Behavioral Types and Partially Informed Decision Makers in Communication Games

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Ying Chen

Dissertation Director: Professor David Pearce

May 2005

ABSTRACT

Behavioral Types and Partially Informed Decision Makers in Communication Games

Ying Chen

2005

This dissertation consists of two chapters on communication games.

The first chapter introduces two behavioral types into a model of strategic communication based on Crawford and Sobel (1982). With positive probability, the sender is an *honest* type who always tells the truth, and the receiver is a *naive* type who always follows whatever message is sent to her. We establish existence and uniqueness of monotonic equilibrium (a sequential equilibrium in which the sender's message strategy is non-decreasing in the state of the world) under certain assumptions. In this important class of equilibria, we find that only the most informative equilibrium in the C-S model is robust to the perturbation of the behavioral types. In a monotonic equilibrium, the dishonest sender always distorts the messages in the direction of his bias. If the message space is discrete and the dishonest sender has an upward bias, then his messages will cluster around the top few messages. Interestingly, the sophisticated receiver's strategy is *not* monotonic in the messages she receives even in a monotonic equilibrium. The existence of monotonic equilibrium may fail when

the message space is a continuum. Following Manelli (1996), we show that adding incentive compatible communication (cheap-talk extension) restores existence.

The second chapter incorporates partially informed decision makers into communication games. We analyze three extensive form games in which the expert and the decision maker (DM) each privately observe a signal about the state of the world. In game 1, the DM reveals her private signal to the expert before the expert reports to her. In game 2, the DM keeps her signal private. In game 3, the DM strategically communicates to the expert before the expert reports to her. We find that the DM's expected equilibrium payoff is *not* monotonically increasing in the accuracy of her private signal because the expert may reveal less information when facing a well-informed DM. Whether the DM extracts more information from the expert in game 1 or in game 2 depends on the parameters. Allowing the DM to communicate strategically to the expert first does not help her extract more information.

Contents

List of Figures

Acknowledgments

I am indebted to David Pearce for his continuous support and advice throughout the dissertation writing process. I am grateful to Stephen Morris for highly valuable conversations. I also thank Vincent Crawford, Ezra Friedman, Dino Gerardi, Alvin Klevorick, Joel Sobel for helpful comments and suggestions.

to my parents

Chapter 1

Perturbed Communication Games with Honest Senders and Naive Receivers

1.1 Introduction

It is common in economic life that a decision maker does not have all the relevant information for making a good decision and therefore needs access to better informed people for knowledge and expertise. Here are a few examples: government officials consult experts before making policies; investors constantly get suggestions and recommendations from financial analysts; employers routinely ask for references from job applicants.

To a large extent, the information in many of these situations is transmitted through what economists call "cheap talk", that is, unmediated communication that does not directly affect either side's payoff. Since the interests of the two parties are usually not perfectly aligned, the informed party has an incentive to mislead the decision maker to believe something other than the truth. One central concern is whether information can ever be conveyed through cheap talk.

In Crawford and Sobel's (1982) seminal article, they model the situations described above as sender-receiver games in which the sender, after privately observing the state of the world, sends a costless message to the receiver who then chooses an action that affects both players' payoffs. They find that due to the difference in interests, information can never be fully transmitted in equilibrium. However, when the two parties' interests are close enough, cheap talk can be effective – there exist equilibria in which the sender introduces noise into his messages and the receiver extracts a certain amount of imprecise information from them.

Although cheap talk *can* convey some information when the sender's bias is small enough, it *need not do so* – there are always equilibria in which cheap talk is ineffective and no information is conveyed from the sender to the receiver. The problem of multiplicity of equilibria in cheap talk games provides a major motivation for our study. In this paper, we address the problem by introducing behavioral types into a game of information transmission through costless messages based on Crawford and Sobel (C-S for short) (1982). Specifically, with some positive probability, the sender is an *honest* type who always truthfully reports his observation of the state of the world and the receiver is a *naive* type who always follows whatever message is sent to her. Otherwise, the players are the same as in the original C-S model and act fully strategically to maximize their expected payoffs.[1]

Experimental[2] and anecdotal[3] evidence supports the observation that some players follow these behavioral rules. We will not attempt to explain why some people "mechanically" tell the truth or follow the messages. Some players may have preferences that induce them to adopt the "behavioral rules" as their optimal strategies in this one-shot interaction. Alternatively, some players may lack the strategic sophistication to figure out how to play the game optimally and therefore act "irrationally" in ways *as if* their preferences were the same as the former kind of players. In this paper, we take the existence of the behavioral types as given and explore its implications. For our purpose, we view them in the following way: they have different preferences from the non-behavioral players that make "telling the truth" or "following the sender's messages" dominant strategies for them.

Introducing behavioral types into the C-S model fundamentally changes the way the game is played. The messages are no longer pure "cheap talk" because they

[1] The strongest results in Crawford and Sobel (1982) are obtained for the "uniform-quadratic" case in which the prior on the state of the world is uniform on a bounded interval and the players' payoff functions have the quadratic form. This paper adopts the "uniform-quadratic" assumption. At this point, it is not clear how to extend the results to more general settings.

[2] See Forsythe, Lundholm and Rietz (1999) for an experimental study which finds that in a cheap talk game, the subjects (as senders) display a tendency to reveal the true state of the world in their costless disclosures and they (as receivers) show a certain amount of gullibility even when their opponents have a clear incentive to lie.

[3] For example, news stories have documented that rumors on a stock's prospects that float on internet message boards can drive the price of the stock up or down drastically because some investors believe those messages despite the fact that they come from anonymous sources with dubious motives.

enter the sender's payoff function through the fixed responses of the naive receiver. Moreover, this "cost" of sending a particular message varies with the true state of the world. When the dishonest sender decides what message to send, he weighs the strategic response of the sophisticated receiver against the fixed response of the naive receiver. This imposes a lot of structure on the incentive constraints that the sender faces and drives many of the results we derive in this paper.

The existence of the honest sender implies that every message is sent on the equilibrium path if we assume that the message space is the same as the state space. This enables us to always find the receiver's posterior belief by using Bayes' rule. The question of what kind of restrictions to impose on beliefs off the equilibrium path, which is a prominent issue in signaling and cheap talk games, simply does not arise in our model. By developing some of the techniques for solving signaling games with a continuum of types,[4] we are able to translate the equilibrium conditions into a differential equations system in terms of the non-behavioral types' strategies and solve for the equilibrium numerically.

Perturbing the C-S model with the behavioral types provides a new perspective on the equilibrium selection problem in cheap talk games, which are notorious for having a serious problem of multiplicity of equilibria. In the C-S model, the size of the sender's bias determines how much information can be transmitted in the "most informative" equilibrium. However, due to the circularity embedded in the solution

[4]See, in particular, Mailath (1987), (1992).

concepts we typically use, there are always other equilibria in which the messages are less informative. This is highlighted in the example of babbling equilibrium in which the sender's messages are uncorrelated with his private information and are taken to be meaningless. The receiver who responds to each message with the same action does not have any incentive to deviate because the messages contain no useful information. In turn, the sender's "babbling" is justified since the receiver "ignores" what he says. By letting the probabilities of the behavioral types approach zero, we can conduct a formal robustness test for the multiple equilibria that exist in the C-S model. We find that if we focus on the class of monotonic equilibria (these are equilibria in which the sender's message strategy is non-decreasing in the state of nature), *then we rule out all but the most informative equilibrium in the C-S model.* To see why, let's suppose, without loss of generality, the sender has an upward rather than downward bias. Then, in a monotonic equilibrium in the perturbed model, the dishonest sender with the lowest signal always has the opportunity to report that he is the lowest type and be believed by the receiver. This immediately implies that the expected equilibrium payoff for the lowest type of sender has to be at least as high as the payoff he gets if identified by the receiver as the lowest type. This condition on expected equilibrium payoffs holds in the limit as the probabilities of the behavioral types go to zero. Because among all the equilibria in the C-S model, only the most informative one satisfies this condition, all the other equilibria are eliminated by the perturbation. This result is different from what we get when we apply prominent refinements of cheap talk equilibria in the literature to the C-S model. None of the equilibria in the C-S model is "neologism-proof" (Farrell (1993)) or "announcement-proof" (Matthews, Okuno-Fujiwara and Postlewaite (1991)). In contrast, the solution concept of "credible message rationalizability" (Rabin (1990)) does not rule out any equilibrium in the C-S model. In section 1.5.2, we provide a more detailed discussion of the refinements mentioned above. An exception is Kartik's (2003) paper which considers a model of communication through both costly signals and cheap talk and uses it as a refinement of equilibria in the C-S model. Similar to our result, Kartik (2003) selects the most informative equilibrium. There are some interesting connections between Kartik's model and ours and a more detailed comparison of the two models is provided at the end of section 1.3.

In many situations of strategic information transmission, the messages that the sender uses have pre-existing, commonly understood focal meanings. However, in most models of cheap talk, the real meaning of a message, i.e., the information it contains, arises endogenously in equilibrium and the focal meaning (which is established in a much larger environment than the game) is completely irrelevant. In fact, given an equilibrium in a cheap talk game, "any permutation of messages across meanings gives another equilibrium."[5] This makes perfect sense from a purely theoretical standpoint because the focal meanings are completely extrinsic to the game. But it is also counter-intuitive. Imagine a situation in which people communicate through

[5] Farrell (1993).

a natural language or some other system where messages have well-understood focal meanings. Although rational agents may choose not to interpret the messages literally, we would still expect the messages' focal meanings to influence the communication process. In our model, this is indeed the case. With just a small fraction of players being "literal-minded," exaggeration arises naturally in a monotonic equilibrium – the dishonest sender always distorts his messages in the direction of his bias in the attempt to manipulate the receiver's beliefs. Moreover, if the sender has an upward bias and the message space is discrete, then it emerges in equilibrium that messages will cluster around the top few messages. Interestingly, in equilibrium, *the action that the sophisticated receiver takes is not monotonic in the messages she receives*, even when the dishonest sender's reporting strategy is increasing in the state of nature. Roughly speaking, a higher message may lead the receiver to choose a lower action if she believes that with a relatively high probability, the message is sent by the dishonest sender (who exaggerates his claims in equilibrium).

Another contribution of this paper is the analysis of the relationship between the game with a continuous message space and its finite approximating versions. Economists often choose to model games with an uncountable number of actions because it is easier to work with a continuum than with a large finite grid. Indeed, in this paper, tools from calculus enable us to reduce an infinite number of incentive constraints to an analytically simple form – a system of differential equations. However, when the message space is a continuum, the existence of monotonic equilibrium may fail. Fortunately, we can show that the non-existence problem does not arise if the message space is discrete. Approximating the infinite game with similar games that have finite message spaces resolves the non-existence problem. To find the limit distribution of the equilibrium outcomes in the converging sequence of finite games, we use a device that is introduced by Manelli (1996), the "cheap-talk extension" game.[6] In a cheap-talk extension of our sender-receiver game, in addition to the regular message the sender reports, he makes a costless, non-binding suggestion of action to the receiver. Manelli (1996) establishes upper hemi-continuity of sequential equilibria between the finite approximating games and the cheap-talk extension of the limit continuous game. We also establish lower hemi-continuity between the two in our model. The one-to-one mapping between the limit distribution of the equilibrium outcomes in the finite approximating games and the equilibrium outcome in the cheap-talk extension of the continuous game allows us to find the former by characterizing the latter. This exercise highlights the root of the non-existence problem in the continuous game. When the message space is discrete, the sender is able to convey different information by using different messages that are at the top and right next to each other in the message space. As the message space approaches the continuum, those messages collapse to one, and the sender can no longer use them to convey different information to the receiver, thereby resulting in non-existence.

