 12 is supported by candidates' sufficiently high policy motivation, that is, by small k. In the equilibrium, even though incompetent candidates know that choosing an extreme platform gives a higher winning probability, they stick to the median platform because of their policy motivation. It can easily be seen that the candidates' strategies supported in the equilibrium of Theorem 12 is the best for the voter among all possible strategies candidates may choose.

2.3.2. Low Uncertainty

In this subsection, we examine the equilibrium when there is low uncertainty. More precisely, we focus on the case in which m, the probability that the state is median, is higher than $1/2$. We show that in such a case, when a candidate's motivation to obtain the office is higher than some threshold, the equilibrium involves polarization from competent candidates and moderation from incompetent ones.

Theorem 13. *Let $m > 1/2$. There is a perfect Bayesian equilibrium that satisfies the following if and only if, we have $k \geq \kappa(c, m)$, where $\kappa(c, m)$ is a unique number determined by (c, m) and $\kappa(c, m) < 2 - 2c$ for all (c, m).*

1. Competent candidates mildly polarize, that is,

$$g_0(0) > 0, \;\; g_0(-1) = g_0(1) > 0, \;\; \textit{and } g_{-1}(-1) = g_1(1) = 1.$$

2. Incompetent candidates always choose the median platform, that is, $g(0) = 1$.

3. The voting strategy is such that $W(-1) > W(0)$ *and* $W(1) > W(0)$.

We say that competent candidates "mildly" polarize because they still put strictly positive probability on choosing the median platform. Note that, we do not exactly specify the voting strategy v, since there are multiple voting rules that attain the indifference for the type of sender, which we explain later.

In the equilibrium, competent candidates observing the median state (type $(C, 0)$ candidates) mix between all the policies. Incompetent candidates choose the moderate policy for certainly. The voter, who is indifferent between two candidates proposing an extreme and the median, respectively, mix strategies.[11] This requires that the way the voter mixes strategies makes type $(C, 0)$ candidates indifferent between all platforms, and that the way type $(C, 0)$ candidates mix strategies on policy announcements makes the voter indifferent between candidates.

[11]Kartik and McAfee (2007)'s model also shares this property. In their model, this property and the assumption that voters flip a fair coin when there is a tie induces ex post equilibrium, in which the same behavior of candidates remains an equilibrium, regardless of one candidate announces first, second, or both announce simultaneously. The equilibrium in our model, however, is not ex-post equilibrium, since the voter is more likely to vote for an extreme. This implies that an extreme platform is more preferred when the opponent chooses an extreme and vice versa.

To see this point, keeping the other types' strategies fixed as stated in the theorem, suppose that type $(C, 0)$ candidates' strategy puts a very small probability on choosing an extreme platform, say platform 1. Then, choosing the platform works as a very strong signal about type $(C, 1)$ or state being one and, hence, it makes the voter to vote for the extreme. On the other hand, if type $(C, 0)$ candidates choose an extreme platform with high probability, that platform loses signaling power and, thus, the voter prefers to vote for the median. In the adequate degree of mixing, the expected value of the state after observing a platform pair of (0,1) becomes exactly $1/2$. On the other hand, to make type $(C, 0)$ candidates indifferent between the median and extremes the winning probability of extreme platforms must be larger than that of the median platform because extreme platforms attain lower policy utility than the median platform for those types of candidates.

As is mentioned above, there are multiple voting strategies that support the equilibrium, and it does not need to be $W(-1) = W(1)$. The way that the voting strategy v supports the candidates' strategies is to make two relations hold, that is, $U(-1, 0) = U(0, 0)$ and $U(1, 0) = U(0, 0)$. However, we have three variables, $v(-1, 1)$, $v(-1, 0)$, and $v(0, 1)$ to attain this, which is the reason for the multiplicity. If we further impose the symmetry assumption of $v(-1, 1) = 1/2$, it is possible to get $W(-1) = W(1)$.

We have some comparative statics results and limiting behavior of the equilibrium when we take extreme values of parameters.[12] In order to do this, focus on the symmetric equilibrium such that $v(0,-1) = v(0,1) > 0$, and denote $g_0(0)$ and $v(0,-1)$ by $g(k,c)$ and $\beta(k,c)$, respectively, as functions of k and c.[13] It can be shown that they are differentiable and

$$\frac{\partial g(k,c)}{\partial c} > 0, \ \frac{\partial g(k,c)}{\partial k} = 0, \text{ and } \lim_{c \uparrow 1} g(k,c) = 0, \tag{2.5}$$

and

$$\frac{\partial \beta(k,c)}{\partial c} > 0, \ \frac{\partial \beta(k,c)}{\partial k} > 0, \text{ and } \lim_{k \uparrow \infty} \beta(k,c) = \frac{1}{2}. \tag{2.6}$$

The relations in (2.5) say that as we increase the high probability of competence, incompetent candidates' strategies skews toward the median. This is explained as follows. If we fix the candidates' strategies, in which type I candidates are moderate and type $(C,0)$ candidates mildly polarize, an increase in the probability of competence makes the voter think that an extreme platform choice is likely to be made by a type $(C,0)$ candidate. This shifts her belief about the state

[12]As will be shown in the Appendix, there are two different voting rules that support the candidates' strategies in Theorem 2.3 as an equilibrium. One involves $v(-1,0) = v(1,0) = 0$ and another involves at least one of those being strictly higher than zero. In the comparative statics, we focus on the latter because the former equilibrium exists only for a smaller range of parameters compared to the latter.

[13]The existence of the symmetric equilibrium and the uniqueness of those values, g and β, are proved in the Appendix.

toward the center when she observes the pair of platform choices of an extreme and the median. This effect must be offset by making type $(C, 0)$ candidates put less weight on the extreme so that the voter's expected state being $1/2$ $(-1/2)$ when she observes the platform pair $(0, 1)$ ((-1,0)) and, thereby, make her indifferent to the choice of median. It implies that a higher probability of competence (higher c) has positive effects on the voter's welfare.

On the other hand, (2.6) shows that a high motivation to obtain office increases the probability that the median platform will win, thereby, preventing candidates from polarizing more. It also shows that a high probability of competence increases the probability that the median will win. To see this, note that a high probability of competence implies that the opponent candidate is more likely to choose an extreme. This means that the winning probability of the median decreases, which must be offset to support the equilibrium.

The comparative statics analysis gives the relation between those parameters and voter's expected payoff, which is denoted by $\mathbb{E}(u)$ as a corollary. Because, on the equilibrium of the theorem, $g_0(0)$ is constant with respect to k, the voter's expected payoff is also constant with respect to k. On the other hand, the voter's payoff is increasing with c, mainly from the direct effect of competence, and partly from the indirect effect of making type $(C, 0)$ candidates put more weight on the median. See Figure 2.1.

The condition for the existence of the equilibrium in Theorem 13 is explained as follows. The equilibrium requires that the voter is indifferent to choice between the median and extremes. If an extreme platform can be chosen only by candidates observing the same state, the voter votes for a candidate who chooses the extreme because it is a perfect signal about the state. Therefore, to support the equilibrium of Theorem 13, some types must choose the extreme platform without actually observing the same extreme state so that the voter is indifferent between voting for an extreme and the median. When m is high and, thus, the state is likely to be the median, mixing from type $(C, 0)$, is enough to generate the indifference. On the other hand, when m is small and thus extremes are likely, it is impossible to generate the indifference if incompetent candidates always choose the median. Then we need not only type $(C, 0)$ but also type I, incompetent, to choose an extreme to generate the indifference, and those cases are treated in the next theorems.

Note the difference in players' behavior between the equilibria in Theorems 12 and 13. When the motivation for office is sufficiently high, i.e., $k > \kappa$, there is an equilibrium in which competent candidates try to utilize the effect of signaling competence to increase the probability of winning, by choosing extreme platforms. On the other hand, when the motivation for office is sufficiently low, i.e., $k \leq 2$, there is an equilibrium in which they simply cast aside the opportunity of pretending to be competent. Note that the parameter range that ensures the existence of equilibria in Theorem 12 or 13 covers all the cases of $m > 1/2$.

The fact that we have multiplicity of equilibria in a parameter range specified in Theorems 12 and 13 stems from the fact that the model has strategic complementarity among candidates. To see this, think of the case in which candidates' motivation for office is in the range where we have both types of equilibria, that is, $2 \geq k \geq \kappa(c, m)$. In the equilibrium of Theorem 12, even if an extreme platform has a much higher winning probability than the median, competent candidates observing the true state of 0 (type $(C, 0)$ candidates) do not polarize, while they do in the equilibrium of Theorem 13. This comes from the fact that in the former equilibrium, because the opponent, who always chooses the median when he is type $(C, 0)$, is likely to choose the median. This implies that the winning probability of the median is high compared to the case in which even type $(C, 0)$ candidates may choose an extreme, which keeps them from deviating to an extreme platform. On the other hand, in the equilibrium of Theorem 13, even a type $(C, 0)$ candidate may choose an extreme, which makes the winning probability of the median smaller, up to the point that the extreme and the median are indifferent to type $(C, 0)$ candidates.

2.3.3. High Uncertainty

In this subsection, we examine the case of high uncertainty, that is, $m \leq 1/2$. It is shown that the equilibria in such a case involve more polarization than in the case of low uncertainty.

The next two theorems deal with the cases in which a candidate's motivation for office is not too high and not too low. In there, the voter votes for an extreme candidate for certainty. In the equilibrium of Theorem 14, a competent candidate mildly polarizes in a sense that when he observes the true state of the median, he sometimes chooses an extreme policy and sometimes chooses the median, while incompetent candidates converge to the median. In the equilibrium of Theorem 15, on the other hand, a competent candidate completely polarizes in a sense that when he observes the true state of the median, he never chooses the median policy, while incompetent candidates converge to the median.

Theorem 14. *Let $m \leq 1/2$. There is a perfect Bayesian equilibrium that satisfies the following conditions if and only if $k \in (2-2c, 2-2cm)$.*

1. Competent candidates mildly polarize, that is,

$$g_0(0) > 0, \ g_0(-1) = g_0(1) > 0, \ \text{and } g_{-1}(-1) = g_1(1) = 1.$$

2. Incompetent candidates choose the median platform, that is, $g(0) = 1$.

3. The voter votes for an extreme candidate, that is, $v(-1,0) = 1$ and $v(1,0) =$ 1.

Theorem 15. *Let $m \leq 1/2$. There is a perfect Bayesian equilibrium that satisfies the following conditions if and only if $k \in [2-2c, 2-2cm]$.*

1. Competent candidates completely polarize, that is,

$$g_0(-1) + g_0(1) = 1, \ \text{and } g_{-1}(-1) = g_1(1) = 1.$$

2. Incompetent candidates choose the median platform, that is, $g(0) = 1$.

3. The voter votes for an extreme candidate, that is, $v(-1, 0) = 1$ *and* $v(1, 0) = 1$.

The equilibria given by Theorems 12, 14, and 15, respectively, are similar in the players' strategies, and the range of parameter values that ensure their existence. Between the three equilibria, only the behaviors of competent candidates observing the true state of the median (type $(C, 0)$ candidates) differ. Again, the reason for multiplicity of equilibria is the strategic complementarity between candidates generated by the voting strategy. A candidate has more incentive to polarize as more and more the opponent polarizes, because the median platform, which has a smaller probability of winning, becomes unlikely to win and is, thus, less attractive.

The equilibria characterized in Theorems 14 and 15 are supported by the candidates' intermediate degree of motivation for office. Since incompetent candidates are risk averse from our functional assumption of policy utility, choosing an extreme platform is very risky for them, while it is not for competent candidates observing the true state of the median. Therefore, as long as extreme platforms give them a higher winning probability, competent candidates, knowing that the true state to be the median, may be willing to be extreme because it does not hurt much. To support such a behavior, candidates' motivation for office should not be too high for incompetents to polarize, and not be too low for type $(C, 0)$ candidates to polarize. Also note that, given such behavior of candidates, the voter does not want to vote for an extreme unless it is likely that the state is the extreme, i.e., m is small.

The next result characterizes an equilibrium for the case in which k is high and probability of competence is not too high. In the equilibrium, competent candidates completely polarize while incompetent candidates mix between all policies.

Theorem 16. *Let $m \leq 1/2$. There is a perfect Bayesian equilibrium that satisfies the following conditions if and only if $c < \frac{1}{2(1-m)}$, and $k > \kappa(m,b)$, where $\kappa(c,m)$ is the unique number determined by (c,m) such that $\kappa(c,m) < 2-2cm$ for all (c,m).*

1. *Competent candidates completely polarize, that is,*

$$g_0(-1) + g_0(1) = 1, \text{ and } g_{-1}(-1) = g_1(1) = 1.$$

2. *Incompetent candidates mildly polarize, that is,*

$$g(0) > 0 \text{ and } g(-1) = g(1) > 0.$$

3. *The voting strategy is such that $W(-1) > W(0)$ and $W(1) > W(0)$.*

Again, the set of equilibria characterized in Theorem 16 contains the symmetric case, that is, $g_0(-1) = 1/2$, $v(0,1) = v(0,-1)$, and $v(-1,1) = 1/2$. This equilibrium has a similarity towards that of Theorem 13. In the equilibrium of

Theorem 16, the way the voter mixes strategies makes incompetent candidates indifferent between all platforms, while it makes type $(C,0)$ candidates indifferent in Theorem 13. On the other hand, the way incompetent candidates mix makes the voter indifferent between candidates for any combination of platform choices. That incompetent candidates are indifferent between the median platform and extreme platforms means a higher winning probability of extremes and competent candidates are willing to be extreme.

We can do some comparative statics in this equilibrium, and get somewhat similar results as in Theorem 13. Again, focus on the symmetric equilibrium where $v(0,-1) = v(0,1) > 0$ and denote $v(0,-1)$ and $g(0)$ by $\beta(k,c)$ and $g(k,c)$, respectively as functions of k and c. Then we have

$$\frac{\partial g(k,c)}{\partial c} > 0, \ \frac{\partial g(k,c)}{\partial k} = 0, \text{ and } \lim_{c \uparrow 1} g(k,c) = 0.$$

and

$$\frac{\partial \beta(k,c)}{\partial c} > 0, \ \frac{\partial \beta(k,c)}{\partial k} > 0, \text{ and } \lim_{k \uparrow \infty} \beta(k,c) = \frac{1}{2}.$$

Those comparative statics results come from almost the same reasoning as in the case of Theorem 13. Fixing other parameter values, increase in candidates' office motivation reduces the winning probabilities of extreme platforms. We can also see that those comparative statics results give the same relation between the voter's payoff and the degree of office motivation, as in the case of Theorem 13.

This equilibrium, however, cannot be supported when competence probability is very high. This is because in such a case, even when incompetent candidates put a high probability of choosing extremes, the voter still strictly prefers to vote for an extreme because the median state is unlikely. Then, we cannot support incompetent candidates' strategy in a way that the median is chosen with a strictly positive probability when motivation for office is high.

So far, we have focused mainly on equilibria where all the platforms can be chosen with strictly positive probabilities (no off-equilibrium platform choice). The following theorem demonstrates that there are no other equilibria than those characterized so far, if we confine our focus on the ones where there is no off-equilibrium platform choice.

Theorem 17. *Any equilibrium such that all the three platforms may be chosen is one of those characterized in Theorems 12 through 16.*

We briefly mention the possibilities of other types of equilibria. By appropriately choosing voter's off-equilibrium belief, it is possible to construct the following types of equilibrium. 1. An equilibrium such that only one platform is chosen by candidates. 2. An equilibrium such that only two platforms (0 and -1, or 0 and 1) may be chosen by candidates. We did not focus on those types of equilibria because we do not perceive those to be plausible, compared to the equilibria characterized above. Indeed, we can exclude the first type of equilibrium by applying usual D1 criterion. Because the type of candidate who has the strongest incentive to deviate to an off-equilibrium platform is competent observing the same state, the voter votes for the deviating candidate and, thereby, breaks the equilibrium. For the second type of equilibrium, although we do not find the usual application of D1 criterion is enough to invalidate it, we find that appropriately combining D1 with the requirement that the voter believes that the deviation is unilateral invalidates it.[14] All the equilibria shown in this section survive such refinements.

The equilibrium characterized in Theorem 11, which survives our refinement criteria, may be given a reasonable justification, at least for some parameter range. To see this, suppose that the median state is unlikely and candidates are likely to be competent. If the median platform is off-equilibrium choice, the type of candidates that have the strongest incentive to deviate is incompetent type. Then, after the voter observes one candidate chooses median while another chooses an extreme, her expectation about the true state is still sufficiently extreme, and, hence, she votes for the extreme candidate, which prevents candidates from deviate. Note also that the range of parameter values that we can render such justification, low m and high c, approximately corresponds to the range where equilibria of Theorem 12 through 16 are not supported.

2.4. Conclusion

This study examined a signaling game where a fraction of candidates have competence, which is unobservable to voters. We show that candidates have an

[14]It is discussed in Section 2.6.

incentive to polarize to make them appear competent and, thereby, increase their probability of winning. Depending on the parameter specifications, the equilibrium varies from the equilibrium in which only competent candidates polarize to the one in which every candidate polarizes. The general insight is that being extreme is advantageous for winning the election, because it makes the candidate appear competent.

An important aspect of our model is that candidates are assumed to commit to policies during the election, and the commitment is assumed to be credible. Some justification for this assumption can be made. For example, real-life implementation of a policy requires preparation, when there is a time lag between proposing a platform and actually implementing it and, thus, it is impossible to change policies flexibly. In our model, if the commitment is not credible,[15] the voter knows that, after the election, competent candidates will implement the best policy contingent on the revealed state of the world and incompetent candidates will implement the median. There are two types of equilibrium in such a setting. When candidates' policy motivation is sufficiently high, something reminiscent of the equilibrium of Theorem 12 emerges. The only difference is that now, one of the three platform works as a signal about incompetence or competence observing the true state of median. The name of a platform does not matter because now the game becomes a cheap-talk game. On the other hand, when candidates' policy

[15] Osborne and Slivinski (1996) and Besley and Coate (1997) consider models in which candidates cannot make commitments at all.

motivation is sufficiently small, all the equilibrium that may be announced in an equilibrium (on-equilibrium) should have the same winning probability.

In this chapter, in order to focus on the roles of vertical difference between candidates (competence), we assumed that every candidate shares the same policy preference, conditional on state. It may be interesting to extend the model to the case in which candidates have biased policy preferences that are common knowledge among all players. In such a case, there may be an effect such that proposing a platform that is opposite to a candidate's policy bias serves as a stronger signal about his competence. This type of extension needs to enlarge the policy space as in the standard continuum policy space model. However, in such an extension, the number of possible combinations of policy announcements becomes large and it is necessary to construct a large number of equilibrium beliefs, contingent on policy choices over the state, which is a highly difficult task that requires some simplifying assumptions on the way the voter beliefs are formed. For the same reason, extending our model into the one with continuum of state space is also difficult.

Finally, although our equilibrium heavily depends on mixing of strategies of players, it is possible to purify their strategies without changing the main insights. One way of doing this is to assume that parameter k for each candidate, which represents his degree of office motivation, has a non-degenerate distribution and also is private information about him. Then there is a cut-off value above which he chooses an extreme platform and below which he chooses the median. Similarly, it is possible to purify the voter's strategy by adding some preference parameter such as the degree of risk aversion, that is private information.