Two papers in the literature are closely related to ours. Crawford (2003) considers

[6] Jackson, Simon, Swinkels and Zame (2002) apply a similar idea to a broader class of games of incomplete information.

a binary, asymmetric, zero-sum game in which one of the players can costlessly signal his intention of play. Standard equilibrium analysis predicts that when the players have opposite interests, pre-game communication like this involves babbling in which the signals are completely ignored. However, by introducing "mortal types" (these include "truth tellers" and "liars" on the sender's side and "believers" and "inverters" on the receiver's side), Crawford (2003) finds that misrepresentation of intentions can be successful sometimes – when the probabilities of the mortal types are high enough, even the sophisticated receiver can be fooled ex post. Ottaviani and Squintani (2002) introduce honest senders/naive receivers into a C-S kind of sender-receiver game. However, instead of maintaining the bounded state/message space of the C-S model as this paper does, their model assumes that the state/message space is unbounded and generates different results. Because there is no bound on how distorted a message can be, there always exists a fully separating equilibrium in which the sender adopts a strictly increasing message strategy.

Incomplete information about the players' preferences plays an important role in a number of other studies on information transmission, in both static and dynamic settings. Morgan and Stocken (2003) analyze how uncertainty about stock analysts' incentives affects stock recommendations. Sobel (1985) and Benabou and Laroque (1992) model the dynamics of a sender's "credibility" in a long term relationship when he can be either a "friend" or an "enemy" to the receiver. Morris (2001) explains that an advisor whose preferences are identical with the decision maker's may have a reputational incentive to lie and be "politically correct" because he does not want to be perceived as being biased.

The rest of the chapter is organized as follows. Section 1.2 describes the model. Section 1.3 lays out the theory that links the game that has a continuous message space with its finite approximations. Section 1.4 characterizes the class of monotonic equilibria. Section 1.5 presents a detailed examination and interpretation of the model's predictions, which include the case where the probabilities of behavioral types are non-negligible as well as the asymptotic case as the probabilities approach zero. For most part of the paper, we focus on the important class of monotonic equilibria. Section 1.6 discusses what happens when we relax this restriction. Section 1.7 concludes.

1.2 The Model

The benchmark is the classic model of strategic information transmission introduced by Crawford and Sobel (1982).

They consider the following game. There are two players, called a sender (S) and a receiver (R). At the beginning of the game, S privately observes the realization of a random variable, ω, called the "state of the world" and then sends a costless message, m, to R. Upon receiving m, R chooses an action, a, which affects both players' payoffs. For the rest of the paper, we are going to focus on the leading case of this sender-receiver game, known as the "uniform-quadratic" case. It is assumed that the state space $\Omega = [0,1]$ and the prior probability distribution of ω is the uniform distribution on the interval $[0,1]$. Both players have quadratic utility functions. Specifically, the von Neumann-Morgenstern utility functions for S and R are given respectively by

$$\begin{aligned} u^S(\omega, a, b) &= -(a - \omega - b)^2 \\ u^R(\omega, a, b) &= -(a - \omega)^2 \end{aligned} .$$

The "bias" of the sender, b, parameterizes the divergence of interest between the two parties. Without loss of generality, we are going to assume that $b \geq 0$. This means that for any ω, the sender's ideal action is always higher than the receiver's ideal action.

Note that unlike typical signaling models where differential signaling costs are exogenously given, the message m does not directly enter either player's payoff function. So we are dealing with "cheap talk." One of the main insights of the C-S paper is that even if the messages do not have any exogenous costs, R's equilibrium action rule generally creates endogenous signaling costs which allow equilibria with partial sorting. Specifically, they find that all the equilibria take a very simple form in which S partitions the state space Ω into subintervals and introduces noise into his messages by reporting, in effect, which element of the partition his observation of ω actually lies in. Moreover, for any $b > 0$, there is an upper bound, denoted by $N(b)$, on the size of an equilibrium (i.e., the number of subintervals of an equilibrium partition). There exists one equilibrium of each size from 1 through $N(b)$. Intuitively, the closer

S and R's interests are, the more information can be transmitted in equilibrium in the sense that $N(b)$ is non-increasing in b.

We depart from the C-S model by introducing two behavioral types into the model. On the sender's side, there is an "honest" type who always reports truthfully his observation of ω. On the receiver's side, there is a "naive" type who always blindly follows whatever message is sent to her. Their utility functions are:

$$\begin{aligned} u_h^S(m,\omega) &= -(m-\omega)^2 \\ u_n^R(a,m) &= -(a-m)^2. \end{aligned}$$

With $u_h^S(m,\omega)$, it is a dominant strategy for the sender to tell the truth (i.e., choose $m=\omega$) and with $u_h^R(a,m)$, it is a dominant strategy for the receiver to follow the message(i.e. choose $a=m$).[7]

Those players who have the same preferences as in the original C-S model we call "dishonest senders" and "sophisticated receivers", or "strategic types" in general.

Denote by θ the probability that the sender is honest and by λ the probability that the receiver is naive. We say that the sender has a two-dimensional type space $T=\Omega\times P$ where $P=\{honest, dishonest\}$. The receiver's type space is $Q=\{naive, sophisticated\}$. The two elements of the sender's type have independent probability distributions and are also independent of the receiver's type distribution.

[7] Of course, there are many other utility functions with which choosing "$m=\omega$" and "$a=\omega$" are dominant strategies. Choosing different utility functions won't change our results.

Let ρ denote the probability distribution on T. We assume that the message space M is the same as the state space Ω. Therefore $M = \Omega = [0,1]$. This seems to be a natural assumption with the existence of the naive receiver in the model. The receiver's action space is $A = \mathbb{R}$.

Here we have an extensive form game of incomplete information. It is not a standard signaling game because the receiver, as well as the sender, has private information on her type. We would like to "convert" it into a signaling game in order to apply some of the results in the literature that have been established for signaling games. This requires nothing more than a little redefinition because the naive type of receiver responds to the messages in a predetermined way and the only important role that the receiver's private information plays is to change the payoff function of the sender. Specifically, consider the following signaling game summarized by $\Gamma[(T,\rho), M, A, U^S, U^R]$.[8] In this game, player S first observes his type t from set T and then sends a signal m from the set M. Player R receives the signal m, infers player S's probable type and then selects an action a from the set A. The game ends and each player i receives payoff $U^i(t,m,a,b)$ $(i = S, R)$, which is defined as follows.

Define $I(p)$ as the indicator function such that

$$I(p) = \begin{cases} 1 & \text{if } p = \mathit{honest} \\ 0 & \text{if } p = \mathit{dishonest} \end{cases}$$

[8] We borrow this notation from Manelli (1996).

Also, define

$$\begin{aligned} U^S(t,m,a,b) &= \left(-\lambda(m-\omega-b)^2 - (1-\lambda)\,(a-\omega-b)^2\right) \\ &\quad (1-I(p)) - (m-\omega)^2 I(p) \\ \text{where } t &= (\omega,p) \in T. \end{aligned}$$

And

$$U^R(t,m,a,b) = -(a-\omega)^2.$$

To complete the specification of the game, assume that player R has prior beliefs ρ about the possible types t of player S and ρ is common knowledge.

Clearly, the only difference between the signaling game Γ and the game described earlier is that the naive receiver is not explicitly modeled in Γ. However, since we are interested in only the strategic types' equilibrium behavior, this difference is inconsequential and for our purpose, we can treat the two games as equivalent.

In the rest of the paper, we use Γ to refer to the game as described above and Γ^n to refer to a similar game with a different message space M^n to be introduced in the next section, that is, $\Gamma^n = \Gamma[(T,\rho), M^n, A, U^S, U^R]$. Game Γ and Γ^n are parameterized by λ, θ and b. Obviously, when $\lambda = \theta = 0$, we are back to the C-S model. We use Γ_{C-S} and Γ^n_{C-S} to refer to the C-S model with message space M and M^n, respectively. For notational convenience, $u^S_d(\omega,m,a,b)$ is used to denote the dishonest sender's payoff function: $u^S_d(\omega,m,a,b) = -\lambda(m-\omega-b)^2 - (1-\lambda)\,(a-\omega-b)^2$.

1.3 Monotonic Equilibrium in Infinite Game Γ, Finite Approximating Games Γ^n and the Cheap-talk Extension Game $\Gamma(L)$

Since the strategies for the honest sender and the naive receiver are given exogenously, we only need to find the equilibrium strategies for the strategic types. Let $m(\omega) : \Omega \to M$ be the dishonest sender's (pure) reporting strategy and $a(m) : M \to A$ be the sophisticated receiver's (pure) action strategy. (Note that since the sophisticated receiver has a strictly concave utility function, she never plays a mixed strategy in equilibrium.) In this section, we are going to focus on a very important class of equilibria, which we call monotonic equilibria.

Definition 1.1. *A monotonic equilibrium in* Γ *is a sequential equilibrium*[9] *in which* $m(\omega)$ *is weakly increasing in* ω.

So, we require the dishonest sender's strategy to be pure and non-decreasing in his observation of ω. (The honest sender's reporting strategy, by definition, is increasing in ω.) Given that both the sender and the receiver prefer higher actions for higher states, this seems to be a natural assumption to make, especially with

[9]Kreps and Wilson (1982) define sequential equilibrium only for finite games. Manelli (1996) adapts their definition to infinite signaling games and we are going to use his definition for the infinite games considered in this paper. The definition requires that the sender selects a best reponse for any type realization and that the receiver selects a best response to any message on and off the equilibrium path. No "inferior response" is allowed in equilibrium, even for a set of types/messages of measure zero. In Γ, since every message is sent on the equilibrium path, we are in effect using the solution concept of Bayesian Nash equilibrium with the restriction of interim optimality.

a fraction of receivers taking the messages literally. (We are going to discuss in section 1.6.2 what happens when we allow the sender to randomize and to have a non-increasing reporting strategy.) Since every $m \in M$ is sent on the equilibrium path, the sophisticated receiver's beliefs on the type distribution of the sender can always be derived by Bayes' rule.

The advantage of studying the continuous game Γ where $\Omega = M = [0,1]$ is that equilibria are relatively easy to characterize. The incentive constraints can be translated into a system of differential equations with boundary conditions whereas if Ω and/or M are discrete, the incentive constraints are much more complicated to analyze. However, there is no preexisting theorem that guarantees the existence of monotonic equilibrium (or, for that matter, sequential equilibrium) in game Γ. Indeed, we shall see that for certain parameter configurations of λ, θ and b, there is no monotonic equilibrium in Γ.