2.5. Appendix: Proofs

2.5.1. Proof of Theorem 12

If direction: Suppose that candidates' strategies are as described in the theorem. For a pair of off-equilibrium platform choices, i.e., -1 and 1 are chosen, let it satisfy $\mathbb{E}[\varphi(\theta,-1,1)] = \mathbb{E}[\varphi(\theta,1,-1)] = 0$. We will show that no player has an incentive to deviate.

Note that we have

$$U(1,1) - U(0,1) = U(-1,-1) - U(0,-1) = \frac{1}{2}ck + 1 - c \geq 0,$$

and

$$U(0,0) - U(-1,0) = U(0,0) - U(1,0) = -\frac{1}{2}k + 1 \geq 0.$$

Thus no competent candidates have an incentive to deviate.

It is straightforward to see that type I candidates do not have an incentive to deviate.

For the voter, the voting rule is optimal if $\sum_{\theta\in\Theta}\theta\varphi(\theta,1,0) \geq 1/2$, where the left hand side is 1. Those imply that no player has an incentive to deviate, and thus it is an equilibrium.

Only if direction: It is straightforward from "if" direction. *Q.E.D.*

2.5.2. Proof of Theorem 13

If direction: Suppose that the players strategies are as follows: $v(0,-1) = v(0,1) = \beta,\ v(1,-1) = 1/2,\ g(0) = 1,\ g_{-1}(-1) = g_1(1) = 1,\ g_0(0) = d$, and $g_0(-1) = g_{-1}(1) = (1-d)/2$. We will show that we can find $\beta \in [0,1/2)$ and $d \in (0,1)$ such that those strategies constitute an equilibrium. Note that if this is the case, $W(-1) > W(0)$ and $W(1) > W(0)$ follow immediately.

There are two ways of supporting those strategies as an equilibrium, that is, by $\beta = 0$ and $\beta > 0$.

First, we examine how we can support the equilibrium by $\beta > 0$. A strategy of type $(C,0)$ candidates is optimal when $U(1,0) = U(0,0)$ and $U(-1,0) = U(0,0)$. The former (and also the later) can be rewritten as

$$G(d,k) = U(1,0) - U(0,0) = (\frac{1}{2} - \beta)k + c - cd + \beta - 2c\beta + 2cd\beta - 1 = 0. \tag{2.7}$$

On the other hand, the voting strategy is optimal if and only if $\sum_{\theta\in\Theta} \theta\varphi(\theta,1,0) = 1/2$ and $\sum_{\theta\in\Theta} \theta\varphi(\theta,-1,0) = 1/2$, since in such a case, candidates proposing 0 and -1 (and 0 and 1) are equally preferred. Thus we have

$$\frac{\frac{1-m}{2}c(1-c)}{\frac{1-m}{2}c(1-c) + mc(1-c)\frac{1-d}{2} + 2mc^2 d\left(\frac{1-d}{2}\right)} = \frac{1}{2},$$

which can be rewritten as

$$F(d) = cmd^2 + (c-2cm)d + 1 - c - 2m + 2cm = 0. \tag{2.8}$$

We can see that F is a strictly increasing function and $F(1) = cm + 1 - c - m > 0$, we can find some $d \in (0,1)$ if and only only if $F(0) < 0$, which is $m > 1/2$. Let this value of d be $d(c,m)$. To prove the theorem, it is enough to show that given k, we can find β such that $G(\beta, d(c,m), k) = 0$. Since it is a strictly decreasing function of β and $G(\beta, d, k) < 0$ for all $\beta \geq 1/2$, we can find such a β if $G(0, d(c,m), k) > 0$, which holds if and only if $\frac{k+2c-2}{2c} > d(c,m)$. Thus it is an equilibrium when $k > \kappa = 2cd(c,m) + 2 - 2c > 2 - 2c$, because when those are satisfied, it is easy to see that the other types of candidates' strategies are optimal.

Second, we examine how we can construct an equilibrium such that $\beta = 0$. First, type $(C, 0)$ candidates' strategy is an optimal when $U(1,0) = U(0,0)$ and $U(-1,0) = U(0,0)$. The former (and also the later) can be rewritten as

$$G(\beta, d, k) = U(1,0) - U(0,0) = \frac{1}{2}k + c - cd - 1 = 0,$$

and thus $\widehat{d} = \frac{k+2c-2}{2c}$, and thus it must hold that $k \in (2-2c, 2)$. Denote this value by $d(k)$. On the other hand, the voting strategy, $\beta = 0$, is optimal if $\sum_{\theta \in \Theta} \theta \varphi(\theta, 1, 0) \geq 1/2$ and $\sum_{\theta \in \Theta} \theta \varphi(\theta, -1, 0) \geq 1/2$, since in such case, a candidate proposing -1 (and 1) is preferred to a candidate proposing 0. This can be rewritten as

$$F(d) = cmd^2 + (m - 2cm)d + 1 - c - 2m + 2cm \geq 0, \tag{2.9}$$

and the solution of the equation $F\left(\widehat{d}\right)=0$ is given by

$$\widehat{d}=\frac{-(m-2cm)+\sqrt{(m-2cm)^{2}-4cm\left(1-c-2m+2cm\right)}}{2cm}. \tag{2.10}$$

Then we have an equilibrium when $\frac{k+2c-2}{2c}\geq d\left(c,m\right),$ because in such a case $F\left(\widehat{d}\right)\geq 0.$

Only if direction: Suppose that the parameter values are outside of the two ranges, and we are given players' strategies such that $g\left(0\right)=1,\ g_{-1}\left(-1\right)=g_{1}\left(1\right)=1,\ g_{0}\left(0\right)>0$, and $g_{0}\left(-1\right)=g_{0}\left(1\right)>0$. We must have $U\left(-1,0\right)-U\left(0,0\right)=0$ and $U\left(1,0\right)-U\left(0,0\right)=0$ to be an equilibrium. If we regard $U\left(-1,0\right)-U\left(0,0\right)$ and $U\left(1,0\right)-U\left(0,0\right)$ as functions of $v\left(0,-1\right),\ v\left(0,1\right),\ v\left(-1,1\right)$, and $g_{0}\left(-1\right)$, we have

$$\begin{aligned}
\frac{\partial W_{-1}\left(x,y,z,a\right)}{\partial x} &= \frac{\partial W_{1}\left(x,y,z,a\right)}{\partial y}<0,\\
\frac{\partial W_{-1}\left(x,y,z,a\right)}{\partial y} &= \frac{\partial W_{1}\left(x,y,z,a\right)}{\partial x}<0,\\
\text{and }\left|\frac{\partial W_{-1}\left(x,y,z,a\right)}{\partial z}\right| &= \frac{\partial W_{1}\left(x,y,z,a\right)}{\partial z},
\end{aligned}$$

where $W_{-1}=U\left(-1,0\right)-U\left(0,0\right),\ W_{1}=U\left(1,0\right)-U\left(0,0\right),\ x=v\left(0,1\right),\ y=v\left(0,-1\right)$, and $v\left(-1,1\right)=z,\ g_{0}\left(-1\right)=a$. Since we have W_{-1} and W_{1} are decreasing functions of x and y, it is enough to show that $W_{-1}\left(0,0,z,a\right)\geq 0\ W_{1}\left(0,0,z,a\right)\geq 0$ cannot hold simultaneously for any combinations of z and a such that both $\sum_{\theta\in\Theta}\theta\varphi\left(\theta,-1,0\right)\leq -1/2$ and $\sum_{\theta\in\Theta}\theta\varphi\left(\theta,1,0\right)\geq 1/2$ hold. This follows from a simple calculation combined with the facts that $\left|\frac{\partial W_{-1}(x,y,z,a)}{\partial z}\right|=\frac{\partial W_{1}(x,y,z,a)}{\partial z}$ and that we cannot have $W_1\left(0,0,\frac{1}{2},a\right)\geq 0$ and $W_1\left(0,0,\frac{1}{2},a\right)\geq 0$ simultaneously for all a in a way that $\sum_{\theta\in\Theta}\theta\varphi\left(\theta,-1,0\right)\leq -1/2$ and $\sum_{\theta\in\Theta}\theta\varphi\left(\theta,-1,0\right)\geq 1/2$, which is shown in the "if" part of the proof.

Comparative statics results follow from (2.7) and (2.8). *Q.E.D.*

2.5.3. Proof of Theorem 14

It is proved in the same way as in Theorem 13. Note that when $m\leq 1/2$, function F defined by (2.8) satisfies $F\left(0\right)\geq 0$. *Q.E.D.*

2.5.4. Proof of Theorem 15

If direction: Suppose that k is in the range, $m\leq 1/2$, players' strategies are as follows: $v\left(0,-1\right)=v\left(0,1\right)=0$, $v\left(1,-1\right)=1/2$, $g\left(0\right)=1$, $g_{-1}\left(-1\right)=g_1\left(1\right)=1$, and $g_0\left(-1\right)=g_{-1}\left(1\right)=1/2$. We will show that no player has an incentive to deviate.

First note that

$$\mathbb{E}[u\left(0,\theta\right)]=-(1-m)\text{ and }\mathbb{E}[u\left(-1,\theta\right)]=\mathbb{E}[u\left(1,\theta\right)]=-\left(2-m\right).$$

For a type $\langle 0,\varnothing\rangle$ candidate, his expected payoff from choosing 0 is

$$\frac{(1-c)}{2}k-cm-\left(1-b\right)\left(1-m\right).$$

On the other hand, his payoff from deviating to $x = 1$ is

$$\left(\frac{2-c}{2}\right)k - c - (1-c)(2-m) = \left(\frac{2-c}{2}\right)k - 2 + c + m - cm.$$

Thus his no-deviation condition is $2 - 2cm \leq k$.

For type $\langle 1, 0\rangle$, his expected payoff from choosing -1 is $\left(\frac{2-c}{2}\right)k - 1$, whereas his expected payoff from deviating to $x = 0$ is $\frac{(1-c)}{2}k - c$. Thus his no-deviation condition is $k \geq 2 - 2c$.

Those strategies can be supported as an equilibrium when $\sum_{\theta\in\Theta} \theta\varphi(\theta, -1, 0) \leq -1/2$ and $\sum_{\theta\in\Theta} \theta\varphi(\theta, 1, 0) \geq 1/2$, which can be written as

$$\frac{\frac{1-m}{2}[c(1-c)]}{\frac{1-m}{2}[c(1-c)] + m[\frac{1}{2}c(1-c)]} \geq \frac{1}{2} \rightarrow m \leq 1/2.$$

Only if direction: It is straightforward from the proof of "if" direction. *Q.E.D.*

2.5.5. Proof of Theorem 16

If direction: Suppose that the players strategies are as follows: $v(0, -1) = v(0, 1) = \beta$, $v(1, -1) = 1/2$, $g(0) = 1$, $g_{-1}(-1) = g_1(1) = 1$, $g_0(-1) = g_0(1) = 1/2$, and $g(-1) = g(1) = (1-d)/2$, and $g(0) = d$. We will show that we can find $\beta \in [0, 1/2)$ and $d \in (0, 1)$ such that those strategies constitute an equilibrium. Note that if this is the case, $W(-1) > W(0)$ and $W(1) > W(0)$ follow immediately.

First, we examine when we can construct an equilibrium in a way such that $\beta > 0$. Then, obviously, we must have $\mathbb{E}[U(-1, \theta)] = \mathbb{E}[U(0, \theta)]$ and $\mathbb{E}[U(1, \theta)] =$

$\mathbb{E}[U(0,\theta)]$, since in such case, candidates proposing 0 and -1 (and 0 and 1) are equally preferred.

$$G(\beta, d, k) = \mathbb{E}_\theta[U(1,\theta)] - \mathbb{E}_\theta[U(0,\theta)] = 0. \tag{2.11}$$

Also, we have $\sum_{\theta\in\Theta}\theta\varphi(\theta,1,0) = 1/2$ and $\sum_{\theta\in\Theta}\theta\varphi(\theta,-1,0) = 1/2$. Those can be rewritten as

$$\frac{\frac{1-m}{2}[c(1-c)d + (1-c)^2 d\frac{1-d}{2}] - \frac{1-m}{2}[(1-c)^2 d\left(\frac{1-d}{2}\right)}{\Psi} = \frac{1}{2}, \tag{2.12}$$

, where

$$\begin{aligned}\Psi &= \frac{1-m}{2}[c(1-c)d + (1-c)^2 d\frac{1-d}{2}] \\ &\quad + m[c(1-c)\frac{1}{2}d + (1-c)^2 d\left(\frac{1-d}{2}\right)] + \frac{1-m}{2}[(1-c)^2 d\left(\frac{1-d}{2}\right)]\end{aligned}$$

This can be rewritten as

$$F(d) = 2c + d - 2cm - cd - 1 = 0. \tag{2.13}$$

Then it follows immediately from $F\left(\widehat{d}\right) = 0$ that

$$\widehat{d} = \frac{1-2c+2cm}{1-c} \in (0,1) \text{ and } c < \frac{1}{2(1-m)}. \tag{2.14}$$

Therefore, to prove the theorem, it is enough to show that given k, equation (2.11), as a function of β, has a solution in $(0, 1/2)$ when $d = \widehat{d}$. It can be easily seen that

$G(\beta, d, k) < 0$ for all $\beta \geq 1/2$, d, and k. Let κ be the value of k that satisfies $G\left(0, \widehat{d}, \kappa\right) = 0$. Since G is a strictly increasing function of k when $\beta = 0$, we can find such κ. Then it follows that when $k > \kappa$, we can find β such that $G\left(\beta, \widehat{d}, k\right) = 0$ from $(0, 1/2)$. It is easy to see that $\kappa < 2 - 2cm$.

Next, we examine when we can construct an equilibrium in a way such that $\beta = 0$. First, type $\langle 0, \varnothing \rangle$ candidates' strategy is optimal when $\mathbb{E}[U(-1, \theta)] = \mathbb{E}[U(0, \theta)]$ and $\mathbb{E}[U(1, \theta)] = \mathbb{E}[U(0, \theta)]$. The former (and also the later) can be rewritten as

$$
\begin{aligned}
(2.15) \qquad & G(\beta, d, k) = \mathbb{E}_\theta[U(1, \theta)] - \mathbb{E}_\theta[U(0, \theta)] \\
= \quad & \frac{1}{2}k + m[cd - d] + \frac{(1-m)}{2}\left[\frac{3}{2}cd - 2d - \frac{3}{2}c - \frac{1}{2}c^2 + \frac{1}{2}c^2 d\right] = 0.
\end{aligned}
$$

and thus $d = \frac{k+2c-2}{2c}$, and thus it must hold that $k \in (2 - 2c, 2)$. Denote this value by $d(k)$. On the other hand, the voting strategy is optimal if $\sum_{\theta \in \Theta} \theta \varphi(\theta, 1, 0) \geq 1/2$ and $\sum_{\theta \in \Theta} \theta \varphi(\theta, -1, 0) \geq 1/2$, since in such case, a candidate proposing 1 (and -1) is preferred to a candidate proposing 0. This can be rewritten as $F(d) \geq 0$. It is easy to see that the inequality holds for all d so long as $m \leq 1/2$, and thus the equilibrium can be supported when $k \in (2 - 2c, 2)$.

The proof of the "only if" direction is almost the same as in Theorem 13 and thus omitted. *Q.E.D.*

2.5.6. Proof of Theorem 17

Given an equilibrium strategies $(g, g_\theta, v, \varphi)$, denote by $\Upsilon(g, g_\theta, v, \varphi)$ the set of platforms that are chosen with strictly positive probabilities, i.e., $\Upsilon(g, g_\theta, v, \varphi) = \{x| \max\{g(x), \sum_\theta g_\theta(x)\} > 0\}$ (we sometimes abbreviate and just write Υ). In the following proofs, we often use the necessary conditions for having $U(x,\theta) \geq U(y,\theta)$ for $|x-\theta| > |y-\theta|$, which we can derive by explicitly writing down the inequality by using (2.1). We do not provide the explicit form every time it appears because it is space consuming. One example is

$$
\begin{aligned}
&U(-1,\theta) - U(0,\theta) = \{c(1-g_\theta(1)) + (1-c)(1-g(1))\}\left(\frac{1}{2} - v(0,-1)\right)k \\
&+\{cg_\theta(1) + (1-c)g(1)\}(v(-1,1) - v(0,1))k \\
&+\sum_{x\in\{-1,0,1\}}\{cg_\theta(x) + (1-c)g(x)\}\times \\
&\{v(0,x)u(0,\theta) + v(x,0)u(x,\theta) - v(-1,x)u(-1,\theta) - v(x,-1)u(x,\theta)\}
\end{aligned}
$$

from which we can see that the difference in expected office rent between proposing -1 and 0 does not depend on $g_\theta(0)$, $g_\theta(-1)$, $g(0)$ and $g(-1)$. This property is going to be implicitly used very often.

Lemma 10. *If* $\Upsilon(g, g_\theta, v, \varphi) = X$, *the followings cannot hold.*

1. $U(-1,-1) = U(0,-1) = U(1,-1)$.

2. $U(-1,1) = U(0,1) = U(1,1)$.

Proof. Suppose that $U(-1,-1) = U(0,-1) = U(1,-1)$. It is easy to see that we have to have $k > 1$, and it can be easily verified the following: 1. $U(-1,-1) = U(0,-1)$ implies either a. $v(0,-1) > 1/2$ or x. $v(0,1) > v(-1,1)$. 2. $U(1,-1) = U(-1,-1)$ implies either b. $v(1,-1) > 1/2$ or y. $v(1,0) > v(-1,0)$. 3. $U(0,-1) = U(1,-1)$ implies either c. $v(1,0) > 1/2$ or z. $v(1,-1) > v(0,-1)$. We will show that each combination leads to a contradiction.

Case abc. If $v(1,0) \geq v(-1,0)$, it must hold that $\{(C,-1)\} = P(-1)$, since $\mathbb{E}[U(-1,\theta)] < \mathbb{E}[U(1,\theta)]$ from it and b. In such a case, from the voter's optimization condition, $v(-1,0) = 1$, which is a contradiction. Hence we must have $v(-1,0) > v(1,0)$. However, this contradicts a and c.

Case abz. By using the same logic used above, we can derive $v(-1,1) > v(0,-1)$. However, this contradicts z.

Case ayz. Same as in case abz.

Case ayc. From the same kind of reasoning used above, we must have $v(1,-1) < 1/2$. Then, to have $U(-1,-1) = U(1,-1)$, we must have $v(1,0) > v(0,-1)$, which contradicts y.

Case xbc, xyc, xbz, and xbz. The fact that x is the only possibility for having $U(-1,-1) = U(0,-1) = U(1,-1)$ and we must have the symmetric condition for having $U(-1,1) = U(0,1) = U(1,1)$ implies that we have either $U(1,1) > U(-1,1)$ or $U(1,1) > U(0,1)$.

Think of the case in which $U(1,1) > U(-1,1)$. Then, unless $U(-1,0) > U(1,0)$, we have $\{(C,-1)\} = P(-1)$, and thus $v(-1,0) = v(-1,1) = 1$, a contradiction. Hence $U(-1,0) > U(1,0)$. We can check that this, x, and $U(-1,-1) = U(1,-1)$ imply that $g_0(0) > 0$ and hence $U(0,0) \geq U(-1,0)$. This implies that $U(-1,1) \geq U(0,1)$, since if it is not the case, we must have $\{(C,-1)\} = P(-1)$. Then, we have $U(1,1) > U(0,1)$ and $g_1(1) = 1$, but $U(-1,1) \geq U(0,1)$ and $U(0,0) \geq U(-1,0)$ requires $g_0(1) > g_1(1) = 1$, which is a contradiction.