One way to get around this non-existence problem is to discretize the message space. In Theorem 1.1 below, we establish the existence of monotonic equilibrium in game Γ^n. Game Γ^n is the same as Γ except that Γ^n has the discrete message space M^n defined as follows.

Pick $n+1$ different real numbers $m_i (i = 0, 1, 2, .., n)$ from the interval $[0, 1]$ where $0 = m_0 < m_1 < m_2 < ... < m_{n-1} < m_n = 1$.

Define $M_n = \{m_0, m_1, m_2, ..., m_{n-1}, m_n\} = \{0, m_1, m_2, ..., m_{n-1}, 1\}$.

Theorem 1.1. *For any $b > 0$ and $\lambda, \theta \in (0, 1)$, there exists a monotonic equilibrium* *for game* Γ^n *if* $\max\{m_i - m_{i-1}\}_{i=1,\dots,n} \leq b$.

Theorem 1.1 states that for a game with discrete message space where the distances between adjacent messages are small enough, there always exists an equilibrium where the dishonest sender's reporting strategy is weakly increasing in his observation of ω.

Outline of proof. For details, see appendix.

The proof has two main steps.

First, by Kakutani's fixed point theorem, there exists a fixed point for the "restricted" best response correspondence where the dishonest sender's strategy is restricted to be weakly increasing in ω.

Second, we show that for the fixed point we found in the first step, the reporting strategy that corresponds to the fixed point is still a best response even without the monotonicity constraint, thus establishing the existence of monotonic equilibrium. □

Theorem 1.1 guarantees that there exists a monotonic equilibrium in finite games Γ^n that are close to the infinite game Γ. As we take an increasingly finer discretization of $[0, 1]$, the corresponding sequence of monotonic equilibrium outcomes will converge (in a subsequence) to a limit distribution. But the limit distribution may not be feasible in the limit game, i.e., no strategy profile in the limit game can generate this distribution, the result being non-existence.

To illustrate, consider the following example.

Example 1.1. *Suppose* $b = 0.05$.

In the C-S model, we have an equilibrium partition consisting of 3 *subintervals:* $[0, \frac{2}{15}], [\frac{2}{15}, \frac{7}{15}], [\frac{7}{15}, 1]$.

In our perturbed model, to make matters simple, we consider the limit case as λ *and* θ *go to* 0. *(In the following charaterization, the equilibria are found as the limit equilibria of* $\Gamma^n(\lambda, \theta)$ *when* λ *and* $\theta \to 0$.*)*

Consider a sequence of finite approximating games Γ^n *with* $M^n = \{0, \frac{1}{n}, \frac{2}{n}, ..., \frac{n-1}{n}, 1\}$. *For* Γ^n, *we are going to have equilibrium*[10] *strategies:*

$$m(\omega) = \begin{cases} \frac{n-2}{n} & \text{if } \omega \in [0, \frac{2}{15}) \\ \frac{n-1}{n} & \text{if } \omega \in [\frac{2}{15}, \frac{7}{15}) \\ 1 & \text{if } \omega \in [\frac{7}{15}, 1] \end{cases}$$

$$a(m) = \begin{cases} \frac{1}{15} & \text{if } m = \frac{n-2}{n} \\ \frac{3}{10} & \text{if } m = \frac{n-1}{n} \\ \frac{11}{15} & \text{if } m = 1 \end{cases}$$

These strategies, together with the prior distribution on T, *induce a joint probability distribution on* $\Omega \times M^n \times A$ *such that* $pr\{\omega \in [0, \frac{2}{15}), m = \frac{n-2}{n}, a = \frac{1}{15}\} = \frac{2}{15}$, $pr\{\omega \in [\frac{2}{15}, \frac{7}{15}), m = \frac{n-1}{n}, a = \frac{3}{10}\} = \frac{1}{3}$ *and* $pr\{\omega \in [\frac{7}{15}, 1], m = 1, a = \frac{11}{15}\} = \frac{8}{15}$. *As* $M^n \to [0, 1]$, *these outcomes will converge weakly to the distribution where* $pr\{\omega \in [0, \frac{2}{15}), m = 1, a = \frac{1}{15}\} = \frac{2}{15}$, $pr\{\omega \in [\frac{2}{15}, \frac{7}{15}), m = 1, a = \frac{3}{10}\} = \frac{1}{3}$ *and* $pr\{\omega \in [\frac{7}{15}, 1), m = 1, a = \frac{11}{15}\} = \frac{8}{15}$.

But no strategy pair of the limit game Γ *can generate this distribution. In order*

[10] As we will show in Proposition 1.5, this constitutes the unique limit equilibrium outcome of Γ^n as $\lambda, \theta \to 0$ for $b = 0.05$.

to generate this distribution, the dishonest sender must send $m = 1$ for all $\omega \in [0,1]$ and the sophisticated receiver must respond to $m = 1$ with $a = \frac{1}{15}$ when $\omega \in [0, \frac{2}{15})$, with $a = \frac{3}{10}$ when $\omega \in [\frac{2}{15}, \frac{7}{15})$ and with $a = \frac{11}{15}$ when $\omega \in [\frac{7}{15}, 1]$, which is obviously impossible.

In any finite approximating game, the "top three" messages in the finite message space are used to convey different information, but there are no "top three" messages in the infinite game Γ and therefore equilibrium breaks down in the limit.

The non-existence problem as illustrated above has been investigated in the literature. Manelli (1996) provides an ingenious way to "solve" the non-existence problem in infinite signaling games similar to Γ. The idea is simple: consider a variant of game Γ in which the sender, in addition to his message m, makes a non-binding suggestion of response $a \in A$ to the receiver. Adding such cheap-talk suggestions to the original game Γ restores existence of equilibrium. Going back to Example 1.1, we can see that this variant of game Γ has an equilibrium in which the dishonest sender sends $m = 1$ and suggests what the receiver should do (when $\omega \in [0, \frac{2}{15})$, suggests $a = \frac{1}{15}$, when $\omega \in [\frac{2}{15}, \frac{7}{15})$, suggests $a = \frac{3}{10}$, when $\omega \in [\frac{7}{15}, 1]$, suggests $a = \frac{11}{15}$), and the sophisticated receiver follows that suggestion.

The above idea can be formalized like this. Imagine the following (artificial) game. After observing his type $t \in T$, the sender sends, in addition to $m \in M$, a costless message $l \in L$. The receiver receives both m and l before choosing an action $a \in A$. The additional messages l are pure cheap talk – they do not affect the payoffs of any type of the players. Given the utility functions U^S and U^R, this means that there is no restriction on how the honest sender sends l, although he still tells the truth about ω through m. It also implies that the naive receiver simply ignores l in her decision and still chooses $a = m$. Following the terminology in Manelli(1996), we shall call this game the cheap-talk extension of game Γ, denoted by $\Gamma(L) = \Gamma((T, \rho), M, L, A, U^S, U^R)$.

In general, incorporating cheap talk into a game enlarges the set of equilibria. Of course, the set of available cheap-talk messages plays an important role in determining what results we get. As the space of L gets richer, the possibilities of communication increase and potentially, so does the the set of equilibrium outcomes. For the purpose of our analysis, we are going to focus on what is called *the canonical cheap-talk extension of* Γ in which the cheap-talk message space L is equal to the action space A. Our interpretation of the canonical cheap-talk extension game $\Gamma(A)$ is that the sender, in addition to m, also makes a non-binding recommendation of action $a \in A$ to the receiver. Further, we can focus on what is called a simple equilibrium in $\Gamma(A)$. A *simple equilibrium* of $\Gamma(A)$ is a sequential equilibrium in $\Gamma(A)$ in which the sender sends a message m and suggests a response a to the receiver, who then follows the suggestion on the equilibrium path.[11] We can without loss of generality study simple equilibrium in the canonical cheap-talk extension game $\Gamma(A)$ because Proposition 1

[11] A simple equilibrium has two requirements in addition to the sequential equilibrium condition. 1. The sophisticated receiver follows the sender's suggestions on the equilibrium path. 2. The sophisticated receiver's responses do not depend on the suggestions received off the equilibrium path.

in Manelli's (1996) paper guarantees that any sequential equilibrium outcome of any cheap-talk extension game $\Gamma(L)$ can be obtained as a simple equilibrium outcome of $\Gamma(A)$. To facilitate comparison with Γ, the equilibrium outcomes of $\Gamma(L)$ are defined on $T \times M \times A$.

Next, we state Theorem 1.2 which provides the link between finite games Γ^n and infinite game Γ through the cheap-talk extension game $\Gamma(A)$.

Theorem 1.2. *Let* $\Gamma^n = \left[(T,\rho), M^n, A, U^S, U^R\right], n = 1, 2, ...$ *be a sequence of games converging to the limit game* $\Gamma = \left[(T,\rho), M, A, U^S, U^R\right]$. *Suppose* $\hat{E}^n$ *is an equilibrium outcome of the game* Γ^n *and* $\left(\hat{E}^n\right) \to \hat{E}$. *Then,* $\hat{E}$ *is a simple equilibrium outcome of the canonical cheap-talk extension of the limit game* $\Gamma(A)$.

Proof. : Application of Theorem 2 in Manelli (1996). □

Since there always exists a monotonic equilbrium in the finite approximating games, as we have established in Theorem 1.1, it follows from Theorem 1.2 that there always exists a simple equilibrium in $\Gamma(A)$ where the dishonest sender has a reporting strategy $m(\omega)$ that is non-decreasing in ω.

We would like to interpret the cheap-talk extension game $\Gamma(A)$ in a non-literal fashion. We don't believe that it is a realistic depiction of the economic situations we want to model. It plays a largely auxiliary role in our analysis, enabling us to find the limit distribution of the equilibrium outcomes in the finite approximating games by working within the continuous framework.[12]

1.4 Characterization of Monotonic Equilibrium

In this section, we are going to first characterize monotonic equilibrium in the infinite game Γ, assuming existence. Then, we will discuss when monotonic equilibrium fails to exist in Γ and how adding cheap talk suggestions enlarges the set of equilibria and restores existence. We will also prove the uniqueness of monotonic equilibrium outcomes in the canonical cheap-talk extension game $\Gamma(A)$ for any fixed parameter values λ, $\theta \in (0,1)$ and $b > 0$.

[12] A recent paper by Kartik (2003) considers a model of communication through multiple channels. Specifically, there is one dimension of signaling where the signals have misrepresentation costs and there is a second dimension where the signals are costless. Formally, Kartik's model is very close to the cheap-talk extension of the continuous sender-receiver game we consider, with the probability of the honest sender being 0. The message m and the cheap-talk suggestion l in our model correspond to the "costly" and "costless" signals in Kartik's model, respectively.

Despite the formal similarity, the pure cheap-talk messages have completely different interpretations in the two papers. In Kartik's model, the cheap-talk messages are an intrinsic part of the communication process and they have direct interpretation in economic applications. In our model, the cheap-talk suggestions are a technical device that we use to solve the non-existence problem.

Kartik's model has the problem of multiple equilibria and he invokes an equilibrium refinement – the monotonic D1 criterion that imposes restrictions on beliefs off the equilibrium path to rule out "unreasonable" equilibria. Due to the existence of the honest type of sender in our model, every message is sent on the equilibrium path and we have uniqueness in the class of monotonic equilibrium without imposing any *ad hoc* restrictions on the receiver's beliefs. Both Kartik's model and ours select the most informative equilibrium in the C-S model as the only equilibrium that is robust to the perturbations we introduce. However, when the probabilities of the behavioral types are strictly positive, our models provide different predictions (for example, the non-monotonicity of $a(m)$). We also discuss what equilibria look like when the message space is discrete and the connection between the game with an infinite message space and its finite approximating versions.

1.4.1 Monotonic Equilibrium in Γ

The two propositions below provide a full characterization of monotonic equilibria in Γ. Proposition 1.1 summarizes a series of properties on the continuity, differentiability and shape of the equilibrium strategies $m(\omega)$ and $a(m)$. Proposition 1.2 presents the differential equations system translated from the incentive constraints under the differentiability conditions. Following the propositions, we give an example which illustrates what equilibrium strategies look like for a specific set of parameter values. Throughout this section, the maintained assumptions are $b > 0$, $\lambda, \theta \in (0, 1)$ and that $m(\omega)$ is weakly increasing in ω.

Before we proceed to the propositions, here are a few useful notations and facts.

Define $g(\cdot)$ as the inverse function of $m(\cdot)$ whenever $m(\cdot)$ is invertible.

Since ω is uniformly distributed on $[0, 1]$, the random variable m, defined by $m(\omega)$, has the following distribution function: $F_m(x) = pr(m(\omega) \leq x)$. When $m(\cdot)$ is invertible, $F_m(x) = pr(\omega \leq m^{-1}(x)) = g(x)$.

In addition, when $g(\cdot)$ is differentiable, the density function for m is well defined: $f_m(x) = g'(x)$.

Given that the sophisticated receiver's payoff function is $-(a - \omega)^2$, her optimal choice of action when receiving a message m is equal to the conditional expectation of ω, $E(\omega|m)$.

If $m < m(0)$, the sophisticated receiver infers that m was sent by the honest sender with probability 1. Hence $a(m) = E(\omega|m) = m$ when $m < m(0)$.

If $m \geq m(0)$ and $g(m)$ is well defined and differentiable, $a(m) = E(\omega|m) = \frac{\theta m + (1-\theta) g'(m) g(m)}{\theta + (1-\theta) g'(m)}$.[13]

Proposition 1.1. *In a monotonic equilibrium in game Γ, the equilibrium strategies $m(\omega)$ and $a(m)$ have the following properties:*

1. *$m(\omega)$ is continuous for $\omega \in [0,1]$ and $0 < m(0) < b$;*

2. *$m(\omega)$ is strictly increasing for $\omega \in [0, \overline{\omega})$; $m(\omega) = 1$ for $\omega \in [\overline{\omega}, 1]$;*

3. *$m(\omega)$ is differentiable on $(0, \overline{\omega})$;*

4. *$a(m)$ is continuous on $[0,1)$ and differentiable on $(m(0), 1)$;*

5. *If $m \leq m(0)$, $a(m) = m$; if $m \in (m(0), 1)$, $m > a(m) > g(m)$ and $a(m) < g(m) + b$; if $m = 1$, $a(m) = \frac{\overline{\omega}+1}{2}$.*

Proposition 1.1 establishes continuity and differentiability for the dishonest sender's message strategy $m(\omega)$. It says that in a monotonic equilibrium, there is separation

---

[13]This is how the posterior is computed. Given the honest sender's message strategy $m^h(\omega)$ ($m^h(\omega) = \omega$) and the dishonest sender's message strategy $m(\omega)$, we have two random variables m^h and m^d defined by $m^h = m^h(\omega)$ and $m^d = m(\omega)$.

The c.d.f. and the density function of m^h are the same as those of ω, which is uniformly distributed on $[0,1]$. Also, as shown in the main text, the c.d.f. of m^d is given by $g(\cdot)$ and its density function is $g'(\cdot)$.

With probability θ, the receiver observes the realization of m^h and with probability $(1-\theta)$, she observes the realization of m^d. Therefore, given the probability distributions of m^h and m^d, when the receiver observes m, she infers that with probability $\frac{\theta}{\theta+(1-\theta)g'(m)}$, she has observed the realization of m^h (i.e. the sender is honest) and with probability $\frac{(1-\theta)g'(m)}{\theta+(1-\theta)g'(m)}$, she has observed the realization of m^d (i.e., the sender is dishonest). Hence, the conditional expectation of ω given m is equal to $\frac{\theta m+(1-\theta)g'(m)g(m)}{\theta+(1-\theta)g'(m)}$.

for the dishonest sender with observation below a threshold value $\overline{\omega}$. For ω higher than $\overline{\omega}$, the dishonest sender pools at the upper bound of the message space, $m = 1$.

The action strategy of the sophisticated receiver, $a(m)$, is continuous except for a possible jump at $m = 1$. It is also differentiable on $(m(0), 1)$.

The message sent by the dishonest sender with the lowest observation $\omega = 0$ in equilibrium is bounded away from 0 but strictly below his ideal point b. At $m(0)$, the sophisticated receiver's response $a(m) = m$. This ensures the continuity of $a(m)$ at $m = 0$ since $a(m) = m$ for $m < m(0)$. It also implies that the posterior probability attached by the sophisticated receiver to the sender being dishonest is zero when the message received is $m(0)$. This in turn implies that $g'(m(0)) = 0$ (alternatively, $m'(m(0)) = \infty$).

For m between $m(0)$ and 1, the Proposition says that the message sent by the dishonest sender with observation ω is always strictly higher than ω and the equilibrium response he gets from the sophisticated receiver is also higher than ω, but lower than his ideal point $\omega + b$.

A detailed, step-by-step proof of Proposition 1.1 is in the appendix. Here we give an outline of the proof.

First, it is straightforward to see that $m(\omega)$ is continuous because otherwise there exists an $m' > m(0)$ such that no dishonest sender sends m' in equilibrium, $a(m') = m'$ and it is a profitable deviation for some type of dishonest sender to send m' and induce $a = m'$ from both types of receivers, a contradiction.

Next, under the assumptions that $m(\omega)$ is strictly increasing, $g(m)$ is differentiable and $a(m)$ is continuous, we show that $a(m)$ is differentiable and satisfies $\frac{da}{dm} = -\frac{\lambda(m-g(m)-b)}{(1-\lambda)(a(m)-g(m)-b)}$. The approach we use in the proof is similar to Mailath (1987). Basically, we expand $u_d^S(\omega', m(\omega''), a(m(\omega'')), b)$ in a Taylor series and use incentive compatibility constraints for ω'' close to ω' to establish that $\frac{du_d^S(\omega',m(\omega'),a(m(\omega')))}{da} \lim_{\omega''\to\omega'} \frac{a(m(\omega''))-a(m(\omega'))}{m(\omega'')-m(\omega')} + \frac{du_d^S(\omega',m(\omega'),a(m(\omega')))}{dm} = 0$, which under the assumptions implies that $-2(1-\lambda)(a(m)-g(m)-b)a'(m) - 2\lambda(m-g(m)-b) = 0$. Notice that this is just the F.O.C. for the reporting strategy $m(\omega)$ to be (locally) optimal, given that the sophisticated receiver's strategy is $a(m)$. By computing the S.O.C. for the optimality of $m(\omega)$, we find that $\frac{g'(m)(m-a(m))}{(a(m)-g(m)-b)} \leq 0$, which implies that either (i) $m = a(m)$ or (ii) $m > a(m) > g(m)$ and $a(m) < g(m) + b$ must hold.

The F.O.C. and the S.O.C. are derived under a number of monotonicity, continuity and differentiability constraints. We need to show that they are satisfied in equilibrium so that the F.O.C. and the S.O.C. apply.

Step by step, we show that if $m(\omega)$ is strictly increasing, then the continuity of the dishonest sender's equilibrium payoff function (which is implied by the incentive constraints) implies that $a(m)$ is continuous. The continuity of $a(m)$ implies the differentiability of $g(m)$. Also, we show that pooling of the dishonest sender cannot happen at any message m below 1 because otherwise we can find a message $m'(> m)$ close to m where $m(\omega)$ is strictly increasing and the S.O.C. for optimality of $m(\omega)$ is violated. Therefore, for $m \in [m(0), 1)$, $m(\omega)$ is strictly increasing, $g(m)$ is differentiable, $a(m)$ is continuous so that the F.O.C. and the S.O.C. apply.

It is easy to show that $a(m(0)) = m(0)$. Suppose not. Then $a(m(0))$ (a weighted average of $m(0)$ and $0) < m(0)$ and there exists $\varepsilon > 0$ such that it is a profitable deviation for the dishonest sender with observation 0 to send $m(0) - \varepsilon$ and induce both types of the receiver to respond with $m(0) - \varepsilon > a(m(0))$.

To show that $m(0)$ is bounded away from 0 and is strictly lower than b, we again use proof by contradiction: if $m(0) = 0$ or if $m(0) \geq b$, then there always exists a profitable deviation for the dishonest sender with observation $\omega = 0$.

From the F.O.C. $a'(m) = -\frac{\lambda(m-g(m)-b)}{(1-\lambda)(a(m)-g(m)-b)}$, we know that whether $a(m)$ is increasing or decreasing depends on the signs of $(m - g(m) - b)$ and $(a(m) - g(m) - b)$. Since $a(m(0)) = m(0) < b$, the right derivative of $a(m)$ at $m(0)$ is negative and $a(m)$ starts to decrease at $m(0)$. For the range that $a(m)$ is decreasing, it is obvious that $a(m) < g(m) + b$ since $g(m)$ is increasing. In addition, if $a(m)$ is increasing, it must be the case that $m > g(m)+b$ and $a(m) < g(m)+b$. Hence, for $m \in (m(0), 1)$, we have $m > a(m) > g(m)$ and $a(m) < g(m) + b$. Since $a(m) = \frac{\theta m+(1-\theta)g'(m)g(m)}{\theta+(1-\theta)g'(m)}$, $a(m) > m$ implies that $g'(m) \neq 0$ and $m'(\cdot)$ is well defined for $m \in (m(0), 1)$.

Proposition 1 is a summary of these results.

Now define $g(1) = \lim_{m\to 1^-} g(m)$ and $g'(1) = \lim_{m\to 1^-} g'(m)$.

For a fixed strategy $m(\omega)$ that satisfies properties $(1) - (3)$ in the previous proposition, define

$$E_{m(\omega)}(\omega|m)=\begin{cases} m & \text{if } m\in[0,m(0)) \\ \frac{\theta m+(1-\theta)g'(m)g(m)}{\theta+(1-\theta)g'(m)} & \text{if } m\in[m(0),1) \\ \frac{1+g(1)}{2} & \text{if } m=1. \end{cases}$$

Fixing the dishonest sender's strategy $m(\omega)$, $E_{m(\omega)}(\omega|m)$ is the sophisticated receiver's conditional expectation of ω when she receives a message m. $E_{m(\omega)}(\omega|m)$ is a function in m.