Therefore, think of the case in which $U(1,1) = U(-1,1) > U(0,1)$. From the same argument already used, we must have $U(0,0) > U(1,0) = U(-1,0)$. However, $U(1,-1) = U(-1,-1)$ and $U(-1,1) = U(1,1)$ imply $g_1(0) > 0$, which is a contradiction. $\square$

The following lemma demonstrates that in all equilibrium such that $\Upsilon(g, g_\theta, v, \varphi) = \{-1,0,1\}$, it must hold that $g_{-1}(-1) = 1$ and $g_1(1) = 1$.

Lemma 11. *If* $\Upsilon(g, g_\theta, v, \varphi) = \{-1,0,1\}$,

1. $U(-1,-1) > \max\{U(0,-1), U(1,-1)\}$.

2. $U(1,1) > \max\{U(0,1), U(1,1)\}$.

Proof. 1. First, to get a contradiction, suppose that $U(-1,-1) = U(0,-1) > U(1,-1)$. Then we must have either A. $v(0,-1) > 1/2$ and $v(0,1) < v(-1,1)$, or B. $v(0,-1) < 1/2$ and $v(0,1) > v(-1,1)$, because otherwise we have $\{(C,-1)\} = P(-1)$ or $\{(C,0)\} = P(-1)$, which leads to $v(-1,0) = v(-1,1) = 1$ and thereby $U(-1,-1) > U(0,-1)$, a contradiction.

Case A. In such a case, not to have $\{(C,-1)\} = P(-1)$, which leads a contradiction, we have two possibilities as follows: 1. $U(-1,0) \geq U(0,0)$ and $U(-1,1) \leq U(0,1)$. 2. $U(-1,0) < U(0,0)$ and $U(-1,1) > U(0,1)$. We induce contradiction for each case. Think of the first case. From $U(-1,-1) = U(0,-1)$ and $U(-1,0) \geq U(0,0)$, we must have $g_0(1) > g_{-1}(1) = 0$, which implies $U(1,0) \geq U(0,0)$. This and $U(1,1) \leq U(0,1)$ imply $g_0(-1) > g_1(-1) = 0$ and $v(1,-1) - v(0,-1) > 0$. However, those imply that $U(1,0) > U(-1,0)$, which contradicts $g_0(-1) > 0$. Think of the second case. $U(-1,1) > U(0,1)$ and $U(-1,-1) = U(0,-1)$ implies $g_0(1) > 0$. Then we have to have $U(1,0) \geq U(0,0)$, which implies either $v(1,-1) - v(0,-1) > 0$ or $v(1,0) > 1/2$. Moreover, $U(1,0) \geq U(-1,0)$ implies that we have to have either $v(1,-1) < 1/2$ or $v(-1,0) - v(1,0) > 0$. It can be easily checked that only the case of $v(1,0) > 1/2$ and $v(0,-1) > 1/2$ are compatible with A. However, it such a case, we have $U(-1,-1) - U(1,-1) < |U(-1,1) - U(1,1)|$ and thus $\mathbb{E}[U(-1,\theta)] < \mathbb{E}[U(1,\theta)]$ and $g(1) = 0$. It implies that $v(-1,1) = 1$, which is a contradiction.

Case B. From $g_{-1}(1) = 0$, we have $U(-1,1) - U(0,1) < [U(-1,-1) - U(0,-1)] = 0$ from a simple calculation. Then, we have to have $U(-1,0) - U(0,0) > 0$, since otherwise, $\{(C,-1)\} = P(-1)$ follows. However, $U(-1,0) - U(0,0) > 0$ and $U(-1,-1) - U(0,-1) = 0$ imply $g_{-1}(1) > g_0(1) > 0$, which is a contradiction.

The proof for the case $U(-1,-1) = U(1,-1)$ uses the same kind of argument, and thus omitted. The proof for the second statement is the same. $\square$

Lemma 12. *1.If $\Upsilon(g, g_\theta, v, \varphi) = \{-1, 0, 1\}$, then $g(-1) = g(1)$.*

2. If $\Upsilon(g, g_\theta, v, \varphi) = \{-1, 0, 1\}$, $0 < g_0(0) \leq 1$ implies $g(0) = 1$.

3. If $\Upsilon(g, g_\theta, v, \varphi) = \{-1, 0, 1\}$, $g(0) = 1$ implies $g_0(-1) = g_0(1)$.

Proof. 1. First, take an equilibrium such that $\Upsilon(g, g_\theta) = \{-1, 0, 1\}$ and $g_0(0) > 0$. Suppose $g(-1) > 0$. Then we have $\mathbb{E}[U(-1, \theta)] \geq \mathbb{E}[U(0, \theta)]$, which is possible only when $U(-1, -1) + U(-1, 1) \geq U(0, -1) + U(0, 1)$, because we already have $U(0, 0) \geq U(-1, 0)$. This and $g_1(1) = 1$ imply that $v(-1, 1) > v(0, 1) > 0$. Because $v(-1, 1) > 0$ holds only when $\sum_{\theta \in \Theta} \theta \varphi(\theta, -1, 1) \leq 0$, we must have $g(-1) \leq g(1)$. This implies $\mathbb{E}[U(1, \theta)] \geq \mathbb{E}[U(0, \theta)]$, which is possible only when $U(1, -1) + U(1, 1) > U(0, -1) - U(0, 1)$. This and $g_{-1}(-1) = 1$ imply that $v(1, -1) > v(0, -1) > 0$. However, this implies $\sum_{\theta \in \Theta} \theta \varphi(\theta, -1, 1) \geq 0$ and thus $g(-1) \geq g(1)$. Hence we have $g(-1) = g(1)$.

Next, take an equilibrium such that $\Upsilon(g, g_\theta, v, \varphi) = \{-1, 0, 1\}$ and $g_0(0) = 0$. Suppose that $g(1) > g(-1)$. Then, $\sum_{\theta \in \Theta} \theta \varphi(\theta, -1, 1) < 0$ follows and we have $v(-1, 1) = 1$. If, moreover, $g_0(1) \geq g_0(-1)$, we have $\left|\sum_{\theta \in \Theta} \theta \varphi(\theta, -1, 0)\right| > \left|\sum_{\theta \in \Theta} \theta \varphi(\theta, 1, 0)\right|$ and hence either $v(0, 1) = 1$ or $v(-1, 0) = 1$ follows. In both cases, we have $U(-1, 0) > U(1, 0)$, which is a contradiction. Hence we have $g_0(1) < g_0(-1)$, and thus $U(-1, 0) \geq U(1, 0)$. Suppose that $U(-1, 0) = U(1, 0)$. Then, using $v(-1, 1) = 1$, we can calculate that $U(1, 1) - U(-1, 1) + U(1, -1) - U(-1, -1) = U(1, 0) - U(-1, 0) + 4c > 0$, and thus $\mathbb{E}[U(-1, \theta)] - \mathbb{E}[U(1, \theta)] > 0$. This implies $g(1) = 0$, which is a contradiction. Next, suppose that $U(-1,0) > U(1,0)$. Then we have $g_0(-1) = 1$. However, in such a case, we can get $\mathbb{E}[U(-1,\theta)] - \mathbb{E}[U(1,\theta)] > 0$, which contradicts $g(1) > 0$. Those imply that $g(-1) = g(1)$.

2. Take an equilibrium such that $\Upsilon(g, g_\theta) = \{-1, 0, 1\}$. We prove the statement by showing that $\mathbb{E}[U(0,\theta)] - \mathbb{E}[U(1,\theta)] > U(0,0) - U(1,0)$ and $\mathbb{E}[U(0,\theta)] - \mathbb{E}[U(-1,\theta)] > U(0,0) - U(-1,0)$. Suppose that $\mathbb{E}[U(0,\theta)] - \mathbb{E}[U(-1,\theta)] \leq U(0,0) - U(-1,0)$. It must hold that $U(0,0) - U(-1,0) = 0$, since otherwise we have $\{(C,-1)\} = P(-1)$, which leads to a contradiction.

If $\mathbb{E}[U(0,\theta)] - \mathbb{E}[U(-1,\theta)] < U(0,0) - U(-1,0) = 0$, we have $\{(C,0)\} = P(0)$ and hence $v(0,1) = v(0,-1) = 1$. This implies that $g_0(0) = 1$. In such case, it can be calculated that $\mathbb{E}[U(0,\theta)] - \mathbb{E}[U(-1,\theta)] > 0$, which is a contradiction. Therefore, we must have $\mathbb{E}[U(0,\theta)] - \mathbb{E}[U(-1,\theta)] = 0$. Then if moreover $U(1,0) > U(0,0)$ or $U(1,0) < U(0,0)$ holds, we have $\{(C,-1)\} = P(-1)$ for the former case and $\{(C,1)\} = P(1)$ for the later case, which leads to a contradiction. Hence we have $U(1,0) = U(0,0)$ and $\mathbb{E}[U(1,\theta)] - \mathbb{E}[U(0,\theta)] = 0$. Since we must have $\left|\sum_{\theta\in\Theta} \theta\varphi(\theta,-1,1)\right| = 0$ to support such an equilibrium under $g_1(1) = g_{-1}(-1) = 1$, it must hold that $bg_0(1) + (1-b)g(1) = bg_0(-1) + (1-b)g(-1)$. However, then, we can calculate that $\mathbb{E}[U(1,\theta)] + \mathbb{E}[U(1,\theta)] - 2\mathbb{E}[U(0,\theta)] < 0$. This shows that we cannot have both $\mathbb{E}[U(1,\theta)] - \mathbb{E}[U(0,\theta)] = 0$ and $\mathbb{E}[U(-1,\theta)] - \mathbb{E}[U(0,\theta)] = 0$ simultaneously, which is a contradiction.

3. Because if $g_0(-1) > g_0(1)$, in addition to $g(0) = 1$, we have both

$$\left|\sum_{\theta\in\Theta}\theta\varphi(\theta,1,0)\right| > \left|\sum_{\theta\in\Theta}\theta\varphi(\theta,-1,0)\right| \text{ and } \left|\sum_{\theta\in\Theta}\theta\varphi(\theta:1,-1)\right| > 0,$$

that lead $v(1,0) \geq v(-1,0)$ and $v(1,-1) > 1/2$, respectively, which contradicts $g_0(-1) > 0$. □

Thus, combining Lemma 10 through 12, it follows that any equilibrium such that every platform may be chosen, i.e., $\max\{g(x), \sum_\theta g_\theta(x)\} > 0$ for all x, candidates' strategies are one of those characterized in Section 2.3. To see that we also have $W(-1) > W(0)$ and $W(1) > W(0)$, suppose that $W(-1) \leq W(0)$. Then, it follows that $U(0,0) > U(-1,0)$ and hence $g_0(-1) = 0$, and $g(-1) = 0$. Then $\{(C,-1)\} = P(-1)$ and thus $v(-1,0) = v(-1,1) = 1$, which is a contradiction. *Q.E.D.*

2.6. Appendix: Equilibrium Refinement

Here, we discuss some refinement issue. Towards this end, we apply the D1 refinement, proposed by Cho and Kreps (1987). In our context, it requires that the voter does not attribute a deviation to a particular type of candidate if there is some other type that is willing to make the deviation for a strictly larger set of possible voting rules.

In our model, however, the simple application of D1 criterion is not powerful enough to refine equilibrium, because we cannot derive a set inclusion relation between the set of voting rules that induces a particular type of candidate to deviate. To get around this, we modify the usual D1 criterion by applying the idea first invented by Bagwell and Ramey (1991). It uses the fact that the types of two candidates are strongly correlated and type of candidates are also correlated with the state. For example, it is not possible that the voter perceives a candidates' type to be $(C, 1)$ at the same time while perceiving another candidate to be type $(C, 0)$. Once we require that after one candidate deviates the voter should still believe that another candidate follows the equilibrium strategy, this property of correlated types should impose some conditions on off-equilibrium belief.[16] Our modified D1 criterion makes use of this intuition.

To formally define the modified D1 criterion of our model, given an equilibrium $(g, g_\theta, v, \varphi)$, let $\Upsilon(g, g_\theta, v, \varphi)$ be the set of platforms that are chosen with strictly positive probabilities. Moreover, let $V(t)$ be the equilibrium payoff for type t candidates in equilibrium $(g, g_\theta, v, \varphi)$, which is defined by (2.1) for competent candidates and (2.2) for incompetent candidates. Also, denote by $P(x)$ the set of types of candidates that choose a particular platform with strictly positive probability, that is, $(C, \theta) \in P(x)$ if $g_\theta(x) > 0$ and $I \in P(x)$ if $g(x) > 0$.

Then given a tuple $(g, g_\theta, v, \varphi)$, define $\Xi_{(g,g_\theta,v,\varphi)}$ be the set of belief functions such that for all pair of x, y such that $x \notin \Upsilon$, $y \in \Upsilon$ and $I \notin P(y)$, we have

[16]Bagwell and Ramey (1991) construct two periods oligopoly model in which the incumbents' pricing choices, which may signal their production cost, are followed by the entrant's entering decision. In their refinement of what they call unprejudiced beliefs, the entrant assumes a single deviation happens after off-equilibrium choices of incumbents.

$\sum_{\theta\in\Theta}\theta\varphi'(\theta,x,y) \in co\{\theta|(C,\theta) \in P(y)\}$, where coA means convex hull of set A. Thus $\Xi_{(g,g_\theta,v,\varphi)}$ is the set of belief functions such that the voter's belief about the type of the non-deviating candidate is not affected by the deviating candidate, that is, the deviation is unilateral. To get the sense of this condition, suppose that $\{(C,1)\} = P(1)$ and $0 \notin \Upsilon$, and thus platform 1 is chosen only by candidate type $(C,1)$ and platform 0 is never chosen in an equilibrium. Then as long as the voter assumes that the deviation is unilateral, after seeing zero probability platform choices of $(0,1)$, she believes that platform 1 is chosen by a type $(C,1)$ candidate, because only that type chooses 1, and hence the true state is 1. Moreover, let Ψ be the set of voting rule that can be justified by some element from $\Xi_{(g,g_\theta,v,\varphi)}$, that is, each voting rule in Ψ maximizes the voter's expected payoff, under some belief in $\Xi_{(g,g_\theta,v,\varphi)}$.

Next, given an equilibrium (g,g_θ,v,φ), for each pair of type $t\in T$ and an off-equilibrium platform choice $p\notin\Upsilon$ let $D_t(p)$ and $D_t^+(p)$ be sets of functions from $X\times X$ to $[0,1]$ that are defined, respectively, as follows:

$$
\begin{aligned}
D_t(p) &= \{v^0 : U(p,\theta) \geq V(t)\} \cap \Psi \text{ if } t=(C,\theta) \\
&= \{v^0 : \mathbb{E}_\theta[U(p,\theta)] \geq V(t)\} \cap \Psi \text{ if } t=I. \\
D_t^+(p) &= \{v^0 : U(p,\theta) > V(t)\} \cap \Psi \text{ if } t=(C,\theta) \\
&= \{v^0 : \mathbb{E}_\theta[U(p,\theta)] > V(t)\} \cap \Psi \text{ if } t=I,
\end{aligned}
$$

where $U(p,\theta)$ is calculated by (2.1) with voting rule v^0 and candidates' strategies g and g_θ of the equilibrium. Note that these sets can be empty.

Above definitions say that, $D_t(p)$ is the set of voting rules that make type t candidates weakly prefer to deviate to off-equilibrium platform p, when the voter's belief and voting strategy satisfy the particular restriction and the opponent is supposed to follow equilibrium strategy (g, g_θ). Similarly, $D_t^+(p)$ is defined with strict preference. An electoral equilibrium that satisfies D1 criterion is the equilibrium satisfying the following condition, in addition to the conditions for a Bayes-Nash equilibrium:

Definition 4. *An equilibrium $(g, g_\theta, v, \varphi)$ satisfies the D1 criterion if for all pair of x and y such that $x \in \Upsilon(g,\ g_\theta, v, \varphi)$, $y \notin \Upsilon(g,\ g_\theta, v, \varphi)$, the followings are satisfied:*

1. It holds that $\varphi(\theta, x, y) = 1$ if $\{(C,\theta)\} = P(x)$, and $\sum_{\theta\in\Theta} \theta\varphi(\theta, x, y) \in$ int co$\{\theta | (C,\theta) \in P(x)\}$ if $\sharp P(x) > 1$.

2. If $D_I(y) \in D_t^+(y)$ for some t then $\varphi(\theta, x, y) = 1$ if $D_{t'}(y) \in D^+_{(C,\theta)}(y)$ for all $t' \neq (C,\theta)$ otherwise $\sum_{\theta\in\Theta} \theta\varphi(\theta, x, y) \in$ int co$\{\theta | D_{(C,\theta)}(y) \notin D_t^+(y)$ for all $t\}$.

In the definition, int$A$ is the interior of set A. Although the first condition is not at all related to usual D1 requirement, we include it as a natural requirement of the voter's belief formation after the unilateral deviation of a candidate. The second condition corresponds to usual D1 requirement adequately translated into our model. Note that we have no restriction on off-equilibrium beliefs when incompetent candidates are not excluded from the possibility of being the deviator, because the voter is free to think that the deviator should be incompetent in which case his deviation tells nothing about the state. We have the following result, which demonstrates that in any equilibrium that satisfies D1, extreme platforms signal competence and have strictly higher winning probabilities.

Theorem 18. *An equilibrium that satisfies D1 is one of those characterized in Section 2.3.*

Proof. We first show that there is no equilibrium that satisfies D1 such that Υ is a singleton. Towards this end, suppose that there is an equilibrium $(g, g_\theta, v, \varphi)$ that satisfies $D1$ such that Υ is a singleton. First think of the case in which $\Upsilon = \{1\}$. Then, we have

$$D_t(-1) \subset D^+_{(C,-1)}(-1) \text{ for all } t \in \{I, (C,1), (C,0)\}.$$

Then D1 implies that if the voter observes platform pair $(-1, 1)$, her off-equilibrium belief puts the whole mass on the event that the deviating candidate to be type $(C, -1)$. However, then it must be the case that $v(-1, 1) = 1$. This implies that type $(C, -1)$ candidates have an incentive to deviate to 0, which is a contradiction. The same proof applies to the case in which $\Upsilon = \{-1\}$.

Next, think of the case in which $\Upsilon = \{0\}$. Then we have

$$D_t(1) \subset D^{+}_{(C,1)}(1) \text{ for all } t \in \{I, (C,0), (C,-1)\}.$$

Then D1 implies that if the voter observes platform pair $(1,0)$, off-equilibrium belief puts the whole mass on the event that the deviating candidate being type $(C,1)$. However, then it must be the case that $v(1,0) = 1$ and $v(0,1) = 0$. Those imply that type $(C,1)$ candidates have an incentive to deviate to 1, which is a contradiction.