Proposition 1.2. *The strategy profile* $(m(\omega), a(m))$ *constitutes a monotonic equilibrium strategy profile in game* Γ *if and only if* $m(\omega)$ *and* $a(m)$ *satisfy properties* $(1)-(5)$ *in Proposition* 1.1 *and* $g(m)$, $a(m)$ *are a solution to the following system of differential equations with two boundary conditions at* $m=m(0)$ *and* $m=1$*:*

1. *For* $m\in[m(0),1)$,

$$a(m)=E_{m(\omega)}(\omega|m)$$

2. *For* $m\in[m(0),1)$,

$$\lambda(m-g(m)-b)+(1-\lambda)(a(m)-g(m)-b)a'(m)=0$$

3.

$$a(m(0))=m(0)$$

4.

$$\lim_{m \to 1^-} \left(a\left(m\right) - g\left(m\right) - b\right)^2 = \left(\frac{1+g(1)}{2} - g\left(1\right) - b\right)^2$$

Sketch of Proof. Necessity is quite clear since all the conditions come from the incentive constraints. Condition 4 is an implication of the continuity of the dishonest sender's equilibrium payoff: $u_d^S\left(\overline{\omega}, m\left(\overline{\omega}\right), a\left(m\left(\omega''\right)\right)\right) \to u_d^S\left(\overline{\omega}, m\left(\overline{\omega}\right), a\left(m\left(\overline{\omega}\right)\right)\right)$ as $\omega'' \to \overline{\omega}$. The function $a\left(m\right)$ doesn't have to be continuous at $m = 1$. It may jump at $m = 1$ and in that case, $a\left(1\right) = \frac{1+g(1)}{2} > g\left(1\right) + b$.

For sufficiency, we need to check that there is no profitable deviation once these conditions are satisfied. It is clear that there is no profitable deviation for the receiver. For the dishonest sender, we need to show that $m\left(\omega\right)$ globally maximizes his utility, for all $\omega \in [0, 1]$. For details of the proof, see appendix. □

We cannot solve the differential equations system analytically. The following example gives us a concrete idea as to what equilibrium strategies look like.

Example 1.2. *Suppose* $b = 0.5, \lambda = 0.2, \theta = 0.2.$

The figures below illustrate the strategies for the strategic players in a monotonic equilibrium of Γ *with* $b = 0.5$ *and* $\lambda = \theta = 0.2.$

As we can see from figure 1.1, the dishonest sender adopts an increasing message strategy $m\left(\omega\right)$*: for* $\omega \in [0, 0.125)$*, he reports a (strictly) higher message when he observes a (strictly) higher state of the world; for* $\omega \in [0.125, 1]$*, he reports the highest message in the message space(* $m = 1$*). Observe that* $m\left(\omega\right)$ *is above the* $45°$ *line, so* *the dishonest sender always exaggerates his claims no matter what his observation of ω is.*

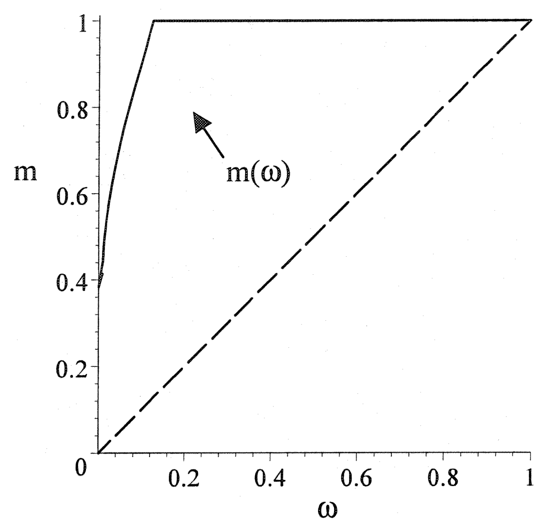

Figure 1.1: $m(\omega)$

Figure 1.2 *shows that the sophisticated receiver's strategy* $a(m)$ *has an interesting shape.*

Since $m(0) = 0.3835$, *if the sophisticated receiver receives a message* m *that is below* 0.3835, *she infers that it was sent by the honest sender with probability* 1 *and in that case, her optimal action would be equal to* m. *Therefore, for* $m \in [0, 0.3835)$, $a(m)$ *coincides with the* $45°$ *line.*

On the other hand, if she receives a message $m \in [0.3835, 1)$, *she infers that it could be sent by either the honest sender with observation* $\omega = m$ *or the dishonest sender with observation* $\omega = g(m)$. *In that case, the optimal action for the sophisti-*

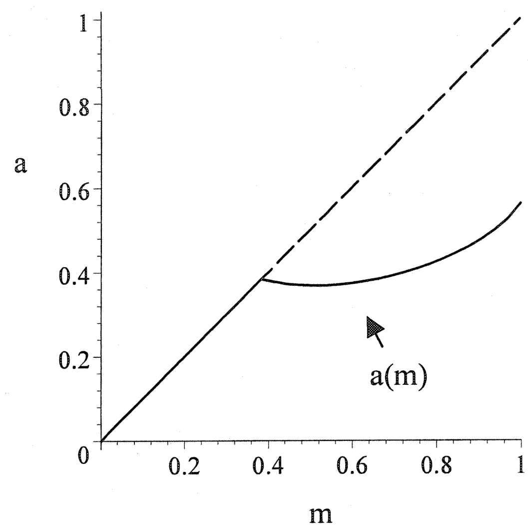

Figure 1.2: $a(m)$

cated receiver is equal to a weighted average of m *and* $g(m)$*. Intuitively, if* m *is high enough (in this example, if* $m > 0.3835$*), the sophisticated receiver doubts the truthfulness of the message she receives. Since the dishonest sender always exaggerates, the sophisticated receiver's optimal response is not as high as* m*. For* $m \in (0.3835, 1)$*,* $a(m)$ *is below the* $45°$ *line.*

Since the dishonest sender with observation of ω *between* 0.125 *and* 1 *reports* $m = 1$ *and the honest sender reports* $m = 1$ *only when* $\omega = 1$*, the sophisticated receiver infers that the message is sent by the dishonest sender with probability* 1 *when she receives* $m = 1$*. Therefore,* $a(1) = \frac{1+g(1)}{2} = \frac{1+0.125}{2} = 0.5625$*. For the parameter values we have chosen in this example,* $a(m)$ *is continuous at* $m = 1$*.*

One intriguing feature of $a(m)$ *is its non-monotonicity. In this example,* $a(m)$

coincides with the $45°$ *line and is strictly increasing for* $m \in [0, 0.3835)$. *However,* $a(m)$ *becomes a decreasing function in* m *when* m *exceeds* 0.3835 *until* $a(m)$ *reaches a local minimum at* $m = 0.5147$. *Then,* $a(m)$ *starts to rise again.*

From condition 2 *in Proposition 1.2, we know that* $a'(m) = \frac{-\lambda(m-g(m)-b)}{(1-\lambda)(a(m)-g(m)-b)}$. *Since we have established in Proposition 1.1 that* $a(m) < g(m) + b$, *the sign of* $a'(m)$ *is the same as the sign of* $(m - g(m) - b)$. *In this example, when* $\omega \in [0, 0.0147)$ *(and correspondingly,* $m \in [0.3835, 0.5147)$*), we have* $m(\omega) < \omega + b$ *and* $a'(m) < 0$, *which implies that a higher message leads to a lower action by the sophisticated receiver. Intuitively, when* $m(\omega) < \omega + b$, *the dishonest sender has an incentive to distort his messages further to induce the naive receiver to choose an action that is even more favorable to him. But more distortion is not optimal in equilibrium because it will also result in a worse response from the sophisticated receiver. On the other hand, when* $\omega \in [0.0147, 0.125)$ *(and correspondingly* $m \in [0.5147, 1)$*), we have* $m(\omega) > \omega + b$ *and* $a'(m) > 0$. *This is the opposite of the previous case. For the dishonest sender, more distortion will result in a more favorable response from the sophisticated receiver but it will also induce the naive receiver to choose an action which is further away from the sender's ideal point. These two marginal effects work in opposite directions and cancel out at the optimum.*

The sender's bias is quite large in this example ($b = 0.5$) *and there is no informative equilibrium in the C-S model. However, as we see from this example, when a positive fraction of the players are the behavioral types, the dishonest sender follows a* *strictly increasing reporting strategy for* $\omega \in [0, 0.125)$ *and there is some information being transmitted in equilibrium.*

The problem of non-existence of monotonic equilibrium in Γ arises when there is no solution to the differential equations system as specified in Proposition 1.2 and this happens, loosely speaking, when λ and θ are close to 0 and b is small.

To see why, let's take a closer look at the differential equations system 1– 4 in Proposition 1.2. To find a solution to the system, we can imagine fixing $m(0) = m_0$ where m_0 is between 0 and b. The initial conditions at m_0 are $g(m_0) = 0$ and $a(m_0) = m_0$. Together with these initial value conditions, the differential equations $a(m) = \frac{\theta m + (1-\theta) g'(m) g(m)}{\theta + (1-\theta) g'(m)}$ and $\lambda(m - g(m) - b) + (1-\lambda)(a(m) - g(m) - b) a'(m) = 0$ will determine the trajectories of $g(m)$ and $a(m)$ on $[m_0, 1)$. If we have chosen the "right" m_0, then the paths of $g(m)$ and $a(m)$ will lead to $\lim_{m \to 1^-} (a(m) - g(m) - b)^2 = \left(\frac{1+g(1)}{2} - g(1) - b\right)^2$ and we arrive at a solution to the differential equations system, i.e., we find a monotonic equilibrium in Γ. However, this is not feasible for all parameter values. When λ and θ are close to 0, no matter what m_0 we choose, $g(1)$ is going to be close to 0. From Proposition 1.1, we know that $g(m) < a(m) < g(m) + b$ for $m < 1$, and therefore $g(1) \leq \lim_{m \to 1^-} a(m) \leq g(1) + b$. Clearly, when $g(1)$ is close to 0 (which implies that $a(1) = \frac{1+g(1)}{2}$ is close to $\frac{1}{2}$) and b is small, condition 4 cannot be satisfied and there is no solution to the differential equation system implied by the equilibrium conditions, resulting in the non-existence of monotonic equilibrium.

1.4.2 Monotonic Equilibrium in the Canonical Cheap-talk Extension Game $\Gamma(A)$

In section 1.3, we have laid out the theory that guarantees the existence of monotonic equilibrium in the canonical cheap-talk extension game $\Gamma(A)$. What makes the difference when we add cheap talk suggestions to the original game Γ? As we shall see below, many of the properties that we established for equilibrium strategies $m(\omega)$ and $a(m)$ in Γ are still true in $\Gamma(A)$. It is still the case that $m(\omega)$ is strictly increasing for $m < 1$ in equilibrium. Cheap-talk suggestions have no effect on how information is transmitted when $m(\omega)$ is strictly increasing. However, they may be effective at $m = 1$ where the dishonest sender with different observations of ω pools. Mathematically, the existence of monotonic equilibrium is restored in $\Gamma(A)$ because adding cheap talk suggestions may change the boundary condition at $m = 1$ so that there exists a solution to the new differential equations system implied by the equilibrium conditions in $\Gamma(A)$.

In $\Gamma(A)$, the cheap-talk space is $L = A$ and the sender sends both a regular message $m \in M$ and a cheap-talk suggestion $l \in A$ to the receiver. So the strategy for the sender consists of two parts: a message strategy $s(t) : T \to M$ and a cheap-talk suggestion strategy $l(t) : T \to A$. For notational convenience and consistency, we use $m(\omega)$ to denote the dishonest sender's message strategy, i.e., $m(\omega) = s((\omega, dishonest))$. Also, we use $m^h(\omega)$ to denote the honest sender's message strategy. By definition, $m^h(\omega) = \omega$. But the cheap-talk suggestion strategy for the honest sender, $l((\omega, honest))$, is determined endogenously in equilibrium. Under our assumptions on the payoff functions, the naive receiver ignores the cheap-talk suggestions and always chooses $a = m$. As to the sophisticated receiver's strategy in $\Gamma(A)$, it is a mapping from $M \times L$ to A. Let's call it $r(m, l)$.

Define a monotonic simple equilibrium as a simple equilibrium in $\Gamma(A)$ where $m(\omega)$ is weakly increasing in ω. As we have established in section 1.3, we can without loss of generality focus on simple equilibria in $\Gamma(A)$. The main advantage of studying simple equilibria is clarity: we can pin down exactly what the cheap-talk suggestions are on the equilibrium path, which makes our analysis easier. The following lemma is an illustration of this property.

Lemma 1.1. *Suppose in a monotonic simple equilibrium in* $\Gamma(A)$, $\{\omega : m(\omega) = m'\}$ *is a singleton for* $m' \in [m(0), 1]$. *Let* $\omega' = m^{-1}(m')$. *We must have* $l((\omega', dishonest)) = l((m', honest))$ *and they are equal to* $E_{m(\omega)}(\omega | m')$.

Proof. By contradiction. Suppose $l((\omega', dishonest)) \neq l((m', honest))$. Since we are in a simple equilibrium and $\{\omega | m(\omega) = m'\}$ is a singleton, it must be the case that $l((m', honest)) = m'$ and $r(m', m') = m'$, i.e., the honest sender with observation $\omega = m'$ sends message $m = m'$ and also recommends the receiver to take action $a = m'$ and the sophisticated receiver follows that recommendation in equilibrium. If $m' > b$, it is a profitable deviation for the dishonest sender with observation $\omega = m' - b$ to send (m', m') and incur the response $a = m'$ from both the sophisticated and the naive receivers. If $m' \leq b$, then sending (m', m') is a profitable deviation for the dishonest sender with observation $\omega = 0$, a contradiction.

Let $l' = l((\omega', dishonest)) = l((m', honest))$. In equilibrium, $r(m', l') = E_{m(\omega)}(\omega|m')$. Since in a simple equilibrium the rational receiver follows the sender's suggestion on the equilibrium path, we must have $l' = E_{m(\omega)}(\omega|m')$. □

From this lemma, we know that adding cheap-talk suggestions doesn't change the incentive constraints when $m(\omega)$ is strictly increasing. Therefore, the properties we derived from the incentive compatibility constraints in Γ still apply here.

In a monotonic simple equilibrum in $\Gamma(A)$, only the honest sender sends messages that are below $m(0)$. For $\omega < m(0)$, $m^h(\omega) = \omega$ and $l((\omega, honest)) = \omega$. In this case, the cheap-talk suggestions are in effect ignored by the sophisticated receiver. In light of lemma 1.1, we can see that when $m(\omega)$ is strictly increasing, the cheap-talk suggestions are also effectively ignored by the sophisticated receiver. Hence, under cases (i) $m < m(0)$ and (ii) $m \geq m(0)$ and $m(\omega)$ is strictly increasing, we have a well defined function $a(m)$ that gives the sophisticated receiver's equilibrium response to m when the pair (m, l) that has been received is on the equilibrium path. In other words, $a(m') = r(m', l')$ if $l' = E_{m(\omega)}(\omega|m')$.

Since the incentive constraints do not change in $\Gamma(A)$ under the assumption that $m(\omega)$ is strictly increasing, the results we found in game Γ are still true in $\Gamma(A)$. Specifically, in a monotonic simple equilibrium in $\Gamma(A)$, if $m(\omega)$ is strictly increasing, then $g(m)$ and $a(m)$ are differentiable, $a'(m) = -\frac{\lambda(m-g(m)-b)}{(1-\lambda)(a(m)-g(m)-b)}$, $a(m) < g(m)+b$ and either $m = a(m)$ or $m > a(m) > g(m)$.

What happens when $m(\omega)$ is not strictly increasing, in other words, when the dishonest sender with different observations of ω send the same message? How does adding cheap-talk suggestions affect the information transmission process in this case? Imagine we have $\omega_0 < \omega_I$ and $m(\omega) = m'$ for $\omega \in [\omega_0, \omega_I]$. In game Γ, this means that all $\omega \in [\omega_0, \omega_I]$ are pooled together and $a(m') = \frac{\omega_0+\omega_I}{2}$ whereas in game $\Gamma(A)$, the dishonest sender with different observations of ω on $[\omega_0, \omega_I]$ may use different cheap-talk suggestions to separate themselves as long as incentive constraints are satisfied.

If we focus on the part where $m(\omega) = m'$ for $\omega \in [\omega_0, \omega_I]$ in $\Gamma(A)$, we are in effect back to a Crawford-Sobel situation: a strategic sender conveys his private information on $\omega \in [\omega_0, \omega_I]$ to a strategic receiver through pure cheap talk l; the behavior of the non-strategic types can be ignored because the naive receiver's response is fixed at $a = m'$ and the honest sender occurs with probability 0. Using the insight from Crawford-Sobel, we know that cheap talk is not necessarily ignored and information can be conveyed in equilibrium through certain partitions of $[\omega_0, \omega_I]$ if b is small relative to the size of the interval $[\omega_0, \omega_I]$. Specifically, suppose there are I subintervals in the partition: $[\omega_0, \omega_1], [\omega_1, \omega_2], ..., [\omega_{I-1}, \omega_I]$, then the equilibrium conditions require that $(1) u_d^S\left(\omega_i, m', \frac{\omega_{i-1}+\omega_i}{2}, b\right) = u_d^S\left(\omega_i, m', \frac{\omega_i+\omega_{i+1}}{2}, b\right)$ and (2) $l(\omega, dishonest) = \frac{\omega_{i-1}+\omega_i}{2}$, $r\left(m', \frac{\omega_{i-1}+\omega_i}{2}\right) = \frac{\omega_{i-1}+\omega_i}{2}$ for $\omega \in (\omega_{i-1}, \omega_i)$, $i = 1,..,I$.[14] Condition (1) is from

[14]There are different ways to specify the equilibrium behavior for the dishonest sender with observation $\omega = \omega_i, i = 1, ..., I-1$ because he is indifferent between actions $\frac{\omega_{i-1}+\omega_i}{2}$ and $\frac{\omega_i+\omega_{i+1}}{2}$. However, different specifications will not change the equilibrium outcome since the set of ω_i's are of measure zero.

Crawford-Sobel. It is the indifference condition for the types on the end-points of the subintervals. Condition (2) comes from the definition of a simple equilibrium in $\Gamma(A)$. It says that for the dishonest sender with observation $\omega \in (\omega_{i-1}, \omega_i)$, he sends message m' and makes a suggestion of action $\frac{\omega_{i-1}+\omega_i}{2}$ to the receiver and the rational receiver follows the suggestion on the equilibrium path.

In the previous section, we have shown that in Γ, pooling for the dishonest sender with different observations of ω can happen only at the highest message $m = 1$ in equilibrium. Here we want to show that this is also true for game $\Gamma(A)$. In other words, even with cheap talk suggestions, it is not possible, in a monotonic simple equilibrium of $\Gamma(A)$, for the dishonest sender with different observations of ω to send the same message if that message is below 1. We can borrow the argument we had for game Γ. Imagine in a monotonic simple equilibrium in $\Gamma(A)$,there exist $\omega_0 < \omega_I$, $m' < 1$ and $m(\omega) = m'$ for $\omega \in [\omega_0, \omega_1]$. Then there exists $\varepsilon > 0$ such that for $m'' \in (m', m' + \varepsilon)$, $m(\omega)$ is strictly increasing. As we have established in lemma 1.1, the equilibrium response of the sophisticated receiver is $a(m'') = E_{m(\omega)}(\omega|m'')$ for $m'' \in (m', m' + \varepsilon)$. Incentive compatibility constraints require that $u_d^S\left(\omega_I, m', \frac{\omega_{I-1}+\omega_I}{2}, b\right) = \lim_{m'' \to m'} u_d^S(\omega_I, m', a(m''), b)$. Since $a(m'') < g(m'') + b$, as shown earler, it must be true that $\frac{\omega_{I-1}+\omega_I}{2} = \lim_{m'' \to m'} a(m'')$, which in turn implies that $m'' < \omega_I$, but this contradicts the property that $m'' \geq a(m'')$. Hence, we have the following lemma.

Lemma 1.2. *In a monotonic simple equilibrium in $\Gamma(A)$, $m(\omega)$ is strictly increasing in ω for $\omega \in [0, \overline{\omega})$, where $\overline{\omega} = \inf\{\omega : m(\omega) = 1\}$.*

So, adding cheap-talk suggestions can only change what happens at $m = 1$ in equilibrium.

Call the system of equations $\{u_d^S(\omega_i, m, \frac{\omega_{i-1}+\omega_i}{2}, b) = u_d^S(\omega_i, m, \frac{\omega_i+\omega_{i+1}}{2}, b)\}_{i=1,...,I-1}$ (A) and say that a sequence $\{\omega_0, \omega_1, ..., \omega_I\}$ where $\omega_i \in \Omega$ and $\omega_I = 1$ is a solution to (A) if either of the following two cases holds: (i) $I = 1$, (ii) $I \geq 2$ and $u_d^S(\omega_i, m, \frac{\omega_{i-1}+\omega_i}{2}, b) = u_d^S(\omega_i, m, \frac{\omega_i+\omega_{i+1}}{2}, b)$ for $i = 1, ..., I-1$.