Next, we show that there is no equilibrium that satisfies D1 such that $\Upsilon = \{0,1\}$ or $\Upsilon = \{0,-1\}$. Towards this end, take an equilibrium such that $\Upsilon = \{0,1\}$. By using $g_\theta(-1) = g(-1) = 0$ for all θ and (1) we can show that $g_{-1}(0) = 1$. To see this, if $g_{-1}(0) < 1$, it must hold that $U(1,-1) - U(0,-1)$, which implies $\left(v(0,1) - \frac{1}{2}\right)k + 3v(0,1) \leq 0$, which in turn implies that both $U(1,0) - U(0,0) > 0$ and $U(1,1) - U(0,1) > 0$ hold, and thus $\mathbb{E}_\theta[U(0,\theta)] > \mathbb{E}_\theta[U(1,\theta)]$. Then only type $(C,-1)$ candidates choose 0, which implies that $v(0,1) = 1$ that contradicts $U(1,-1) - U(0,-1) \geq 0$. Hence, $g_{-1}(0) = 1$. Similarly, we can show that $g_1(1) = 1$. Moreover, we have $g_0(0) > 0$, since otherwise $\sum_{\theta\in\Theta}\varphi(\theta,0,1) < 1/2$ and $v(0,1) = 1$, which is a contradiction. Also, it can be shown that $v(1,0) > 1/2$, since otherwise, we must have $U(0,0) - U(1,0) > 0$ and $\mathbb{E}_\theta[U(0,\theta)] - \mathbb{E}_\theta[U(1,\theta)] \geq 0$, which imply $g_0(0) = 1$ and $g(0) = 1$, respectively. However, this leads $\sum_{\theta\in\Theta}\varphi(\theta,0,1) = 1$ and thereby $v(0,1) = 1$, which is a contradiction. Finally, $g_0(0) > 0$ and $\mathbb{E}_\theta[U(0,\theta)] - \mathbb{E}_\theta[U(1,\theta)] > U(0,0) - U(1,0)$ imply $g(0) = 1$.

We will show that

$$D_t(-1) \subset D^+_{(C,-1)}(-1) \text{ for all } t \in \{I,(C,0),(C,1)\}. \tag{2.16}$$

Note that from $(C,-1) \notin P(1)$, $I \notin P(1)$, and D1, it must hold that $v'(-1,1) = 0$ for all $v' \in \Psi$. Note that whether $v' \in D^+_{(C,-1)}(-1)$ or not depends only on whether $v'(-1,0)$ is strictly larger than some threshold value or not. More precisely, $v' \in D^+_{(C,-1)}(-1)$ if and only if

$$v'(-1,0)\,k - (1 - v'(-1,0)) > \frac{1}{2}k - 1. \tag{2.17}$$

However, it is easy to see that (2.17) is a necessary condition for $v' \in D_t(t)$, from which (2.16) follows. $\square$

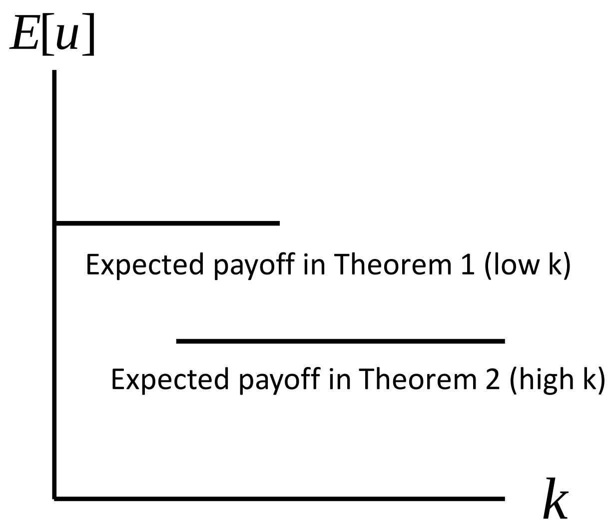

Figure 2.1. Comparative statics on the voter's expected payoff

CHAPTER 3

Cheap Talk and Observational Learning in a Market for Advice

3.1. Introduction

In a recent paper, Che, Dessein, and Kartik (2011) examine how an agent advises a principal on selecting one of multiple projects or an outside option. The agent is privately informed about the projects' benefits and shares the principal's preferences except for not internalizing her value from the outside option. They show that a distortion in advice arises in which the agent biases his recommendation toward "conditionally better-looking projects." The purpose of this chapter is to examine the dynamic consequence of such an advice structure, when a new principal-agent pair, who can observe previous decisions, is formed each period.

In our model, there are two projects and an outside option that the principal can implement, where each project has either a high or low fundamental quality. At each period, idiosyncratic shocks are added to the fundamental qualities and are observed only by an agent who advises the principal. Subsequent players observe principals' earlier choices and can make inference about the two projects' fundamental qualities. The model can be interpreted as stating that product/project have a fundamental common quality and an idiosyncratic match quality. Examples include markets for advice (e.g. constants, or financial brokers), and the academic hiring process, which takes place every year. With a slight modification of the model, it is also possible to capture situations that either products changing over time or products being fixed, but information gradually accruing over time about each product, with agents always fully-informed about all currently available information.

The learning process evolves in a way that if a project is chosen by a principal, public belief is updated toward favoring the fundamental quality of the project, while disfavoring that of other projects. Pandering distortion in advice arises once sufficiently unbalanced public beliefs about the fundamental qualities are formed. After that, successors should extract information from the distorted advice structure. Natural questions are (i) what are the robust long-run outcomes of the learning process in sequential communication? (ii) do principals settle on one choice, and if so, under what conditions?

We show that, depending on the idiosyncratic shocks in the early periods, it is possible that effective communication can be perpetually sustained. There are two scenarios for this to happen; in one scenario, the communication structure eventually becomes the phase in which an agent simply recommends the better project, and a principal, without reaching an accurate belief of the fundamental qualities, rubber-stamps the recommendations. In another scenario, an agent distorts his recommendation, as in Che, Dessein, and Kartik (2011)'s pandering equilibrium.

The first scenario is sustained only when two projects share the same fundamental quality. In such a case, as long as agents keep recommending the better project, or as long as there is no distortion in recommendations, both projects are recommended with the same frequency, which tells principals that the relative qualities of two projects are equal. However, the absolute levels of the two projects' fundamental qualities are never revealed. In this sense, we have perpetual communication with incomplete learning. This scenario provides the optimal communication structure for the principal if both projects have a high fundamental quality.

The second scenario arises only when one of the projects has a high fundamental quality and another has a low fundamental quality. In such a case, agents start to distort their advice in the long run, but the public belief can converge to the true state. Then if the value from the outside option for the principal is low enough to the extent that effective communication is sustained even when the principal is sure that one project has a higher fundamental quality than another, we have perpetual communication with complete learning. This contrast with the first scenario stems from the fact that a pandering equilibrium can convey information about the absolute levels of the fundamental qualities to the successive periods. We also show that this property leads to a possibility of informational cascade.

We also show that, depending on the realizations of idiosyncratic shocks in early periods, at some stage a principal may ignore her agent's advice and act only on the information obtained from previous decisions. This phenomenon, informational cascade, occurs when it is optimal for the principal, having observed the choices of those ahead of her, to follow the behavior of the preceding principal without regards to her agent's advice. Once this stage is reached, her decision is uninfluential to principals in the later periods. Therefore, the next principal draws the same inference from the history of past decisions, as do all later principals. In such a case, effective communication breaks down completely, and it incurs an efficiency loss when both projects have a high fundamental quality.

This chapter brings together two strands of literature. Che, Dessein, and Kartik (2011) is the base model for the current study. They show that the pandering distortion in communication arises when there are alternative actions with different potentials but there is a non-trivial conflict of interest over an outside option. Chakraborty and Harbaugh (2007) is also relevant in examining a model in which comparative statement can be credible across different dimensions. These papers are part of the literature on "cheap-talk games" that was initiated by a very influential paper of Crawford and Sobel (1982), which studies the credibility of cheap talk when there are conflicts of interest between an adviser and a receiver.

The second strand is the literature on observational learning that starts with Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992). They show the possibility of an informational cascade when it is optimal for an individual, having observed the actions of those ahead of him, to follow the behavior of the preceding individual without regards to his own information. Smith and Sorensen (2000) study the observational learning process in a more general model. Lee (1993) characterizes the necessary and sufficient condition for the occurrence of a fully revealing informational cascade, which is defined as an event in which the sequence converges to a limit that is optimal under the true state. Our study also characterizes the conditions under which the public belief about fundamental qualities of projects converges, although there is an idiosyncratic shock in each period that the principal cannot learn from predecessors' actions.

In our model, a new agent-principal pair is formed each period and, hence, they do not care about the future period's choices; all they only care about the decision made in their period. This distinguishes the model from literature on reputation or career concerns in adviser-advisee relationships (e.g., Scharfstein and Stein (1990), Prendergast and Stole (1996), Ottaviani and Sorensen (2006), and Dasgupta and Prat (2008)). Hence, in our model, an equilibrium path is solved forwardly, and the focus is how public belief, that summarizes the past choices, evolves over time and its long run consequences.

Brandenberger and Polak (1996) show that a firm that cares about the market's posterior expectation of its profit will distort investment decision. They also consider a dynamic extension where each firm makes decision sequentially where firms in later periods may learn some information from predecessors' distorted decisions, like in our model. Their study, however, like many applications of herding models, is not one of strategic communication, but rather has an agent making decision himself.

The chapter continues as follows. The following section develops our dynamic extension of Che, Dessein, and Kartik (2011). Section 3.3 characterizes the asymptotic properties of the equilibrium. Section 3.4 presents some extensions of our baseline model. All the proofs are given in the Appendix.

3.2. Model

There is a sequence of periods, $t = 1, 2, ..., T$ (T can be finite of infinite). In each period, there is a distinct principal-agent pair, who care only about the decision made at the period. The principal at each period must make a choice from a set $\{0, 1, 2\}$. It is convenient to interpret option 0 as a status quo or outside option for the principal. Both players enjoy a common payoff if one of the alternative projects is chosen, but this value is private information of the agent. The payoff to both when the principal chooses project $i \in \{1, 2\}$ is b_i^t; the status quo action 0 gives v_0 to principal (for simplicity does not vary over time) and 0 to the agent. Assume that $b_i^t = v_i + u_i^t$, where u_i^t (idiosyncratic shocks) is independently and identically drawn across projects and time from some distribution with non-negative support, and the v_i (fundamental qualities) is drawn at time 0 i.i.d. from some distribution with non-negative support. We assume that the agent at period t privately observes v_i and u_i^t. The project chosen at time t is denoted $a_t \in \{0, 1, 2\}$. At the beginning of period t, the only history observed by the principal-agent pair in that period is $a^t = (a_1, a_2, ..., a_t)$. Finally, there is no verifiable information to observe/reveal, all information is soft. To make the setting very comparable to the leading static example of Che, Dessein, and Kartik (2011), CDK hereafter, suppose that $Support[v_i] = \{0, v\}$ with prior $\pi := \Pr(v_i = v) \in [0,1]$, and u_i^t is drawn uniformly on $[0,1]$. We assume that $v < 1$, which implies that idiosyncratic shocks can overturn the difference in the fundamental qualities. In the following, we sometimes abbreviate the superscript or subscript in u_i^t and b_i^t, when it does not cause confusion.

Remark 3. *Note that by assumption, the history does not add any useful information for an agent-principal pair, since the agent knows (b_1^t, b_2^t). This distinguishes our model from the standard herding literature (Bikhchandani, Hirshleifer, and Welch/Benergee). Moreover, agents have no future concerns here and are not trying to signal ability, which distinguishes from the reputation or career concerns literature.*

Now we describe the strategies of players. The agent at period t's strategy is a function $\alpha^t : \{0,1,2\}^{t-1} \times [0, 1+v]^2 \to [0,1]$ such that it represents the probability that he recommends project one, given history a^{t-1} and a pair of values of projects b_1^t and b_2^t. Note that he never recommends the outside option and, hence, $1 - \alpha^t$ is the probability that he recommends project two. The principal at period t's strategy is a pair of functions q_i^t:$\{0,1,2\}^{t-1} \to [0,1]$ for $i \in \{1,2\}$ such that $q_i^t\left(a^{t-1}\right)$ is the probability that she accepts recommendation i given history

a^{t-1}. With probability $1 - q_i^t(a^{t-1})$ she chooses the outside option.[1] We denote (q_1^t, q_2^t) by $\mathbf{q^t}$. The Bayes Nash equilibrium in each period is defined in an obvious way.[2]

We define some important terms. We denote (b_1^t, b_2^t) by $\mathbf{b^t}$. If $q_j^t(a^{t-1}) = 1$, we say that the principal **rubber-stamps** project j, since she chooses it with probability one when the agent recommends it. We call an equilibrium play at particular period t **truthful given** a^{t-1}, if agents recommends the better project given a^{t-1}, i.e., $\alpha^t = 1$ if $b_1^t \geq b_2^t$ and $\alpha^t = 0$ if $b_1^t < b_2^t$. It is important to emphasize that truthful here is only in the sense of rankings, not in the sense that the agent fully reveals the cardinal values of the projects. An equilibrium play at period t is **influential given** a^{t-1}, if $\min\{q_1^t(a^{t-1}), q_2^t(a^{t-1})\} > 0$. Notice that if an equilibrium play at period t is truthful and influential given a^{t-1}, it must be the case that the principal rubber-stamps both projects.

We maintain the following assumptions on parameter specifications throughout the following:

(A1) $v_0 \in (\frac{v}{3} + \frac{2}{3}, v + \frac{2}{3})$.

(A2) $\mathbb{E}[b_i^1 | b_i^1 \geq b_j^1] > v_0$.

---

[1] In this definition, we exclude the possibility that the principal mixes between accepting projects 1 and 2, after a recommendation. This is actually without loss of generality; see Che, Dessein, and Kartik (2011).

[2] See, for example, Fudenberg and Tirole (1991).

Assumption (A1) is for two purposes. First, it ensures $v_0 \in (0, 1+v)$, which implies that a project has a positive chance of being better and worse for the principal than the outside option; this is without loss of generality because otherwise a project or outside option would not be viable. More importantly, it also implies that the agent strictly prefers any project to the outside option, whereas with positive probability, each project is worse than the outside option for principal. Thus the conflict of interest is entirely about the outside option: agent does not internalize the opportunity cost to the principal of implementing a project.

Secondly, it excludes some trivial cases from the analysis. Actually, we can prove the following claim that demonstrates that when the value of the outside option for the principal is too high, there is no influential equilibrium, while in the opposite case there is a truthful and influential equilibrium even when the principal knows that one project has higher fundamental quality than another project.

Claim 5. *1. If $v_0 \geq v + \frac{2}{3}$, there is no influential period in an equilibrium path.*

2. If $v_0 \leq \frac{v}{3} + \frac{2}{3}$, there is a truthful and influential equilibrium when it is common knowledge that $v_i = v$ and $v_j = 0$.

To see Claim 5-1, note that $v + \frac{2}{3} = \mathbb{E}[u_i + v | u_i + v \geq u_j + v]$. When the value of the outside option to the principal is too high, relative to this value, there is no way to have an influential equilibrium; it is always optimal for the principal to choose outside option. Similarly, to see Claim 5-2, note that $\frac{v}{3} + \frac{2}{3} = \mathbb{E}[u_i | u_i \geq$

$u_j+v]$.[3] Because $\mathbb{E}[u_i+v|u_i+v\geq u_j]>\mathbb{E}[u_i|u_i\geq u_j+v]$, if $v_0\leq \frac{v}{3}+\frac{2}{3}$ it is always optimal for the principal to rubber-stamp the agent's recommendation, even when she knows that $v_i>v_j$. The assumption (A1) excludes these trivial cases.

In order to understand the role of (A2), think about the initial period. Because both projects have the same prior, at the initial period in an influential equilibrium (given an empty history), the agent just recommends the better project. Thus when the agent recommends project i, project i's expected payoff is

$$\mathbb{E}[b_i^1|b_i^1\geq b_j^1]=\pi^2\mathbb{E}[u_i+v|u_i\geq u_j] \tag{3.1}$$

$$+2\pi\left(1-\pi\right)g\left(v\right)\mathbb{E}[u_i+v|u_i+v\geq u_j]$$

$$+2\pi\left(1-\pi\right)\left(1-g\left(v\right)\right)\mathbb{E}[u_i|u_i\geq u_j+v]+\left(1-\pi\right)^2\mathbb{E}[u_i|u_i\geq u_j],$$

where $g\left(v\right)=\Pr[u_i\geq u_j-v]=-\frac{1}{2}v^2+v+\frac{1}{2}$. Hence if π is sufficiently large, it is optimal for the principal to accept project i, i.e., (3.1) becomes higher than v_0, while if π is small, then (3.1) becomes smaller than v_0 and we cannot have an influential equilibrium. Assumption (A2) implies that π is sufficiently large so that we have the truthful equilibrium at period one and the principal rubber-stamps

[3] In order for $\mathbb{E}[u_i|u_i\geq u_j+v]$ to be defined, it must hold that $v\leq 1$.

the agent's recommendation. More generally, it can be shown that there is no influential equilibrium if (A2) is violated.

We further examine the properties of the equilibrium inductively. Although we always have a babbling equilibrium (the equilibrium of no communication), we assume that players always play the influential equilibrium, which is unique if any at each period.

In the following analysis, the pair of real numbers $\mathbf{s}^t = (s^t_{vv}, s^t_{v0}, s^t_{0v}, s^t_{00}) \in [0,1]^4$ is seen as a state variable, where s^t_{xy} represents the probability that public belief $\mathbf{s}^t$ puts to the event $v_1 = x$ and $v_2 = y$, where $x \in \{0, v\}$ and $y \in \{0, v\}$, in the beginning of period t. It is generated by Bayes rule. We have $s^t_{vv} + s^t_{v0} + s^t_{0v} + s^t_{00} = 1$ from the definition. Given public belief $\mathbf{s}^t$, we refer to project $j(i)$ as the **conditionally better-looking (worse-looking) project** if $\mathbb{E}[b^t_j | b^t_j \geq b^t_i, \mathbf{s}^t] > \mathbb{E}[b^t_i | b^t_i \geq b^t_j, \mathbf{s}^t]$. Because the equilibrium in each period is characterized by $\mathbf{s}^t$, we focus on it instead of observed history a^{t-1}.

We characterize an equilibrium path forwardly. We have initial public belief as follows:

$$(s^1_{vv}, s^1_{v0}, s^1_{0v}, s^t_{00}) = (\pi^2, \pi\left(1-\pi\right), \pi\left(1-\pi\right), \left(1-\pi\right)^2).$$

Without loss of generality, suppose that $b^1_1 > b^1_2$ is realized and hence project one is proposed and accepted in the initial period, i.e., $a_1 = 1$. Then the second period's state, the principal's belief at the beginning of the second period, is calculated as follows:

$$(s^2_{vv}, s^2_{v0}, s^2_{0v}, s^2_{00}) = (\pi^2, 2\pi\,(1-\pi)\,g\,(v)\,, 2\pi\,(1-\pi)\,(1-g\,(v))\,, (1-\pi)^2). \tag{3.2}$$

Note that we have $s^2_{v0} > s^2_{0v}$ and hence project one is more likely to have superior fundamental quality than project two from the second period's principal's point-of-view.