Proposition 1.3. *In a monotonic simple equilibrium in $\Gamma(A)$, the equilibrium strategies satisfy the following conditions:*

1. *$m(\omega)$ and $a(m)$ satisfy the properties 1-5 in Proposition 1.1.*
2. *There exist an integer $I \geq 1$ and a sequence $\{\omega_0, \omega_1, ..., \omega_I = 1\}$ which is a solution to (A) such that $g(m)$ and $a(m)$ are a solution to the following system of differential equations:*

   (a) *For $m \in [m(0), 1)$,*

   $$a(m) = E_{m(\omega)}(\omega|m)$$

   (b) *For $m \in [m(0), 1)$,*

   $$\lambda(m - g(m) - b) + (1 - \lambda)(a(m) - g(m) - b)a'(m) = 0$$

   (c) *$a(m(0)) = m(0)$*

   (d) *$g(1) = \omega_0$*

(e) $\lim_{m\to 1^-}(a(m)-g(m)-b)^2=\left(\frac{\omega_0+\omega_1}{2}-g(1)-b\right)^2$

3. $l(\omega, dishonest)=\begin{cases} E_{m(\omega)}(\omega|m=m(\omega)) & \text{if } \omega\in[0,\omega_0) \\ \frac{\omega_{i-1}+\omega_i}{2} & \text{if } \omega\in(\omega_{i-1},\omega_i),\ i=1,...I \end{cases}$

$l(\omega, honest)=E_{m(\omega)}\left(\omega|m=m^h(\omega)\right)$ *if* $\omega\in[0,1)$

$l(\omega, honest)\in\{\frac{\omega_{i-1}+\omega_i}{2}\}$ $i=1,...,I$ *if* $\omega=1$

4. *For* $m<1$, $r(m,l)=l$ *if* $l=E_{m(\omega)}(\omega|m)$ *and* $r(m,l')=r(m,l''),\forall(m,l')$, (m,l'') *if* $l',l''\neq E_{m(\omega)}(\omega|m)$.

For $m=1$, $r(m,l)=l$ *if* $l\in\{\frac{\omega_{i-1}+\omega_i}{2}\}_{i=1,...,I-1}$ *and* $r(m,l')=r(m,l'')$, $\forall(m,l')$, (m,l'') *if* $l',l''\notin\{\frac{\omega_{i-1}+\omega_i}{2}\}_{i=1,...,I-1}$.

The important change in the differential equations system, when we compare it with the one in Proposition 1.2, is the boundary condition at $m=1$, as can be seen in 2.(e). When cheap talk is effective, the dishonest sender with observation $\omega=g(1)=\omega_0$ induces response $a=\frac{\omega_0+\omega_1}{2}$ instead of $\frac{g(1)+1}{2}$ from the sophisticated receiver. A dishonest sender with different observations of $\omega(\geq\omega_0)$ sends the same message $m(\omega)=1$, but he can potentially induce different responses from the sophisticated receiver by making different cheap-talk suggestions. Intuitively, when cheap-talk suggestions are not ignored (and that's precisely when we have the non-existence problem in the original game Γ), they help determine the continuation of the game by indicating the direction to follow.

Condition 3 specifies the sender's suggestion strategy $l(t)$. Condition 4 says that the sophisticated receiver follows the sender's suggestions on the equilibrium path.

Off the equilibrium path, her responses do not depend on the cheap talk suggestions, as the definition of simple equilibrium requires.

As we can see from condition 2, by choosing a different I and/or a different sequence $\{\omega_0, \omega_1, ..., \omega_I = 1\}$, we have different boundary conditions at $m = 1$ and potentially, different solutions and different equilibria in $\Gamma(A)$. If that is the case, not only do we have a multiple equilibria problem, we also have to be concerned with a "lower hemi-continuity" problem. In other words, we need to consider whether a particular monotonic equilibrium outcome we find in the infinite cheap-talk extension game $\Gamma(A)$ is the limit distribution of equilibrium outcomes in some sequence of finite approximating games Γ^n. Fortunately, we can establish uniqueness of monotonic equilibrium outcome in $\Gamma(A)$, which makes the "lower hemi-continuity" problem vacuous.

Theorem 1.3. *Fix parameters $b > 0$ and $\lambda, \theta \in (0,1)$. There is only one monotonic equilibrium outcome in the cheap-talk extension game $\Gamma(A)$.*

Sketch of Proof. For details, see appendix.

There are two main steps in our proof.

First, suppose $(\hat{a}(m), \hat{g}(m))$ and $(\tilde{a}(m), \tilde{g}(m))$ are solutions to 2. (a) and 2. (b) in Proposition 1.3, with initial conditions $(\hat{a}(\hat{m}_0) = \hat{m}_0, \hat{g}(\hat{m}_0) = 0)$ and $(\tilde{a}(\tilde{m}_0) = \tilde{m}_0, \tilde{g}(\tilde{m}_0) = 0)$, respectively. Define $\hat{a}(1) = \lim_{m \to 1^-} \hat{a}(m)$ and $\tilde{a}(1) = \lim_{m \to 1^-} \tilde{a}(m)$. We can show that if $b > \hat{m}_0 > \tilde{m}_0$, then $\hat{a}(1) > \tilde{a}(1)$ and $\hat{g}(1) < \tilde{g}(1)$.

Second, we show that we cannot find two different sequences $\{\hat{\omega}_0, \hat{\omega}_1, ..., \hat{\omega}_{\hat{I}} = 1\}$

and $\{\tilde{\omega}_0, \tilde{\omega}_1, ..., \tilde{\omega}_{\tilde{I}} = 1\}$ that are solutions to (A) and also satisfy 2.*d*. and 2.*e* with $(\hat{a}(1), \hat{g}(1))$ and $(\tilde{a}(1), \tilde{g}(1))$.

Hence, there is only one monotonic simple equilibrium outcome in $\Gamma(A)$. Since any monotonic equilibrium outcome in $\Gamma(A)$ is also a monotonic simple equilibrium outcome in $\Gamma(A)$, we have proved the theorem. □

1.5 Model Predictions

1.5.1 Strictly Positive λ and θ

In the previous section, we established a series of properties on equilibrium strategies when a positive fraction of players are behavioral types. Here we want to explore the economic implications of these results further. Again, we focus on the class of monotonic equilibria.

Distorted claims. Quite intuitively, $m(\omega) > \omega$. Since the dishonest sender has an upward bias, he would like to manipulate the belief of the the receiver so that she would choose a higher action than what is optimal for herself. In the C-S model, such an attempt to fool the receiver is bound to fail in equilibrium because the receiver is strategically sophisticated. In contrast, if a fraction of receivers blindly follow the messages, the dishonest sender can gain from exaggeration. Therefore, in a monotonic equilibrium, the claims of the dishonest sender are always inflated.

Marginal effects of distortion at the optimum. When the dishonest sender with observation ω makes a strategic choice of what message to send, he weighs the response of the sophisticated receiver's against that of the naive receiver's. Consider the case where $m > \omega + b$. In this case, the dishonest sender makes an "extreme exaggeration"– the message he sends is not only higher than the true state of the world, but also exceeds his own ideal point. Clearly, making an even higher claim than m incurs a marginal cost: since $m > \omega + b$, the naive receiver's response to m is already too high for the sender and further distortion will mislead the naive receiver even more. However, there is also a marginal gain from more distortion because a higher message would result in a higher belief for the sophisticated receiver and induce her to choose an action that is closer to the sender's ideal point. At the optimal $m(\omega)$, these two effects cancel out.

Non-monotonicity of the receiver's response. The more interesting case is when $m < \omega + b$. When $m < \omega + b$, the dishonest sender is only making a "moderately" exaggerated claim and he has an incentive to exaggerate even more so that the naive receiver would choose an action closer to his ideal point. The dishonest sender is deterred from doing so at the optimal $m(\omega)$ because a higher message will result in a lower response from the sophisticated receiver ($a'(m) < 0$), which hurts the sender. How can we have a decreasing response function for the receiver when both types of sender are following an increasing reporting strategy? Recall that when the sophisticated receiver receives a message m, she infers that it could be sent by the honest sender with observation $\omega = m$ or by the dishonest sender with observation $\omega = g(m)$. Therefore, the sophisticated receiver's optimal response to message m is a weighted average of m and $g(m)$. Conditional on the sender being the honest or the dishonest type, a higher message implies a higher state of the world in a monotonic equilibrium. However, the sophisticated receiver may infer from a higher message that there is a higher probability that the sender is not telling the truth – the message is "too good to be true." Overall, the expectation (unconditional on the sender being honest or dishonest) of ω may be lower when the message is higher, resulting in a decreasing response function for the sophisticated receiver.[15]

The non-monotonicity of $a(m)$ holds for all parameter values $b > 0$ and $\lambda, \theta \in (0,1)$. Recall that $0 < m(0) < b$. Due to the continuity of $m(\cdot)$, this implies that there always exists a range of ω close to 0 such that in a monotonic equilibrium of Γ, when his observation of ω lies in this range, the dishonest sender makes a moderately exaggerated claim and the sophisticated receiver's response is decreasing in the messages she receives.

1.5.2 Limit Case as λ and θ Approach 0

As the probabilities of the behavioral types approach 0, the perturbed models Γ and Γ^n are arbitrarily close to the original C-S model. Quite naturally, the question arises as to whether all the equilibrium outcomes in the C-S model are limit equilibrium outcomes of Γ and Γ^n as λ and θ go to 0. To avoid the potential non-existence

[15] We show that $a(m)$ may be decreasing in m. However, when taking into account that the receiver is possibly naive, the strategic sender's expectation of the receiver's response to m is equal to $\lambda m + (1-\lambda) a(m)$ and it is increasing in m in a monotonic equilibrium.

problem in Γ, we will look for limit equilibria in the cheap-talk extension game $\Gamma(A)$ as λ and θ approach 0. As shown in the previous section, we have a rather clean characterization of monotonic equilibrium outcomes in $\Gamma(A)$, which gives the limit distribution of equilibrium outcomes in the converging finite games Γ^n. Later in this section, we'll also discuss what equilibria look like in the finite game Γ^n as λ and θ go to 0. With a discrete message space, we derive interesting and intuitive results about how messages are used in communication.

Unique Equilibrium Outcome

Fix $b > 0$. Consider two converging sequences $\{\lambda_i\}_{i=1}^{\infty}$ and $\{\theta_j\}_{j=1}^{\infty}$ where $\lambda_i \longrightarrow 0$ and $\theta_j \longrightarrow 0$. For each $\Gamma_{\lambda_i,\theta_j}(A)$, there is a unique monotonic equilibrium outcome defined on $\Omega \times M \times A$. Call it $\hat{E}_{\lambda_i,\theta_j}$. Since $\hat{E}_{\lambda_i,\theta_j}$ is defined on a compact metric space, as $\lambda_i \longrightarrow 0$, $\theta_j \longrightarrow 0$, it converges (in a subsequence) to a distribution $\hat{E}$ also defined on $\Omega \times M \times A$. Let $\hat{E}_{\Omega\times A}$ be the marginal distribution of $\hat{E}$ on $\Omega \times A$. Following the norm in the literature, we define the equilibrium outcomes in the C-S model only on the payoff relevant space: $\Omega \times A$. Denote by $\hat{E}_{C-S}$ an equilibrium outcome of the C-S model Γ_{C-S}[16]. Say that $\hat{E}_{C-S}$ is a limit monotonic equilibrium outcome of $\Gamma(A)$ if and only if there exist convergent sequences $\{\lambda_i\}_{i=1}^{\infty}$ and $\{\theta_j\}_{j=1}^{\infty}$ such that as $\lambda_i \longrightarrow 0$ and $\theta_j \longrightarrow 0$, $\hat{E}_{\Omega\times A} = \hat{E}_{C-S}$.