In the second period, as long as the agent is using the strategy of recommending the better project, if project one is proposed again project one's expected payoff for the principal is

$$\frac{\mathbf{s}^2 \cdot (\frac{1}{2}(v+\frac{2}{3}), g\,(v)\,\mathbb{E}[u_i + v|u_i + v \geq u_j], (1-g\,(v))\,\mathbb{E}[u_i|u_i \geq u_j + v], \frac{1}{2}\cdot\frac{2}{3})}{\frac{1}{2}s^2_{vv} + s^2_{v0}g\,(v) + s^2_{0v}\,(1-g\,(v)) + \frac{1}{2}s^2_{00}}.$$

On the other hand, if project two is proposed, project two's expected payoff for the principal is as follows:

$$\frac{\mathbf{s}^2 \cdot (\frac{1}{2}(v+\frac{2}{3}), (1-g\,(v))\,\mathbb{E}[u_i + v|u_i + v \geq u_j], g\,(v)\,\mathbb{E}[u_i|u_i \geq u_j + v], \frac{1}{2}\cdot\frac{2}{3})}{\frac{1}{2}s^2_{vv} + s^2_{v0}\,(1-g\,(v)) + s^2_{0v}g\,(v) + \frac{1}{2}s^2_{00}}.$$

Hence the principal at period two rubber-stamps the agent's recommendation if both are higher than the value of outside option to principal, v_0.

More generally, suppose that the players have kept playing the truthful equilibrium up until period $\tau - 1$, and history $a^{\tau-1}$ includes $m-1$ number of i and

$\tau - m$ number of j. Then, at period τ, also suppose that the agent takes the strategy of recommending the better project. Then if project i is proposed, expected payoff of project i is calculated as

$$\frac{\left(\frac{1}{2}\right)^{\tau}\pi^{2}(v+\frac{2}{3})+\pi\left(1-\pi\right)g\left(v\right)^{m}\left(1-g\left(v\right)\right)^{\tau-m}\mathbb{E}[u_{i}+v|u_{i}+v\geq u_{j}]}{\Xi}$$
$$+\frac{\pi\left(1-\pi\right)g\left(v\right)^{\tau-m}\left(1-g\left(v\right)\right)^{m}\mathbb{E}[u_{i}|u_{i}\geq u_{j}+v]\}+\left(\frac{1}{2}\right)^{\tau}\left(1-\pi\right)^{2}\frac{2}{3}}{\Xi},$$

where

$$\begin{aligned}\Xi &= \left(\frac{1}{2}\right)^{\tau}\pi^{2}+\pi\left(1-\pi\right)\left\{g\left(v\right)^{m}\left(1-g\left(v\right)\right)^{\tau-m}+g\left(v\right)^{m}\left(1-g\left(v\right)\right)^{\tau-m}\right\}\\ &\quad+\left(\frac{1}{2}\right)^{\tau}\left(1-\pi\right)^{2}.\end{aligned}$$

We denote this expression by $\varphi\left(\tau,m\right)$. Some properties follow from the definition.[4]

Claim 6. *1.* $\varphi\left(t,m\right)>\varphi\left(t,n\right)$ *for all* $m>n$.

2. $\varphi\left(t,m\right)>\varphi\left(s,m\right)$ *for all* $s>t$.

Think about function N_t from a set of histories to positive integers such that $N_t\left(a^{\tau}\right)$ is the number of times project 1 is contained in the first t elements of

[4]We cannot ensure that $\varphi\left(t+1,m+1\right)>\varphi\left(t,m\right)$ for all t and m. The issue is that recommending a particular project more frequently than another makes a principal believe that two projects have the different fundamentals, which may decrease the expected value of the better looking project.

a^{τ}. For example, if $a^4 = (1,1,2,1)$, $N_3(a^4) = 2$. Then the discussion so far is summarized in the following proposition.

Proposition 6. *Given history $a^{\tau-1}$, the agent at period τ recommends the better project if and only if*

$$\min\{\varphi\left(\tau, N_{\tau-1}\left(a^{\tau-1}\right)+1\right), \varphi\left(\tau, \tau - N_t\left(a^{\tau-1}\right)+1\right)\} \geq v_0. \tag{3.3}$$

This result says that players play the truthful equilibrium up until condition (3.3) is first violated. Let t be the first such period. If both $\varphi(t, N_{t-1}(a^{t-1})+1)$ and $\varphi(t, t - N_{t-1}(a^{t-1})+1)$ become smaller than the value of the outside option v_0, we have no way to make an influential equilibrium at period t. Now think of the case in which

$$\varphi\left(t, N_{t-1}\left(a^{t-1}\right)+1\right) > v_0 \text{ and } \varphi\left(t, t - N_{t-1}\left(a^{t-1}\right)+1\right) < v_0. \tag{3.4}$$

Note that in such a case, project one has been recommended more often than project two, i.e., $N_{t-1}(a^{t-1}) > t - N_{t-1}(a^{t-1})$. From the definition of φ, then we have $\mathbb{E}[b_1 | b_1 \geq b_2, \mathbf{s}^t] > v_0$ and $\mathbb{E}[b_2 | b_2 \geq b_1, \mathbf{s}^t] < v_0$ hold and we may expect to have CDK's "pandering equilibrium." The idea is that the agent recommends project one (the better-looking project) whenever $b_1^t > b_2^t$, but he recommends project two (the worse-looking project) only when it is sufficiently better than project one and, thereby, increases the principal's posterior about b_2^t when he does in fact recommend project two. Thus, pandering toward project one makes a recommendation of project two more acceptable. In turn, the principal must be more likely to follow a recommendation of project one than a recommendation of project two.

Remark 4. *It is important to note that even when the outside option is chosen, players from subsequent periods can infer what project is recommended; if the outside option is chosen at a period, it conveys information that the worse-looking project was recommended at that period. This means that we have the same outcome if we work on the setting in which history consists of advises given, instead of the projects implemented.*

In sum, the influential equilibrium, if any, in period t when (3.4) holds is characterized by a probability that a recommendation of project two is accepted, i.e., $q_2^t \in (0,1)$. The agent proposes project one if and only if $q_2^t < b_1^t/b_2^t$ and proposes project two if and only if $q_2^t \geq b_1^t/b_2^t$. The principal accepts a recommendation of project one with probability one, $q_1^t = 1$, and accepts a recommendation of project two with probability q_2^t. The key of the construction of the equilibrium is that the agent is indifferent between recommending project one and two when $b_1^t/b_2^t = q_2^t$, and also

$$\mathbb{E}[b_1 | b_1 \geq q_2^t b_2, \mathbf{s}^t] > v_0 \text{ and } \mathbb{E}[b_2 | b_2 \geq b_1/q_2^t, \mathbf{s}^t] = v_0, \tag{3.5}$$

and, hence, the principal accepts the recommendation of project one and she is indifferent between project two and the outside option when project two is recommended. In the following, we call this type of equilibrium "a pandering equilibrium." Note that the existence of such q_2^t is not automatically ensured. Analogous argument shows that we may have a pandering equilibrium with $q_1^t < 1$ and $q_2^t = 1$ when $\mathbb{E}[b_1|b_1 \geq b_2, \mathbf{s}^t] < v_0$ and $\mathbb{E}[b_2|b_2 \geq b_1, \mathbf{s}^t] > v_0$.

Definition 5. *Given $\mathbf{s}^t$, an equilibrium at period t is a pandering equilibrium if there is a number $q \neq 1$ such that the agent recommends project one if $b_1^t \geq qb_2^t$ and recommends project two if $qb_2^t \geq b_1^t$. Moreover, if $q < 1$, the principal rubber-stamps project one, and accepts project two with probability q. If $q > 1$, the principal rubber-stamps project two and accepts project one with probability $1/q$.*

In a pandering equilibrium, if $q < 1$ project one is accepted with certainty and project two is accepted with probability q, while if $q > 1$, project two is accepted with certainty and project one is accepted with probability $1/q$. It is convenient for the subsequent analysis to divide the space of public belief into different regions. We define subset S_T of the space of public belief by

$$\{\mathbf{s}^t|\mathbb{E}[b_1|b_1 \geq b_2, \mathbf{s}^t] > v_0 \text{ and } \mathbb{E}[b_2|b_2 \geq b_1, \mathbf{s}^t] > v_0\},$$

and call it the truthful and influential equilibrium phase, or the truthful equilibrium phase for short. Generally s_{vv} should be high, relative to s_{v0}, s_{0v}, and s_{00}, in the truthful equilibrium phase S_T. Also, define subset S_P, which we call the pandering equilibrium phase, by

$$\{\mathbf{s}^t|\mathbb{E}[b_1|b_1 \geq qb_2, \mathbf{s}^t] > v_0 \text{ and } \mathbb{E}[b_2|b_2 \geq b_1/q, \mathbf{s}^t] = v_0 \text{ for some } q < 1$$

$$\text{or } \mathbb{E}[b_1|b_1 \geq qb_2, \mathbf{s}^t] = v_0 \text{ and } \mathbb{E}[b_2|b_2 \geq b_1/q, \mathbf{s}^t] > v_0 \text{ for some } q > 1\}.$$

Obviously, in the pandering equilibrium phase either s_{v0} or s_{0v} should be high, relative to s_{vv} and s_{00}. Those sets are determined by parameter specifications of (v, v_0, π), but we always have $S_T \cap S_P = \varnothing$ (see Claim 10 in the Appendix). When public belief is in S_T, players play the truthful equilibrium, while when it is in S_P, a pandering equilibrium is played. An important observation is that once an equilibrium path, which starts from S_T, goes out from $S_T \cup S_P$, communication breaks down and the uninfluential equilibrium is played perpetually.

Remark 5. *An important observation is that a pandering equilibrium incurs efficiency loss compared to the truthful equilibrium. Obviously, the most desirable behavior of the agent for the principal is that he recommends project one if and only if $b_1 \geq b_2$. It is computed that*

$$\Pr[b_1 \geq qb_2|\mathbf{s}]\mathbb{E}[b_1|b_1 \geq qb_2, \mathbf{s}] + \Pr[qb_2 \geq u_1|\mathbf{s}]\mathbb{E}[b_2|qb_2 \geq b_1, \mathbf{s}]$$

$$> \Pr[b_1 \geq q'b_2|\mathbf{s}]\mathbb{E}[b_1|b_1 \geq q'b_2, \mathbf{s}] + \Pr[q'b_2 \geq b_1|\mathbf{s}]\mathbb{E}[b_2|q'b_2 \geq b_1, \mathbf{s}],$$

for all $q' < q < 1$ and $\mathbf{s}$, which stems from the fact that a pandering equilibrium distorts the recommendation. In this sense q, or inverse of q may be interpreted *as the degree of distortion in advice. This fact makes the following interesting phenomenon happen. Suppose that* $\mathbb{E}[b_j|b_j \geq b_i] - \varepsilon < v_0$ *for a very small* $\varepsilon > 0$*. Therefore, (A2) is barely satisfied and truthful equilibrium exists at the initial period. Then, for a sufficiently small* ε*, there is no truthful equilibrium in the second period. Without loss of generality, suppose that that project one is recommended at the initial period. Note that* $\mathbf{s}^2$ *does not depend on* v_0*. However, the degree of distortion to make* $\mathbb{E}[b_2|qb_1 \geq b_2] = v_0$*, which is the necessary condition for pandering equilibrium, is so high that we may have*

$$\Pr[b_1 \geq qb_2|\mathbf{s}^2]\mathbb{E}[b_1|b_1 \geq qb_2, \mathbf{s}^2] + \Pr[qb_2 \geq b_1|\mathbf{s}^2]\mathbb{E}[b_2|qb_2 \geq b_1, \mathbf{s}^2] < v_0$$

for such q*. In this case we cannot even have a pandering equilibrium. In such a case, communication breaks down at the second period with probability one.*

3.3. Asymptotic Properties of Equilibrium

In this section, we examine the asymptotic properties of the equilibrium. We characterize conditions under which informational cascade may occur or ever-lasting communication is sustained. In the following analysis, we denote by $\mathbf{s}_=$ the public belief that attaches probability one to the event that the two projects have the same fundamental quality, and the relative likelihood of the remaining two events is the same as the initial prior: $\mathbf{s}_= = \left(\frac{\pi^2}{\pi^2+(1-\pi)^2}, 0, 0, \frac{(1-\pi)^2}{\pi^2+(1-\pi)^2}\right)$.

Before examining the asymptotic properties in detail, we give an immediate consequence as Proposition 7. It demonstrates that effective communication cannot be sustained perpetually by totally incorrect public belief, because as long as effective communication is sustained, the relative frequency of recommendations between the two projects at least partially corrects the belief.

Proposition 7. *1. If $v_1 \neq v_2$ is realized, the truthful equilibrium is never played in a sufficiently long-run, i.e., almost surely, there is τ such that $\mathbf{s}^t \notin S_T$ for all $t \geq \tau$.*

2. If $v_1 = v_2$ is realized (either $v_1 = v_2 = v$ or $v_1 = v_2 = 0$), the pandering equilibrium is never played in a sufficiently long-run, i.e., almost surely, there is τ such that $\mathbf{s}^t \notin S_P$ for all $t \geq \tau$.

Note that the threshold τ in the statement of proposition depends on an equilibrium history. Some equilibrium history may take more time than others to perpetually leave a particular phase. Also, note that $\mathbf{s}^t \notin S_T$ ($\mathbf{s}^t \notin S_P$) for all future t does not mean that $\mathbf{s}^t \in S_P$ ($\mathbf{s}^t \in S_P$) holds: it may be the case that communication breaks down perpetually.

We first examine the possibility that communication continues perpetually. Remark 5 in Section 3.2 shows the possibility that communication breaks down with certainty. However, once we impose a certain condition, it is possible that the influential equilibrium is played perpetually.

Theorem 19. *Suppose that parameter specifications are such that* $\mathbf{s}_{=} \in S_T$ *and* $\varphi(2,0) > v_0$. *If* $v_1 = v_2$, *with a strictly positive probability, the truthful equilibrium is played perpetually, i.e.,* $\mathbf{s}^t \in S_T$ *for all* $t \geq \tau$ *for some* τ.

Hence the result shows that informational cascade does not occur with probability one. Basically, the condition $\varphi(2,0) > v_0$ excludes the possibility that an equilibrium path goes out from the truthful equilibrium phase at period 2 (hence the situation described in Remark 5 is excluded). This allows an equilibrium path to stay in the truthful equilibrium phase for the initial few periods and make s_{vv} and s_{00} sufficiently high relative to s_{v0} and s_{0v}. This makes the truthful equilibrium phase an absorbing phase when $v_1 = v_2$.

The question if an equilibrium can perpetually stay in the truthful equilibrium phase has a similar flavor as the *Gambler's ruin problem* where a gambler's objective is to become infinitely rich. In our game, once the principal holds the strong belief that the two projects do not share the same fundamental, i.e., s_{vv} and s_{00} become sufficiently small relative to s_{v0} or s_{0v}, the equilibrium path goes out from the truthful equilibrium phase S_T. This situation corresponds to *ruin* in the gambler's problem. It is known that as long as the odds of the gamble are in his favor, the gambler is not ruined with probability one. If $v_1 = v_2$, the random process of the agents' belief is biased toward the truthful equilibrium phase and, hence, we have a strictly positive probability of having the right convergence of belief without the state going out from S_T, as an analog of the gambler accumulating an infinite fortune, when the odds of the gamble are in his favor.

An important observation is that an equilibrium may perpetually stay on the truthful equilibrium phase with a strictly positive probability even if both projects have low fundamental quality, i.e., $v_1 = v_2 = 0$. This happens because once the equilibrium starts to stay on the the truthful equilibrium phase, public belief ratio s_{vv}/s_{00} cannot be updated. In such a case, although herding to a particular project is not observed, the equilibrium path is only sustained due to incomplete process of information updating. Hence we have *perpetual communication with incomplete learning.* Note that even if principals cannot completely learn the values of fundamental qualities, the first best outcome for them (in a sense of ex-ante from the beginning of each stage) is attained in this case. This observation gives the next proposition.

Proposition 8. *When $v_1 = v_2$, fully correct learning about fundamental qualities (for principal) is impossible. In the long run, when $v_1 = v_2 = v$, the (ex-ante) first best outcome for the principal is attained by perpetual communication with incomplete learning. On the other hand, when $v_1 = v_2 = 0$, the (ex-ante) first best outcome for the principal is attained by correct cascade.*

Theorem 19 leaves open the question of what the asymptotic behavior of the equilibrium is when the two projects have different fundamental qualities, i.e., $v_1 \neq v_2$. We will show that the following condition is crucial for the question:

Condition 1. *Parameter specifications are such that* $(0,1,0,0) \in S_P$ *and* $(0,0,1,0) \in S_P$*, i.e., the existence of pandering equilibrium in the one shot game is ensured when it is common knowledge that* $v_i = v$ *and* $v_j = 0$*.*

Remember that $(0,1,0,0)$ represents the belief that puts the whole mass on the event that only project one has high fundamental quality. Condition 1 requires that even when the principal is completely sure that a project has a superior fundamental quality than another, we have an influential equilibrium. Note first that it requires $v_0 < 1$,[5] because otherwise, the worse-looking project can never be chosen over the outside option. Also, this makes it possible for idiosyncratic shocks to overturn the difference in fundamental qualities between the two projects. When it is common knowledge that $v_1 = v$ and $v_2 = 0$, the condition $\mathbb{E}[b_2|q_2^* b_2 \geq b_1] = v_0$, which determines the acceptance probability of the worse-looking project in a pandering equilibrium, implies $q_2^* = \frac{v}{3v_0-2}$. Then the existence of a pandering equilibrium is ensured if $\mathbb{E}\left[b_1|b_1 \geq \frac{v}{3v_0-2}b_2\right] \geq v_0$. Because $\mathbb{E}\left[b_1|b_1 \geq \frac{v}{3v_0-2}b_2\right]$ is a non-increasing function of v_0, there is a unique v_0^* such that $\mathbb{E}\left[b_1|b_1 \geq \frac{v}{3v_0^*-2}b_2 = v_0^*\right]$. Then a pandering equilibrium exists if and only if $v_0 < \min\{v_0^*, 1\}$. Under the assumption $v_0 < 1$, the outside option is chosen if Condition 1 does not hold.

Now we have the following results, which characterize the asymptotic behavior of the equilibrium when $v_1 \neq v_2$ is realized.

[5]This is satisfied by the assumption $v < 1$ and (A1).

Theorem 20. *1. Suppose that Condition 1 holds. If $v_1 = v$ and $v_2 = 0$ is realized initially (if $v_2 = v$ and $v_1 = 0$), public belief converges to the true state, i.e., $s^t_{v0} \to 1$ ($s^t_{0v} \to 1$) with a strictly positive probability. In such a case, a pandering equilibrium is played perpetually, i.e., $\mathbf{s}^t \in S_P$ for all $t \geq \tau$ for some τ.*

2. Suppose that Condition 1 does not hold. If $v_1 = v$ and $v_2 = 0$ is realized initially (or $v_2 = v$ and $v_1 = 0$), communication breaks down within a finite period of time, i.e., $\mathbf{s}^t \notin S_T \cup S_P$ for some t, almost surely, and the outside option is chosen perpetually.

Theorem 20 holds because in the pandering equilibrium phase, so long as effective communication is sustained, we have complete learning asymptotically. Hence if Condition 1 does not hold, public belief, which converges to the true state, eventually leaves S_P and communication breaks down. On the other hand, if Condition 1 is satisfied, it can be shown that there are equilibrium histories that sufficiently convince the public belief about the true state without leaving $S_T \cup S_P$ even once. Because the existence of a pandering equilibrium is ensured in the limit of the public belief, $(0, 1, 0, 0) \in S_P$, we can sustain communication perpetually.

We next show the possibility of informational cascade: depending on the realizations of initial periods' idiosyncratic shocks, communication may break down even when the true state (fundamental qualities of projects) allows players to gain mutual benefit through communication. We have two types of informational cascade. In one type of cascade, the outside option is chosen perpetually, without having effective communication: this happens when public belief is formed in a way that both projects have low fundamental quality. Another possibility is that public belief is formed in a way that the two project have different fundamental qualities. In such a case, depending on parameter specifications, the outside option may be chosen perpetually. Theorem 21 actually shows that informational cascade occurs with a strictly positive probability after any realizations of fundamental qualities.

Theorem 21. *For all realizations of (v_1, v_2), with a strictly positive probability, communication breaks down within a finite period of time, i.e., $\mathbf{s}^t \notin S_T \cup S_P$ for some t.*

Theorem 21 is easily seen if $\mathbf{s}_= \notin S_T$. In such a case, if an equilibrium path stays in the truthful equilibrium phase S_T for sufficiently long, public belief converges to $\mathbf{s}_=$ and, hence, at some point, communication breaks down. Hence the question is what if this is not the case. Theorem 21 is proved by showing that, even if $\mathbf{s}_= \in S_T$, there is a finite history that sufficiently convinces the public belief that both projects have low fundamental quality, i.e., $v_1 = v_2 = 0$. It is somewhat a striking result given that agents do not internalize the value from the outside option and hence they never recommend the outside option. However, there is a intertemporal pattern of recommendation that may form the belief. In order to explain this, note the following important observation, which we give as a lemma.

Lemma 13. *Suppose that a pandering equilibrium with $q < 1$ is played at period t. If project one (the better-looking project) is recommended, $s_{vv}^{t+1}/s_{00}^{t+1} > s_{vv}^{t}/s_{00}^{t}$, and if project two (the worse-looking project) is recommended, $s_{vv}^{t+1}/s_{00}^{t+1} < s_{vv}^{t}/s_{00}^{t}$. Analogous results hold for the case $q > 1$.*

As long as the truthful equilibrium is played, the public belief cannot update the relative likelihood between the two events of both projects having high fundamental quality and neither having it: s_{vv}^{t}/s_{00}^{t} stays at the same. On the other hand, in a pandering equilibrium in which, say, project two has a smaller acceptance probability it is easier for project two to make up for it when both have low fundamental quality than when both have high fundamental quality: $\Pr[qu_2 > u_1] > \Pr[q(u_2 + v) > u_1 + v]$ for all $q < 1$. This implies that when the worse-looking project is recommended, the public belief is updated toward convincing the principal that the two projects have low fundamental quality, i.e., increasing the relative likelihood s_{00}^{t}/s_{vv}^{t}.

This observation generates the first type of informational cascade. Suppose that in the observed history, the equilibrium reverts to the truthful equilibrium very quickly because the worse looking project is recommended. Then, the public belief is updated toward s_{00} high. Then by repeating this path for many times (hence the equilibrium path fluctuates between two phases), the condition for the existence of influential equilibrium may be violated.[6] A straightforward observation

[6]The existence of a path that makes s_{00} sufficiently high, not only relative to s_{vv} but also s_{v0} and s_{0v}, is shown in the Appendix, when we prove Theorem 21.

is that this type of informational cascade is the most likely to happen when (A2) is barely satisfied and $v_1 = v_2 = 0$, the situation we may call **Correct Cascade,** but it can also happen with non-zero probability when $v_1 = v_2 = v$, the situation we may call **Incorrect Cascade.**

In order to see the possibility of the second type of informational cascade, suppose that in the initial few periods only one project, say project one, is recommended. Then, the principal's belief is shaped towards increasing s_{v0}, relative to $s_{0v},\ s_{vv}$, and s_{00}. Then obviously, whether communication eventually breaks down depends on if we have an influential equilibrium when $v_1 = 1$ and $v_2 = 0$ is common knowledge. If we have this type of cascade, we may say that the correct cascade occurs if $v_1 = v$ and $v_2 = 0$ while the incorrect cascade occurs if not**.**

3.4. Extensions

Delegating authority to the agent: Our baseline model is based on the assumption that the principal is not able to commit ex-ante to a vector of acceptance probabilities. However, it is also interesting to consider the case in which she can make some form of commitment. Among different commitment mechanisms, perhaps the easiest for the principal is to fully delegate the decision to agent. We examine the consequence of such an extension.

CDK show that in a pandering equilibrium phase, the principal is ex-ante better off by delegating authority to the agent.[7] The intuition is as follows: the principal is

[7] They also characterize the optimal commitment mechanism.

indifferent between a project, say project two, and the outside option in a pandering equilibrium. Holding the agent's strategy fixed, the principal's expected payoff is the same whether she plays the equilibrium strategy or rubber-stamps all the recommendations. Delegation commits the principal to play the latter strategy and also has the benefit of eliminating pandering distortion because the agent will always choose the best project (see Remark 5). Furthermore, full delegation is also ex-ante beneficial even when $\mathbf{s}^t \notin S_T \cup S_P$, so long as $v_0 < \mathbb{E}[\max\{b_1, b_2\}|\mathbf{s}^t]$.

To compare the outcome of the game with full delegation and our original game, suppose that the principal at each period can decide whether she wants to fully delegate the decision to the agent, based on her available information $\mathbf{s}^t$. Obviously, when in the truthful equilibrium phase, $\mathbf{s}^t \in S_T$, she is indifferent between delegating or not, because the outcome is the same. In a pandering equilibrium phase $\mathbf{s}^t \in S_P$, she strictly prefers to delegate. Also, even when $\mathbf{s}^t \notin S_T \cup S_P$, she strictly prefers to delegate so long as $v_0 < \mathbb{E}[\max\{b_1, b_2\}|\mathbf{s}^t]$. Only when these are violated, the principal chooses the outside option and communication breaks down perpetually.

An appropriate counterpart for Condition 1 to think about asymptotics in this setting is as follows:

Condition 2: $\mathbb{E}[\max\{b_1, b_2\}|v_1 \neq v_2] \geq v_0$.

A very important observation is that now, the public belief ratio s^t_{vv}/s^t_{00} can never be updated on an equilibrium path. This induces the following theorem that characterizes the asymptotic behavior of an equilibrium when full delegation is possible.

Theorem 22. *Suppose that full delegation is possible.*

1. If $\mathbf{s}_= \in S_T$ and Condition 2 hold, delegation never stops.

2. If $\mathbf{s}_= \notin S_T$, delegation stops and the outside option is chosen within a finite period with a strictly positive probability. If $v_1 = v_2$ is realized, delegation stops within a finite period for sure.

3. Condition 2 is violated, delegation stops and the outside option is chosen within a finite period with a strictly positive probability. If $v_1 \neq v_2$ is realized, delegation stops within a finite period for sure.

The condition $\mathbf{s}_= \in S_T$ is crucial for determining if the truthful equilibrium can be perpetually sustained. Indeed, as long as an equilibrium path stays in the truthful equilibrium phase, $\lim_{t\to\infty} \mathbb{E}[\max\{b_1, b_2\}|\mathbf{s}^t] \geq v_0$ is equivalent to $\mathbf{s}_= \in S_T$.

It is worth noting that in the long run, the availability of the option of full delegation may hurt the principal in a later period. In the original game in which delegation is not possible, the principal can update the belief s^t_{vv}/s^t_{00} and, hence, may have a chance to appropriately terminate communication by finding out that s^t_{vv}/s^t_{00} is sufficiently small, when both projects actually have a low fundamental quality (although she has a chance of wrongfully terminating communication when both projects have high fundamental quality).

Commitment to not observing the history: A different way for the principal to make a commitment is not to observe the history. Indeed, because a pandering equilibrium distorts the advice, the principal is better-off by not observing the history and, thereby, inducing the truthful equilibrium, if she is likely to play a pandering equilibrium. Suppose that the principal in period t can make a commitment of not observing the history at period 0. Then there is a trade-off. The upside of observing the history is that the principal may know whether she should listen to the agent's advice, in other words, she may know $\mathbf{s}^t \in S_T \cup S_P$ or $\mathbf{s}^t \notin S_T \cup S_P$. On the other hand, by not observing the history, she is able to get the truthful recommendation from the agent, which is beneficial if $\mathbf{s}^t \in S_P$.

To see the consequence of allowing this option for the principal, suppose that the principal each period can make a decision of whether to make a commitment of not observing the history, and agents are informed of their decision.[8] Obviously, the period one principal's decision does not matter at all, and hence we regard her as making the commitment, for expositional convenience in the next theorem. It is seen that the period two principal should make the commitment, or her decision does not matter. To see this, note that if $\varphi(2, 0) > v_0$ and hence the period two's players are supposed to play the truthful equilibrium, the principal's decision of

[8] Actually, even if they are not informed of principals' decision, they can infer these.

making the commitment or not does not matter at all. On the other hand, if $\varphi(2,0) < v_0$, and hence they are supposed to play a pandering equilibrium or non-influential equilibrium (see Remark 5), it is strictly better for the principal to make the commitment. Hence, we also regard her as making the commitment. Although their decisions after period three depend on parameter specifications of the model, we have a general result as follows:

Theorem 23. *If $\mathbf{s}_= \in S_T$ and Condition 1 hold, principals in every period make the commitment.*

To see the result, suppose that principals up to period τ make the commitment. This implies that for the principal at period $\tau+1$, even if she observes the history, she cannot update s_{vv}/s_{00}. Then, as long as $\mathbf{s}_= \in S_T$ and Condition 1 hold, for all $\mathbf{s}^{\tau+1}$ that she can form by observing the history, $\mathbf{s}^{\tau+1} \in S_T \cup S_P$ holds and, hence, we have an influential equilibrium. This implies that she is better-off by not observing the history and playing the truthful equilibrium.

On the other hand, if one of the conditions is not satisfied, by observing the history, the principal in a sufficiently later period can know if $\mathbf{s}^t \in S_T \cup S_P$ or $\mathbf{s}^t \notin S_T \cup S_P$, and hence she may know if she should ask advice from the agent or ignore his advice. To see this, suppose that $\mathbf{s}_= \notin S_T$. Then if she makes the commitment, she will rubber-stamp her agent's advice, which gives her strictly negative payoff if $v_1 = v_2 = 0$. This may be avoided by observing the history, and hence she may not want to make the commitment. Characterization of principals' incentives when one of the conditions is violated is left for future research.

Partially internalizing preference on outside option: An important extension of our baseline model is to allow the agent to internalize the principal's preference for the outside option. A simple way to introduce such an idea is to assume that with some probability, an agent is a type who partially internalizes the principal's value for the outside option. Specifically, we think of the case where with some probability p, an agent's payoff from the outside option is κv_0, where $\kappa \in (0, 1)$ determines the degree to which the agent internalizes the value of outside option.[9]

This modification does not cause any qualitative change in asymptotic behavior of the equilibrium, except for the possibility of perpetual communication with incomplete learning. If we assume that with a very small probability an agent partially internalizes the principal's value from the outside option, in a sufficiently long run, the principal can find out if both projects have good fundamental quality, which is information that she can never learn in the original game in the truthful equilibrium phase.

[9]It does not matter if agent's type is private information or not.

Theorem 24. *Suppose that with probability p an agent partially internalizes the value of outside option to a degree of κ. If $v_1 = v_2 = 0$, then $\mathbf{s}^t \notin S_T \cup S_P$ for some t almost surely.*

General distributions for fundamental quality: We discuss the generalization of our results to the case in which projects' fundamental qualities are distributed in continuous ways, rather than the discrete ways as in our original setup. For this purpose, let F be a symmetric (cumulative) and absolutely continuos distribution function for fundamental qualities of projects with support $[\underline{v}_0, \overline{v}_0]^2$, and G be a symmetric (cumulative) and absolutely continuous distribution function for idiosyncratic shocks with support $[\underline{u}, \overline{u}]^2$. Fundamental qualities and idiosyncratic shocks are independent of each other. For simplicity, we impose the assumption that $\overline{u} - \underline{u} \geq \overline{v}_0 - \underline{v}_0$ and, hence, idiosyncratic shocks can overturn the difference in fundamental qualities. The maintained assumption (A1) in our specific model should be modified, and we also need additional assumptions. These are provided as follows:

$(A1')$ $v_0 \in (\mathbb{E}[u_j + \underline{v}_0 | u_j + \underline{v}_0 \geq u_i + \overline{v}_0], \mathbb{E}[u_j + \overline{v}_0 | u_j \geq u_i])$.

$(A3)$ Given (v_j, v_i), $\mathbb{E}[u_j | u_j + v_j > \alpha (b_i + v_i)]$ is non-decreasing in α.[10]

$(A4)$ Given v_i, $\mathbb{E}[u_j + v_j | u_j + v_j \geq u_i + v_i]$ is increasing in $v_j \in [\underline{v}_0, \overline{v}_0]$.

[10]Note that it is only defined for $\alpha \leq \frac{\overline{u} + v_j}{\underline{u} + v_i}$.

The condition (A1′) is analogous to (A1) in Section 3.2. Conditions (A3) and (A4) are automatically satisfied in our original setup developed in the previous section.[11] We maintain (A2) as well. In this setting, the appropriate state variable (public belief) $\mathbf{s}^t$ is now regarded as a probability density function over $[\underline{v}_0, \overline{v}_0]^2$, parameterized by the history a^{t-1}. As in the previous section, we can define S_T as the truthful equilibrium phase, and S_P as the pandering equilibrium phase in analogous ways. From (A2) we have $F \in S_T$. With a slight abuse of notation, we identify a point (v_1, v_2) with a distribution function that puts the entire mass on the point.

Now let v_L be the threshold value of v such that $\mathbb{E}[u_j + v_L | u_j + v_L \geq u_i + \overline{v}_0] = v_0$, which is uniquely determined from (A4). Then, it is easy to see that for all public beliefs $\mathbf{s}$ such that *Support* $[\mathbf{s}_1] \subset [\underline{v}_0, v_L]$ or *Support* $[\mathbf{s}_2] \subset [\underline{v}_0, v_L]$, we have $\mathbf{s} \notin S_T$. Let $\varphi := \overline{v}_0 - v_L$. Then we have the following claim:

Claim 7. *If $|v_i - v_j| \geq \varphi$, the truthful equilibrium is never played in a sufficiently long-run,i.e., there is τ such that $\mathbf{s}^t \notin S_T$ for all $t \geq \tau$, almost surely.*

To see the result, note that if the truthful equilibrium is played perpetually, the relative frequency of recommendations between the two projects tells the principal sufficiently accurate information about $|v_i - v_j|$. From the definition of φ, this implies that $\mathbb{E}[b_i | b_i \geq b_j, \mathbf{s}^t] < v_0$ or $\mathbb{E}[b_j | b_j \geq b_i, \mathbf{s}^t] < v_0$ holds for a sufficiently

[11]For a detailed discussion on this condition, see Che, Dessein, and Kartik (2011).

large t, which contradicts the fact the truthful equilibrium is played perpetually. Hence φ provides an upper bound for the difference in fundamental qualities that allows the truthful communication to be perpetuity sustained.

It is important to note that, in this general distribution case, it is possible to sustain the truthful communication perpetually even when $v_1 \neq v_2$. To see this, think of the case in which $F(\cdot|v_1 = v_2) \in S_T$, and $\varphi(2, 0) > v_0$. As is discussed in the previous section, these conditions make the truthful equilibrium phase an absorbing phase when $v_1 = v_2$. However, actually, it is also an absorbing phase if the difference in fundamental qualities is small. More precisely, because F is absolutely continuous and $F(\cdot|v_1 = v_2) \in S_T$, there is κ such that $F(\cdot|\,|v_i - v_j| < \varepsilon) \in S_T$, or in other words there is κ such that $\mathbb{E}[b_i|b_i \geq b_j, |v_i - v_j| < \varepsilon] > v_0$ and $\mathbb{E}[b_j|b_j \geq b_i, |v_i - v_j| < \varepsilon] > v_0$ for all $\varepsilon < \kappa$. The truthful equilibrium is an absorbing phase if $|v_i - v_j| < \kappa$. This leads the following claim.

Claim 8. *Suppose that* $F(\cdot|v_1 = v_2) \in S_T$, *and* $\varphi(2, 0) > v_0$. *Then, there is* $\lambda > 0$ *such that the truthful equilibrium is played perpetually with a strictly positive probability if* $|v_1 - v_2| \leq \lambda$.

It is also possible to characterize an upper bound for the difference in fundamental qualities that allows the pandering communication to be sustained perpetuity. Towards this end, let v_l be the threshold value of v above which we can find q such that $\mathbb{E}[u_j + \overline{v}_0|u_j + \overline{v}_0 \geq \widehat{q}(u_i + v_l)] \geq v_0$ and $\mathbb{E}[q(u_i + v_l)\,|q(u_i + v_l) \geq u_j + \overline{v}_0] = v_0$, and below which there is not such q. Then, again it is easy to see that for all public belief $\mathbf{s}$ such that *Support* $\mathbf{s}_1 \subset [\underline{v}_0, v_l]$ or *Support* $\mathbf{s}_2 \subset [\underline{v}_0, v_l]$, we have $\mathbf{s} \notin S_L$. More generally, we have the following claim.

Claim 9. *Suppose that* $|v_1 - v_2| > \overline{v}_0 - v_l$. *Then, the pandering equilibrium is never played in a sufficiently long-run, i.e., there is* τ *such that* $\mathbf{s}^t \notin S_P$ *for all* $t \geq \tau$, *almost surely. If, moreover,* $|\overline{v}_0 - v_l| > \varphi$, *communication breaks down within finite a finite period of time, i.e.,* $\mathbf{s}^t \notin S_T \cup S_P$ *for some* t, *almost surely.*

The remaining questions are (i) if the initial realizations of fundamental qualities (v_1, v_2) are such that $(v_1, v_2) \in S_P$, can we sustain the pandering equilibrium perpetually with a strictly positive probability? (ii) if so, can we have fully correct learning? This question turns out to be very subtle. To see this point, in a pandering equilibrium with project two's acceptance probability q, principals in later periods cannot update the relative likelihood of two events that $(v_1, v_2) = (v_1', v_2')$ and $(v_1, v_2) = (v_1', v_2')$ if $v_1' - v_1'' = q\,(v_2' - v_2'')$, because the probability that project two (or one) is recommended is the same between these two events. Hence, a pandering equilibrium cannot convey information about "absolute levels" of fundamental qualities, where the absoluteness is appropriately defined by its acceptance probability q. However, the same information can be conveyed in a pandering equilibrium with different acceptance probability, although it cannot be very informative if the two equilibria are close. Hence the question is if the public belief can accurately obtain the information about the true values of fundamental qualities, which includes the information that cannot be gathered in the pandering equilibrium with the limit acceptance probability q. At this point of time, we give the asymptotic behavior of equilibrium in the case of $(v_1, v_2) \in S_P$ as a conjecture.

Conjecture 1. *Take $(v_1', v_2') \in S_P$. If $v_1 = v_1'$ and $v_2 = v_2'$ are realized initially, public belief converges to the true state, i.e., $\mathbf{s}^t \to (v_1', v_2')$ (note the abuse of notation), with a strictly positive probability. In such a case, the pandering equilibrium is played perpetually, i.e., $\mathbf{s}^t \in S_P$ for all $t \geq \tau$ for some τ.*

3.5. Conclusion

This study has studied a dynamic extension of Che, Dessein, and Kartik (2011)'s cheap-talk model, where a new agent-principal pair is formed each period, and principals in later periods may observe predecessors' actions. The equilibrium paths exhibit different outcomes depending on the idiosyncratic shocks in the early periods. We also discussed some possible extensions of our basic setup.

It may be interesting to extend the model into the case in which the agent is a long-lived player, who meets a new principal in each period. This setting gives an interesting intertemporal trade-off in advice for the agent. Indeed, the agent has an incentive to make principals hold balanced belief about the relative quality of the two projects and, thereby, induce the truthful (and influential) equilibrium. However, it does not necessarily maximize the payoff from the current advice. Also, as is shown in Theorem 21, the agent wants to avoid recommending the worse-looking project in a pandering phase, because it may make both projects' fundamental qualities look worse, which hurts the gain from the future advice. This extension is left for future research.

Another interesting extension would be to generalize the model into a case with many projects. If there are many projects, depending on the realizations of idiosyncratic shocks in initial periods, some projects may be discarded forever. This is because if a project is not recommended in the initial few periods, public belief is formed toward disfavoring the fundamental quality of the project. Hence, after that, it may not be possible to support an equilibrium in a way that the project has a strictly positive probability of acceptance, especially because recommending the worse-looking project shifts the public belief towards disfavoring the absolute levels of other projects' fundamental qualities as well. In such a case, in the long run, we have a perpetual communication in which only a subset of all projects can be recommended.

3.6. Appendix: Proofs

Note that $\mathbb{E}[u_j|qu_j \geq u_i + v] =$

$$\begin{cases} \frac{\int_0^1 \int_{(u_i+v)/q}^1 u_j du_j du_i}{\int_0^1 \int_{(u_i+v)/q}^1 du_j du_i} = \frac{3v^2/q^2+3v/q^2+1/q^2-3}{3/q+6v/q-6} & \text{for } q \in [1+v, \infty], \\ \frac{\int_0^{q_j-v} \int_{(u_i+v)/q_j}^1 u_j du_j du_i}{\int_0^{q_j-v} \int_{(u_i+v)/q}^1 1 du_j du_i} = \frac{2}{3} + \frac{v}{3q} & \text{for } q \in [v, 1+v], \end{cases}$$

and it is not defined for $q < v$. Also, $\mathbb{E}[u_i|u_i + v \geq qu_j] =$

$$\begin{cases} v+\frac{1}{2} & \text{for } q \geq 1/v, \\ \frac{-2v^3+3v^2/q+6v/q-1/q^3+3/q}{-3v^2+6v/q-3/q^2+6/q} & \text{for } q \in (\frac{1}{1+v}, \frac{1}{v}) \\ \frac{6v^2+6v+2}{6v+3} & \text{for } q \leq \frac{1}{1+v}. \end{cases}$$

From above, it is easy to see that $\mathbb{E}[u_i+v|u_i+v \geq u_j] > \mathbb{E}[u_i|u_i \geq u_j+v]$ for all $v \in (0,1)$. Also, we can check that following claim holds, on which following proves are based.

Claim 10. *$\mathbb{E}[u_j|q_j u_j \geq u_i+v]$ is decreasing in q and $\mathbb{E}[u_i|u_i+v \geq q_j u_j]$ is non decreasing in q.*

Proof of Proposition 7:

It is immediately seen from the following proofs and, hence, omitted. *Q.E.D.*

Proof of Theorem 19:

Note that the condition $\left(\frac{\pi^2}{\pi^2+(1-\pi)^2}, 0, 0, \frac{(1-\pi)^2}{\pi^2+(1-\pi)^2}\right) \in S_T$ is rewritten as

$$\frac{\pi^2}{\pi^2+(1-\pi)^2}\left(v+\frac{2}{3}\right)+\frac{(1-\pi)^2}{\pi^2+(1-\pi)^2}\cdot\frac{2}{3} > v_0 \tag{3.6}$$

Suppose that $v_1 = v_2$ is realized, and in the initial τ periods, which we take τ even number, project one and two are recommended in turn, i.e., $a^\tau = (1,2,...,1,2)$. Then from (3.6) and $\varphi(2,0) > v_0$, it is ensured that the truthful equilibrium is kept played. Note that $s_{00}^\tau/s_{vv}^\tau = (1-\pi)^2/\pi^2$, and by taking

τ sufficiently large, we can make $\beta s_{00}^{\tau} > s_{v0}^{\tau} = s_{0v}^{\tau}$ for arbitrary $\beta < 1$. From (3.6), it is enough to show that the probability that an equilibrium path starts from $(s_{vv}^{\tau}, s_{v0}^{\tau}, s_{0v}^{\tau}, s_{00}^{\tau})$ never enter the pandering equilibrium phase is strictly positive. Then for the equilibrium path to enter the pandering equilibrium phase S_P, after the history a^{τ}, at least we must have $s_{vv}^{t}/s_{v0}^{t} < M$ for some M. Note that because $v_1 = v_2$, we have $s_{vv}^{t+1}/s_{v0}^{t+1} = 2g(v)\, s_{vv}^{t}/s_{00}^{t}$ with probability 1/2 and $s_{vv}^{t+1}/s_{v0}^{t+1} = 2\{1 - g(v)\} s_{vv}^{t}/s_{v0}^{t}$ with probability 1/2. Because $g(v) > 1/2$ and $(1 - g(v)) < 1/2$, but $g(v)(1 - g(v)) < 1/4$, we have $\mathbb{E}[s_{vv}^{t+1}/s_{v0}^{t+1} | s_{vv}^{t}/s_{v0}^{t}, v_1 = v_2] > s_{vv}^{t}/s_{v0}^{t}$. This implies that if agents keep recommending the better project irrespective of principals' action, $s_{vv}^{t+1}/s_{v0}^{t+1}$ converges to infinite almost surely, from martingale convergence theorem. Then, we can see that there is s and φ such that $\{s_{vv}^{t}/s_{v0}^{t}\}_{t=s}^{\infty}$ that starts from φ never become smaller than M, as long as principal has been kept recommending the better project. However by extending τ to s, we can reach $s_{vv}^{t}/s_{v0}^{t} > \varphi$, which demonstrates the theorem. *Q.E.D.*

Proof of Theorem 20:

1: Think of the case in which $v_1 = v$ and $v_2 = 0$. At each period t, given state $\mathbf{s}^t$, the players play an equilibrium such that the agent recommends project 1 if $b_1^t > qb_2^t$ (truthful equilibrium corresponds to the case $q = 1$). Such q is characterized by the number that satisfies $\mathbb{E}[b_1^t | b_1^t \geq qb_2^t, \mathbf{s}^t] > v_0$ and $\mathbb{E}[b_2^t | b_2^t \geq b_1^t/q, \mathbf{s}^t] = v_0$. Note that for all $(s_{vv}^t, s_{v0}^t, s_{0v}^t)$, $q \geq v$ and $q \leq \frac{1}{v}$. In the equilibrium, because the true state is $v_1 = 1$ and $v_0 = 0$, the agent recommends project one with probability $\Pr[b_1^t > q_2 b_2^t] = 1 - q + v + \frac{q^2 - v^2}{2q}$.

Think about the stochastic process of $\left\{\frac{s_{0v}^{t+1}}{s_{v0}^{t+1}}\right\}_{t=1}^{\infty}$. Because it is a martingale process, $\mathbb{E}\left[\frac{s_{0v}^{t+1}}{s_{v0}^{t+1}}|\frac{s_{0v}^{t}}{s_{v0}^{t}}\right]=\frac{s_{0v}^{t}}{s_{v0}^{t}}$ does not convergence to any dead wrong belief almost surely, i.e., *Support* $\left(\lim_{t\to\infty}\frac{s_{0v}^{t}}{s_{v0}^{t}}\right)\subset[0,\infty)$. Also, the process $\left\{\frac{s_{0v}^{t+1}}{s_{v0}^{t+1}}\right\}_{t=1}^{\infty}$ converges almost surely to a random variable. Hence it is enough to prove that $\lim_{t\to\infty}\frac{s_{0v}^{t}}{s_{v0}^{t}}=0$.

If we regard stochastic process of $\left\{\frac{s_{0v}^{t+1}}{s_{v0}^{t+1}}\right\}_{t=1}^{\infty}$, the *public likelihood ratio,* as a conditional stochastic process in given state of $v_1=0$ and $v_2=0$, we have

$$\begin{aligned}\mathbb{E}_{v_1=1\text{ and }v_2=0}\left[\frac{s_{0v}^{t+1}}{s_{v0}^{t+1}}|\frac{s_{0v}^{t}}{s_{v0}^{t}}\right] &= \Pr[b_1^t>qb_2^t]\frac{\Pr[u_1>q(u_2+v)]s_{0v}^t}{\Pr[b_1^t>qb_2^t]}\\ &\quad+\left(1-\Pr[b_1^t>qb_2^t]\right)\frac{\Pr[qu_2>u_1+v]s_{v0}^t}{\Pr[b_1^t>qb_2^t]s_{v0}^t}\\ &= \{\Pr[u_1>q(u_2+v)]+\Pr[qu_2>u_1+v]\}\frac{s_{v0}^t}{s_{0v}^t}.\end{aligned}$$

If $q\in(v,1+v)$, we have

$$\Pr[u_1>q(u_2+v)]=\int_0^1\int_{q(u_2+v)}^1 du_1du_2=1-\frac{q}{2}-qv$$

and

$$\Pr[qu_2>u_1+v]=\int_{q/v}^1\int_0^{qu_2-v} du_1du_2=\frac{q}{2}-v-\frac{q^3}{2v^2}+q,$$

and hence

$$\mathbb{E}_{v_1=1\text{ and }v_2=0}\left[\frac{s_{0v}^{t+1}}{s_{v0}^{t+1}}|\frac{s_{0v}^{t}}{s_{v0}^{t}}\right]=[1+q\left(1-v\right)-v-\frac{q^3}{2v^2}]\frac{s_{v0}^t}{s_{0v}^t}<\alpha\frac{s_{v0}^t}{s_{0v}^t}$$

for some $\alpha < 1$. Note that the expectation operator here is not computed by public belief. Rather, it is computed by using the true information of $v_1 = v$ and $v_0 = 0$. Similarly, we can prove that $\mathbb{E}_{v_1=1 \text{ and } v_2=0}\left[\frac{s_{0v}^{t+1}}{s_{v0}^{t+1}}\right] < \alpha\frac{s_{0v}^t}{s_{v0}^t}$ for some $\alpha < 1$ when $q_2 > 1+v$. This implies that for all $(s_{vv}^t, s_{v0}^t, s_{0v}^t)$, we have

$$\mathbb{E}_{v_1=1 \text{ and } v_2=0}\left[\frac{s_{0v}^{t+1}}{s_{v0}^{t+1}}|\frac{s_{0v}^t}{s_{v0}^t}\right] < \rho\frac{s_{0v}^t}{s_{v0}^t} \text{ for some } \rho \in (0,1).$$

This implies that $s_{v0}^{t+1}/s_{0v}^{t+1} \to 0$. Similarly, we can also prove that both $s_{vv}^{t+1}/s_{v0}^{t+1}$ and $s_{00}^{t+1}/s_{v0}^{t+1}$ converges almost surely to 0. Then, because Condition 1 is violated, we eventually have $\mathbf{s}^t \notin S_T \cup S_P$.

2. Before proving the statement, we give the following lemma.

Lemma 14. *If Condition 1 holds, the existence of a pandering equilibrium is also ensured when public belief is such that* $s_{vv} = s_{00} = 0$.

Proof. Suppose that $s_{vv} = s_{00} = 0$, and fix $s_{v0} = s_{v0}^*$. Without loss of generality, suppose that $\mathbb{E}_{s_{v0}^*}[b_1|b_1 \geq b_2] > v_0$ and $\mathbb{E}_{s_{v0}^*}[b_2|b_2 \geq b_1] < v_0$. Obviously, $s_{v0}^* > \frac{1}{2}$. Note that $\mathbb{E}_{s_{v0}^*}[b_1|b_1 \geq qb_2]$ and $\mathbb{E}_{s_{v0}^*}[b_2|qb_2 \geq b_1]$ are increasing and decreasing functions of q, respectively. Because for all $1 > s_{v0}$, we have $\mathbb{E}_{s_{v0}}[b_2|qb_2 \geq b_1] > \mathbb{E}_{s_{v0}=1}[b_2|qb_2 \geq b_1]$, there is q^* such that $\mathbb{E}_{s_{v0}^*}[b_2|q^*b_2 \geq b_1] = v_0$. Denote by q^{**} the threshold acceptance probability of project two when $v_1 = 1$ and $v_2 = 0$ in the pandering equilibrium.

Because there exists a pandering equilibrium when $v_1 = 1$ and $v_2 = 0$, we have $\mathbb{E}_{s_{v0}=1}[b_1|q^{**}b_2 \geq b_1] > v_0$ and $\mathbb{E}_{s_{v0}=1}[b_2|q^{**}b_2 \geq b_1] = v_0$. We can verify that

$$\Pr[u_1 + b > u_2]\mathbb{E}_{s_{v0}=1}[b_1|b_1 \geq b_2] + \Pr[u_2 > u_1 + v]\mathbb{E}_{s_{v0}=1}[b_2|b_2 \geq b_1] >$$

$$\Pr[u_1+v > q^{**}u_2]\mathbb{E}_{s_{v0}=1}[b_1|b_1 \geq q^{**}b_2]+\Pr[q^{**}\left(u_1 + v\right) > u_2]\mathbb{E}_{s_{v0}=1}[b_2|q^{**}b_2 \geq b_1] > v_0.$$

Define function a_1 and a_2 as follows:

$$\alpha_1\left(q, s_{v0}\right) = s_{v0}\Pr[u_1 + v > qu_2]\mathbb{E}_{s_{v0}=1}[b_1|b_1 \geq qb_2]$$

$$+\left(1 - s_{v0}\right)\Pr[u_1 > q\left(u_2 + v\right)]\mathbb{E}_{s_{v0}=0}[b_1|b_1 \geq qb_2],$$

and

$$\alpha_2\left(q, s_{v0}\right) = s_{v0}\Pr[qu_2 > u_1 + v]\mathbb{E}_{s_{v0}=1}[b_2|qb_2 \geq b_1]$$

$$+\left(1 - s_{v0}\right)\Pr[q\left(u_2 + v\right) > u_1]\mathbb{E}_{s_{v0}=0}[b_2|qb_2 \geq b_1].$$

Then we can compute $\frac{\partial\alpha_1(q,s_{v0})}{\partial q} > 0$, $\frac{\partial\alpha_1(q,s_{v0})}{\partial s_{v0}} > 0$, $\frac{\partial\alpha_2(q,s_{v0})}{\partial q} < 0$, and $\frac{\partial\alpha_1(q,s_{v0})}{\partial s_{v0}} < 0$. Define $q\left(s_{v0}\right)$ be a function such that $\alpha_2\left(q\left(s_{v0}\right), s_{v0}\right) = v_0$. Then $q\left(s_{v0}\right)$ is a decreasing function and we have $q\left(1\right) = q^{**}$, $q\left(s^*_{v0}\right) = q^*$, and $q\left(\frac{1}{2}\right) > 1$. Since $\alpha_1\left(q\left(\frac{1}{2}\right), \frac{1}{2}\right) > v_0$ and $\alpha_1\left(q\left(1\right), 1\right) > v_0$, and α_1 is a monotone function of two variables, we must have $\alpha_1\left(q\left(s^*_{v0}\right), s^*_{v0}\right) > v_0$. Hence we have both $\mathbb{E}_{s^*_{v0}}[b_1|b_1 \geq q^*b_2] > v_0$ and $\mathbb{E}_{s^*_{v0}}[b_2|q^*b_2 \geq b_1] = v_0$, which shows that there is a pandering equilibrium with acceptance probability of project two being q^*. $\square$

Suppose that Condition 1 is satisfied and $v_1 = v$ and $v_2 = 0$ is realized. Think of the sequence of idiosyncratic shocks up to period τ such that $b_1^t > b_2^t$ for $t < \tau$. We show that an influential equilibrium is kept played at least until τ. First, think of the case in which (3.6) is satisfied. Suppose that we have an influential equilibrium up until period $t-1$, and reached public belief $\mathbf{s}^t$. Let $\widehat{\mathbf{s}}^t = (0, s_{v0}^t / (s_{v0}^t + s_{0v}^t), s_{0v}^t / (s_{v0}^t + s_{0v}^t), 0)$. Then we have

$$\mathbb{E}[b_1^t | b_1^t \geq q b_2^t, \mathbf{s}^t] > \mathbb{E}[b_1^t | b_1^t \geq q b_2^t, \widehat{\mathbf{s}}^t]$$

and

$$\mathbb{E}[b_2^t | q b_2^t \geq b_1^t, \mathbf{s}^t] > \mathbb{E}[b_2^t | q b_2^t \geq b_1^t, \widehat{\mathbf{s}}^t].$$

Then the existence of pandering equilibrium at t is ensured by Lemma 2.

Next, think of the case in which (3.6) is not satisfied. If an influential equilibrium is kept played, we can form a sequence of q^t and $\mathbf{s}^s$, where q^t is generated by the condition $\mathbb{E}[b_2^t | q^t b_2^t \geq b_2^t, \mathbf{s}^t] = v_0$. Then the proof is done if we can see $\mathbb{E}[b_1^t | b_1^t \geq q^t b_2^t, \mathbf{s}^t] > v_0$ for all $t \leq \tau$. From Lemma 13, we have $s_{vv}^{t-1}/s_{00}^{t-1} < s_{vv}^t / s_{00}^t$. If it becomes

$$\frac{s_{vv}^s}{s_{vv}^s + s_{00}^s}\left(v + \frac{2}{3}\right) + \frac{s_{00}^s}{s_{vv}^s + s_{00}^s}\frac{2}{3} \geq v_0,$$

for some s, it is also satisfied for $t > s$. From Condition 1 and Claim 10, this implies that $\mathbb{E}[b_1^t | b_1^t \geq q^t b_2^t, \mathbf{s}^t] > v_0$ for all $t \geq s$. Hence focus on the case that above is not satisfied. However, in such a case, because $s_{vv}^{t-1}/s_{v0}^{t-1} < s_{vv}^t / s_{v0}^t$, and $s_{vv}^{t-1}/s_{0v}^{t-1} < s_{vv}^t / s_{0v}^t$, we have $\mathbb{E}[b_1^t | b_1^t \geq q b_2^t, \mathbf{s}^t] > \mathbb{E}[b_1^{t-1} | b_1^{t-1} \geq q b_2^{t-1}, \mathbf{s}^{t-1}]$ and

$\mathbb{E}[b_2^t|qb_2^t \geq b_1^t, \mathbf{s}^t] > \mathbb{E}[b_2^{t-1}|qb_2^{t-1} \geq b_2^{t-1}, \mathbf{s}^{t-1}]$ for all q. We know that from Condition 1, by taking τ sufficiently high, we have an influential equilibrium if $\mathbf{s} = \mathbf{s}^\tau$. This and (A1), which shows the existence of influential equilibrium when $\mathbf{s} = \mathbf{s}^1$, imply that we also have an influential equilibrium for $t \in \{2, ..., \tau - 1\}$.

Hence with a strictly positive probability, an equilibrium path enters a pandering equilibrium phase with sufficiently high s_{v0}. Then, by employing the same reasoning as in Theorem 19, we can prove that the equilibrium may perpetually stay in the pandering equilibrium phase with a strictly positive probability. *Q.E.D.*

Proof of Theorem 21:

Think about the following sequence of realizations of idiosyncratic shocks (realizations of fundamental qualities can be anything), and also let players play the (unique) influential equilibrium at each period given the history before their period. For the first period, $b_1^1 > b_1^2$. After the second period, let $b_t^1 > b_t^2$ if players are supposed to play the truthful equilibrium at period t, and let $q^t b_t^2 > b_t^1$ if players are supposed to play a pandering equilibrium with $q^t < 1$ at period t, and $b_t^1 > q^t b_t^2$ if players are supposed to play a pandering equilibrium with $q^t > 1$ at period t. Note that this path of random shocks generates the following path: the truthful equilibrium is followed by a pandering equilibrium with $q^t < 1$, and when a pandering equilibrium is played the worse-looking project is proposed. Note that if the truthful equilibrium is played at period t, we have $s_{00}^{t+1}/s_{vv}^{t+1} = s_{00}^t/s_{vv}^t$, and if a pandering equilibrium is played,

(3.7)
$$\frac{s_{00}^{t+1}}{s_{vv}^{t+1}}=\frac{1-\Pr[u>\gamma u]}{1-\Pr[u+v>\gamma\left(u+v\right)]}\frac{s_{00}^{t}}{s_{vv}^{t}}=\frac{\frac{1}{2}\gamma}{\frac{1}{2}\gamma-\left(1-\gamma\right)v+\frac{\left(1-\gamma\right)^{2}}{2\gamma}v^{2}}\frac{s_{00}^{t}}{s_{vv}^{t}}=\lambda\left(\gamma\right)\frac{s_{00}^{t}}{s_{vv}^{t}},$$

where $\gamma=\max\left\{q,\frac{1}{q}\right\}$. Because $\lambda\left(\gamma\right)>1$ for all $\gamma<1$, s_{00}^{t}/s_{vv}^{t} monotonically increases over time. If we get the point that an influential equilibrium does not exist, the proof is done. Hence suppose that we can continue without having communication to break down in a finite period, in order to get a contradiction.

Take a subsequence k of t such that a pandering equilibrium is played and the truthful equilibrium is played at $t+1$. This implies that $\frac{s_{00}^{k}}{s_{v0}^{k}}>$ is uniformly bounded below. Then think about the sequence q^{k}, where q^{k} is the corresponding acceptance probability in a pandering equilibrium at k. Note that we have

$$\frac{\Pr[qu_{2}\geq u_{1}]}{\Pr[qu_{2}\geq u_{1}+v]+\Pr[q\left(u_{2}+v\right)\geq u_{1}]}>\frac{1}{2} \tag{3.8}$$

for all q. This implies that $\min\{s_{v0}^{t},s_{0v}^{t}\}/s_{00}^{t}$ converges to zero. Also, from the construction, we must have $s_{0v}^{k}<s_{v0}^{k}$. These imply that $\frac{s_{00}^{t}}{s_{0v}^{t}}\rightarrow\infty$ and we can find a sufficiently small uniform lower bound for $\frac{s_{00}^{t}}{s_{0v}^{t}}$.

First, think of the case in which q^{k} converges to 1. From $\frac{s_{00}^{t}}{s_{0v}^{t}}\rightarrow\infty$, we have $\lim_{k\rightarrow\infty}\mathbf{s}^{k}=\left(\widehat{s}_{vv},\widehat{s}_{v0},0,\widehat{s}_{00}\right)$. Note that in a sufficiently long run, a pandering equilibrium reverts to the truthful equilibrium at once because $\left|1-g\left(v\right)\right|<g\left(v\right)$, and

the equilibrium exhibits a cycle. Then we must have

$$\frac{1}{2\left(1-g\left(v\right)\right)}\frac{\widehat{s}_{00}}{\widehat{s}_{v0}}=\frac{\overline{s}_{00}}{\overline{s}_{v0}}\text{ and }\left(\frac{1}{2g\left(v\right)}\right)^{s}\frac{\overline{s}_{00}}{\overline{s}_{v0}}=\frac{\widehat{s}_{00}}{\widehat{s}_{v0}},$$

which implies that there must be s such that $\left(\frac{1}{2g(v)}\right)^{s}=2\left(1-g\left(v\right)\right)$, which can not happen for almost all constellation of parameter specifications.

Therefore, think about the case in which the sequence q^k does not converge to 1. If q^k does not converges to 1, there is $\underline{q}< 1$ such that q^k becomes smaller than $\underline{q}$ for infinitely many times. Because $\frac{s_{00}^{t+1}}{s_{vv}^{t+1}}$ is an increasing sequence, we have $\frac{s_{00}^{t+1}}{s_{vv}^{t+1}}\to\infty$ from (3.7). Then it must hold that $\lim_{k\to\infty}\mathbf{s}^k=(0,\widehat{s}_{v0},0,\widehat{s}_{00})$. However, it implies that eventually the equilibrium path does not allow the existence of the truthful equilibrium, which cannot happen from our choice of sequence of idiosyncratic shocks. *Q.E.D.*

Proof of Theorem 22:

1. Suppose that delegation does not stop until period τ. Think about the decision of the principal at period $\tau+1$. Think of the case that $\mathbf{s}^t\in S_P$. Then, there is q such that $\mathbb{E}[u_1+v|u_1+v\geq qu_2]>v_0$ and $\mathbb{E}[u_2|qu_2\geq u_1+v]=v_0$, which implies that

$$\Pr[u_1+v\geq qu_2]\mathbb{E}[u_1+v|u_1+v\geq qu_2]+\Pr[qu_2\geq u_1+v]\mathbb{E}[u_2|qu_2\geq u_1+v]>v_0.$$

From the discussion in Remark 5, we have

$$\Pr[b_1 \geq b_2|\mathbf{s}]\mathbb{E}[b_1|b_1 \geq b_2, \mathbf{s}] + \Pr[qb_2 \geq b_1|\mathbf{s}]\mathbb{E}[b_2|b_2 \geq b_1, \mathbf{s}]$$

$$= \Pr[u_1 + v \geq u_2]\mathbb{E}[u_1 + v|u_1 + v \geq u_2] + \Pr[u_2 \geq u_1 + v]\mathbb{E}[u_2|u_2 \geq u_1 + v]$$

$$= \Pr[u_1+v \geq qu_2]\mathbb{E}[u_1+v|u_1+v \geq qu_2]+\Pr[qu_2 \geq u_1+v]\mathbb{E}[u_2|qu_2 \geq u_1+v] > v_0,$$

for all $\mathbf{s}^t$ such that $s^t_{vv} = s^t_{00} = 0$. This demonstrate that delegation attains higher payoff for the principal when $\mathbf{s}^t \in S_P$.

On the other hand, even think of the case that $\mathbf{s}^t \notin S_P \cup S_T$. Because $s^t_{vv}/s^t_{00} = \pi^2/(1-\pi)^2$, Condition 2 and $\left(\frac{\pi^2}{\pi^2+(1-\pi)^2}, 0, 0, \frac{(1-\pi)^2}{\pi^2+(1-\pi)^2}\right) \in S_T$ imply

$$\Pr[b_1 \geq b_2|\mathbf{s}]\mathbb{E}[b_1|b_1 \geq b_2, \mathbf{s}] + \Pr[b_2 \geq b_1|\mathbf{s}]\mathbb{E}[b_2|b_2 \geq b_1, \mathbf{s}]$$

$$= \beta\left\{\frac{\pi^2}{\pi^2+(1-\pi)^2}\mathbb{E}[u_1+v|u_1 \geq u_2] + \frac{(1-\pi)^2}{\pi^2+(1-\pi)^2}\mathbb{E}[u_1|u_1 \geq u_2]\right\}$$

$$\begin{aligned} +(1-\beta)\{\alpha\mathbb{E}[u_1+v|u_1+v &\geq u_2] + (1-\alpha)\mathbb{E}[u_1|u_1 \geq u_2+v]\} \\ &> \beta v_0 + (1-\beta)v_0 > v_0, \end{aligned}$$

for some $\beta \in (0,1)$ and $\alpha \in (0,1)$. Hence delegation attains higher payoff for the principal even when $\mathbf{s}^t \notin S_P \cup S_T$. Because the delegation attains the same payoff when $\mathbf{s}^t \in S_T$, the statement follows.

2. Suppose that Condition 2 does not hold. Think of the sequence of the realizations of idiosyncratic shocks such that $b^t_1 > b^t_2$ for even periods and $b^t_1 < b^t_2$ for

odd periods. Apparently, $\mathbf{s}^t$ converges to $\mathbf{s}_=$, which implies the first part. Also, from the same reasoning as in the proof of Theorem 21, $\mathbf{s}^t$ converges to $\mathbf{s}_=$ almost surely when $v_1 = v_2$, so long as communication is sustained. Hence eventually we have $\mathbf{s}^t \notin S_T$ and, hence, delegation stops. The third statement of the theorem is proved in a similar way. *Q.E.D.*

Proof of Theorem 24:

Suppose that $v_1 = v_2 = v$. In the truthful equilibrium phase, if the outside option is not proposed at period t, we have

$$\frac{s_{00}^{t+1}}{s_{vv}^{t+1}} = \{1 - p + p\,\Pr[\max\{u_1, u_2\} \geq \kappa v_0]\}\frac{s_{00}^t}{s_{vv}^t} < \alpha\frac{s_{00}^t}{s_{vv}^t},$$

for $\alpha < 1$, which implies that $\frac{s_{00}^t}{s_{vv}^t} \to 0$. On the other hand, if $v_1 = v_2 = 0$, $\frac{s_{vv}^t}{s_{00}^t}$ becomes zero once the outside option is recommended. Because $\lim_{\tau\to\infty} \Pr[u_t > \kappa v_0 \text{ for all } \tau] = 0$, eventually we have $\frac{s_{vv}^t}{s_{00}^t} = 0$. The result follows from those observations, combined with Theorem 20. *Q.E.D.*

References

[1] Aragones, Enriqueta, and Thomas R. Palfrey. (2001). "Mixed Equilibrium in a Downsian Model with a Favored Candidate." *Journal of Economic Theory,* 103, 131-161.

[2] Aumann, Rovert J., and Sergiu Hart. (2003). "Long Cheap Talk." *Econometrica,* 71, 1619-1660.

[3] Bagwell, Kyle., and Garey Ramey. (1991). "Oligopoly Limit Pricing." *Rand Journal of Economics,* 22, 155-172.

[4] Banerjee, Abhijit V. (1992). "A Simple Model of Herd Behavior." *Quarterly Journal of Economics,* 107, 797-817.

[5] Baliga, Sandeep, and Jeffrey C. Ely. (2010). "Torture." Working Paper, Northwestern University.

[6] Banks, Jeffrey S. (1990). "A Model of Electoral Competition with Incomplete Information." *Journal of Economic Theory,* 50, 309-25.

[7] Besley, Timothy, and Stephen Coate. (1997). "An Economic Model of Representative Democracy." *Quarterly Journal of Economics,* 112, 85-114.

[8] Bikhchandani, Sushil, Hirshleifer, David, and Ivo Welch. (1992). "A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades." *Journal of Political Economy,* 100, 992-1026.

[9] Brandenberger, Adam, and Ben Polak. (1996). "When Managers Cover Their Posteriors: Making the Decisions the Market Wants to See." *Rand Journal of Economics,* 27, 523-41.

[10] Callander, Steven, (2008). Political Motivations. *Review of Economic Studies,* 75, 671-697.

[11] Callander, Steven., and Simon Wilkie, (2007). "Lies, Damned Lies, and Political Campaigns." *Games and Economic Behavior* 60, 262-286.

[12] Calvert, Randall L. (1985). "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence." *American Journal of Political Science,* 29, 69-95.

[13] Caselli, Francesco, and Massimo Morelli. (2004). "Bad Politicians." *Journal of Public Economics,* 115, 151-81.

[14] Chakraborty, Archishman, and Rick Harbaugh. (2007). "Comparative Cheap Talk." *Journal of Economic Theory,* 132, 70-94.

[15] Che, Yeon-Koo, Dessein, Wouter, and Navin Kartik. (2011). "Pandering to Persuade." *American Economic Review,* forthcoming.

[16] Che, Yeon-Koo, and Navin Kartik. (2009). "Opinions as Incentives." *Journal of Political Economy,* 117, 815-860.

[17] Cho, In-Koo, and David Kreps. (1987). "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics,* 102, 179-221.

[18] Crawford, Vincent, and Joel Sobel. (1982). "Strategic Information Transmission." *Econometrica,* 50, 1431-1451.

[19] Dasgupta, Amil, and Andrea Prat. (2008). "Information Aggregation in Financial Markets with Career Concerns." *Journal of Economic Theory,* 143, 83-113.

[20] Davis, Otto A., Hinich, Melvin J., and Peter C. Ordeshook. (1970). "An Expository Development of a Mathematical Model of the Electoral Process." *American Political ScienceReview*, 64, 426-48.

[21] Dewatripont, Mathias., and Jean Tirole. (2005). "Modes of Communication." *Journal of Political Economy,* 113, 1217-1238.

[22] Downs, Anthony, (1957). An Economic Theory of Democracy. New York: Harper.

[23] Dziuda, Wioletta. (2007). "Strategic Argumentation." *Journal of Economic Theory,* Forthcoming.

[24] Eso, Peter, and Yuk Fong. (2008). "Wait and See." mimeo.

[25] Fedderson, Timothey., and Wolfgang Persendorfer. (1997). "Convicting the innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting." *American Political Science Review*, 92, 23–35.

[26] Forges, Francois, and Frederic Koessler (2008). "Long Persuasion Games." *Journal of Economic Theory,* 143, 1-35.

[27] Fudenberg, Drew, and Jean Tirole. (1986). "A Theory of Exit in Duopoly." *Econometrica*, 54, 541-568.

[28] Fudenberg, Drew, and Jean Tirole. (1991). Game Theory. Cambridge, MA: MIT Press.

[29] Fudenberg, Drew, and Jean Tirole. (1991). "Perfect Bayesian Equilibrium and Sequential Equilibrium." *Journal of Economic Theory,* 53, 236-260.

[30] Glazer, Jacob, Ariel Rubinstein. (2006). "On Optimal Rules of Persuasion." *Econometrica* 72, 1715-1736.

[31] Grossman, Sanford J. (1981). "The informational Role of Warranties and Private Disclosure about Product Quality." *Journal of Law and Economics*, 24, 461-483.

[32] Harsanyi, Jhon C. (1973). "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Point." *International Journal of Game Theory*, 27, 245-52.

[33] Heidhues, Paul, and Johan N. M. Lagerlof. (2003). "Hiding Information in Electoral Competition." *Game and Economic Behavior,* 42, 48-74.

[34] Hendricks, Ken, Weiss, Andrew, and Charles A. Wilson. (1988). "The War of Attrition in Continuous Time with Complete Information." *International Economic Review,* 29, 663-680.

[35] Hörner, Johannes, and Andrzej Skrzypacz. (2011). "Selling Information." Cowles Foundation Discussion Paper.

[36] Jensen, Thomas. (2009). "Electoral Competition when Candidates are Better Informed than Voter." mimeo.

[37] Kamenica, Emir, and Matthew Gentzkow. (2010). "Bayesian Persuasion." *American Economics Review,* forthcoming.

[38] Kartik, Navin. (2009). "Strategic Communication with Lying Costs." *Review of Economic Studies*, 76, 1359-1395.

[39] Kartik, Navin, and McAfee, Preston. (2007). "Signaling Character in Electoral Competition." *American Economic Review,* 97, 852-870.

[40] Kartik, Navin, Ottaviani, Marco, and Squintani, Grancesco. (2007). "Credulity, Lies, and Costly Talk." *Journal of Economic Theory*, 134, 93-116.

[41] Kartik, Navin, Squintani, Francesco, and Kartin, Tinn. (2012). "Information Revelation and (Anti-)Pandering in Elections." mimeo.

[42] Kreps, David and Robert Wilson. (1982a). "Sequential Equilibrium." *Econometrica* 50, 863-894.

[43] Kreps, David and Robert Wilson. (1982b). "Reputation and Imperfect Information." *Journal of Economic Theory,* 27, 253-79.

[44] Krishna, Vijai and John Morgan. (2004). "The Art of Conversation: Eliciting Information from Experts through Multi-Stage Communication." *Journal of Economic Theory*, 117, 147-179.

[45] Lee, In-Ho. (1993). "On the Convergence of Informational Cascade." *Journal of Economic Theory,* 61, 395-411.

[46] Leslier, Jean-Francois., and Straeten, Karine. (2004). "Electoral Competition under Imperfect Information." *Economic Thoery*, 24, 419-446.

[47] Loertscher, Simon. (2012). "Location Choice and Information Transmission." mimeo.

[48] Majumdar, Sumon and Sharun W. Mukand. (2004). "Policy Gambles." *American Economic Review,* 94, 1207-1222.

[49] Martinelli, Cesar. (2001). "Elections with privately informed parties and voters." *Public Choice* 108, 147–167.

[50] Milgrom, Paul. (1981). "Good News and Bad News Represenation Theorem and Applications." *Bell Journal of Economics,* 12, 380-391.

[51] Ordever, Janusz A. and Ariel Rubinsten. (1986). "A Sequential Concession Game with Asymmetric Information." *Quarterly Journal of Economics,* 101, 879-888.

[52] Osborne, Martin. J., and Al Slivinsky. (1996). "A Model of Political Competition with Citizen-Candidates." *Quarterly Journal of Economics,* 111, 65-96.

[53] Ottaviani, Marco and Peter Sorensen. (2006) "Professional Advice." *Journal of Economic Theory,* 126, 120-142.

[54] Prendergast, Canice and Lars A. Stole (1996). "Impetuous Youngsters and Jaded Old-timers." *Journal of Political Economy,* 104, 1105-34.

[55] Quement, Marc. (2010). "Persuasion and Rhetorical Moderateness." mimeo.

[56] Scharfstein, David S., and Jeremy C. Stein. (1990). "Herd Behavior and Investment." *American Economic Review,* 80, 465-479.

[57] Schultz, Christian. (1996). "Polarization and Inefficient Policies." *Review of Economic Studies,* 63, 331-44.

[58] Sher, Itai. (2010). "Persuasion and Dynamic Communication." mimeo.

[59] Shin, Hyun S. (1994). "The Burden of Proof in a Game of Persuasion." *Journal of Economic Theory*, 64, 253-264.

[60] Smith, Lones and Peter Sorensen. (2000). "Pathological Outcomes of Observational Learning." *Econometrica,* 68, 371-398.

[61] Sobel, Joel. (1985). "The Theory of Credibility." *Review of Economic Studies,* 52, 557-573.

[62] Stokes, Donald E. (1963). "Spatial Models of Party Competition." American Political Science Review, 57, 368-377.

[63] Verrechia, Robert E. (1983). "Discretionary Disclosure." *Journal of Accounting and Economics,* 5, 179-194.

[64] Wittman, Donald. (1977). "Candidate with Policy Preferences: A Dynamic Model." *Journal of Economic Theory,* 14, 180-89.