Recall that in Crawford and Sobel (1982), the equilibrium of the largest size,

[16] Note that an equilibrium outcome in Γ_{C-S} is also an equilibrium outcome in Γ^n_{C-S} because the equilibrium outcomes in a C-S model are defined only on the payoff relevant space $\Omega \times A$.

$N(b)$, is called the most informative equilibrium in Γ_{C-S}. We have the following proposition.

Proposition 1.4. *Only the most informative equilibrium outcome in* Γ_{C-S} *is a limit monotonic equilibrium outcome of* $\Gamma(A)$.

See appendix for the proof. The intuition behind this proposition is simple. Observe that among all the equilibria in Γ_{C-S}, only in the most informative equilibrium does the sender with the observation $\omega = 0$ have a higher equilibrium payoff than the payoff he could get if the receiver knew that $\omega = 0$. That is, the following inequality holds in the most informative equilibrium and not any other equilibrium: $-(a_1 - 0 - b)^2 \geq -b^2$, where a_1 is the equilibrium action of the receiver when she believes that ω belongs to the lowest subinterval of the equilibrium partition.[17] Recall that in any $\Gamma(A)$ with $\lambda, \theta > 0$, $0 < a(m(0)) = m(0) < b$ in equilibrium. Therefore, the equilibrium payoff for the dishonest sender with $\omega = 0$ is equal to

[17] This condition on equilibrium payoffs is not an algebraic coincidence but has its inner logic. Imagine fixing other parameters in Γ_{C-S} and varying b from large to small values. When b is large, only the babbling equilibrium exists. When b reaches a threshold value b_1, there exists an ω that partitions $[0,1]$ into subintervals $[0,\omega]$, $(\omega,1]$ and $-(a_1 - \omega - b)^2 = -(a_2 - \omega - b)^2$ where a_i is the receiver's best response if she believes that ω belongs to subinterval i. The threshold value b_1 is found by setting $\omega = 0$ and solving $-(0 - 0 - b)^2 = -\left(\frac{1}{2} - 0 - b\right)^2$. When $b = b_1$, the lowest type's equilibrium payoff is the same as the payoff he gets if identified as type 0. As b gets smaller than b_1, an informative equilibrium (with two subintervals) comes into existence and the lowest type prefers being identified as himself than playing the non-informative equilibrium and inducing action $\frac{1}{2}$. As b decreases, the cutoff point ω shifts to the right until it reaches a point where the lowest type is indifferent between being identified as type 0 and playing the informative equilibrium with two subintervals. If b decreases further, an equilibrium with three subintervals comes into existence. Again, in less informative equilibria (equilibria with one or two subintervals), a_1 is too high for the lowest type of sender and his expected equilibrium payoff is lower than the payoff he gets if identified as type 0. This argument carries on as b gets smaller and smaller. As we can see, it applies to more general settings than the uniform-quadratic case. In fact, if condition (M) in the C-S paper (p.1444) is satisfied, this condition on equilibrium payoff holds.

$-\lambda(m(0)-0-b)^2-(1-\lambda)(a(m(0))-0-b)^2 > -b^2$. Consider a sequence of games $\{\Gamma_{\lambda_i,\theta_j}(A)\}$. Since the inequality holds for any $\Gamma_{\lambda_i,\theta_j}(A)$ with $\lambda_i, \theta_j > 0$, it must be the case that in the limit as $\lambda_i \to 0, \theta_j \to 0$, the equilibrium payoff for the dishonest sender with $\omega = 0$ is at least as high as $-b^2$. The result follows.

The proposition says that if we perturb the original C-S model with small uncertainties about the players' preferences, then we have a much sharper prediction on equilibrium outcome than the C-S model does. Only the most informative equilibrium outcome in the C-S model is robust to the introduction of small probabilities of non-strategic behavior.

Standard extensive-form refinements that put restrictions on players' beliefs off the equilibrium path don't have much power in games with costless signals and there are a number of refinements in the literature that directly address the multiplicity problem in cheap talk games. Since the limit case of our model can be used as a refinement of equilibria in a pure cheap talk model, we would like to compare our approach with other refinements in the literature.

The idea of "neologism-proofness" (Farrell (1993)) is the following. Fix a sequential equilibrium in a sender-receiver game and consider a message "My type t is in the set X". This message is not used in equilibrium and therefore considered a neologism. Suppose the receiver hypothesizes that the message is sent by a type in X and so updates her belief by Bayes' rule conditioning on the sender's type being in X. If the receiver's best response given this belief is (strictly) preferred precisely by those types in the set X over what they would get in the putative equilibrium, then the neologism is deemed credible. An equilibrium is neologism-proof if there are no credible neologisms. Matthews, Okuno-Fujiwara and Postlewaite (1991) propose a closely related criterion called "announcement-proofness" which deals with a few conceptual inconsistencies in the "neologism-proofness" criterion. Perhaps the most debatable feature of these two concepts is that the test of credibility of a neologism or an announcement is relative to the putative equilibrium and therefore based on counterfactuals. A solution concept that does not involve counterfactuals is "credible message rationalizability", proposed by Rabin (1990). The premise of this concept is that agents are rational and they have a propensity to speak the truth and believe that others speak the truth but use the game's strategic incentives to check whether such behavior and belief are rational. So the credibility of a message is tested by the strategic incentives of the players with no reference to an equilibrium. All the above concepts have natural predictions in certain examples of sender-receiver games. However, none of them provides a satisfactory answer to the multiplicity problem in the C-S model. On the one hand, there is *no* equilibrium in the C-S model that satisfies the "neologism-proofness" or the "announcement-proofness" criterion. On the other hand, the concept of "credible message rationalizability" does not rule out any equilibrium in the C-S model.

The approach this paper takes is different from those discussed above. Instead of trying to rule out "unreasonable" equilibria from a conceptual standpoint, we incorporate an important missing element in the original C-S model: the possibility of players being honest/naive. Our analysis shows that the introduction of the behavioral types fundamentally changes the way the game is played. Indeed, under some natural assumptions, "unintuitive" equilibria in the C-S model are eliminated.

How Messages Are Used

The C-S model, like most models of cheap talk, deals mainly with the question of how much information can be transmitted in equilibrium and doesn't answer the question of how messages are used in communication. In our model, although the literal meanings of messages have a direct impact only on the behavioral types' actions, they also influence how strategic players encode and interpret messages endogenously. So we can make predictions about what particular messages are used to convey what information in equilibrium.

Again, fix $b > 0$. Consider a sequence of games $\{\Gamma^n_{\lambda_i,\theta_j}\}$ with discrete message space M_n. We assume that $M_n = \{m_0 = 0, m_1, ..., m_n = 1\}$, $\max\{m_i - m_{i-1}\}_{i=1,...,n} \leq b$. So there exists a monotonic equilibrium in $\Gamma^n_{\lambda_i,\theta_j}$ and M_n contains at least $N(b)$ distinct messages.

Let $\{[\omega_{k-1}, \omega_k]\}_{k=1,...,N(b)}$ be the partition of Ω in the most informative equilibrium in Γ^n_{C-S}. Now let λ_i and θ_j approach 0. The next proposition tells us how messages are used to convey information in the limit equilibrium.

Proposition 1.5. *In the limit monotonic equilibrium of* $\Gamma^n_{\lambda_i,\theta_j}$ *as* λ_i *and* θ_j *approach*

0, $m(\omega) = m_{n-(N(b)-k)}$ *if* $\omega \in (\omega_{k-1}, \omega_k)$, *for* $k = 1, 2, ..., N(b)$.

For proof, see appendix. The proposition restates that as the fractions of the behavioral types go to zero, the limit equilibrium outcome of the perturbed model corresponds to the most informative equilibrium outcome in the C-S model. Moreover, with a discrete message space, the dishonest sender who has an upward bias uses only the messages that are at the top in the message space in the limit equilibrium. This clustering of messages in the direction of the informed party's bias is consistent with what we observe in real life. For example, we can analyze the "grade inflation" problem in the context of a sender-receiver game.[18] Think of a university as the sender. It has superior information on its students' academic performances and conveys this information through reports on the students' grades to prospective employers, graduate schools, etc, who make hiring or admission decisions. The university may be biased in its students' favor, creating an incentive to inflate their grades. As predicted in our model, a concentration of high grades (for example, A, A- and B+) is commonly observed.

1.6 Discussion

So far, we have focused on monotonic equilibria. Under the assumptions that the dishonest sender is choosing a pure reporting strategy and the messages he sends to the receiver are non-decreasing in his observation of the state of nature, we find a set

[18] For a related but different analysis, see Ottaviani and Squintani (2002).

of tractable and well-behaved equilibria. Below, we discuss what happens when we relax these assumptions.

1.6.1 Mixed Strategies

In the C-S model, there are plenty of mixed strategy equilibria. For example, there exists a babbling equilibrium in which the sender randomizes over the full support of the message space with the same probability distribution no matter what ω he observes. Then, no message will change the receiver's belief from the prior and no information is conveyed in equilibrium. Call this a "full randomization" babbling equilibrium. For each size from 1 to $N(b)$, there is an equilibrium that has the "full randomization" feature. In an equilibrium like this, all types that belong to the same subinterval in the equilibrium partition randomize over the same set of messages and every message in the message space is sent by some type of sender in equilibrium.

With the introduction of behavioral types, the scope for randomization is much more limited. We do not have a formal proof that mixed strategy equilibria don't exist in the perturbed model, but we can provide some discussion as to why the degree of randomization is small and why the "full randomization" equilibria cannot exist as limit equilibria in the perturbed model.

Due to the strict concavity of her payoff function, the sophisticated receiver never randomizes. Now consider the indifference curves for the dishonest sender with observations ω_1 and ω_2. As illustrated in figure 1.3, the indifference curves cross at

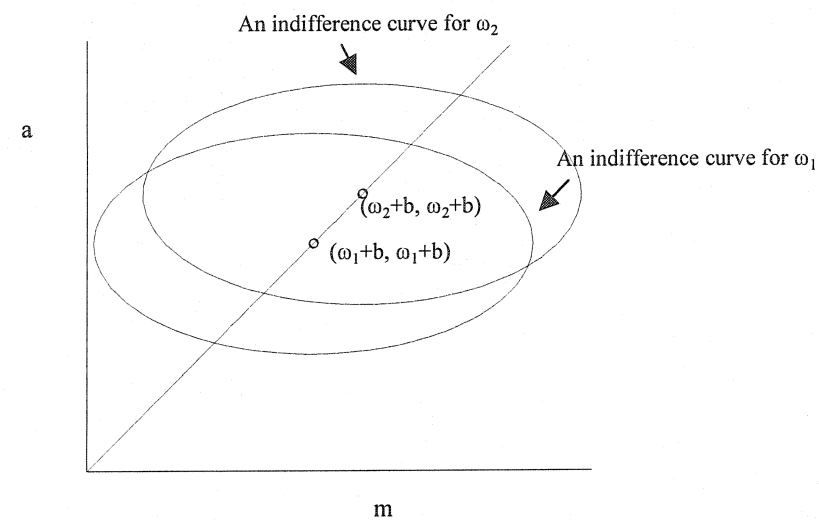

Figure 1.3: indifference curves for ω_1 and ω_2

most twice. Therefore, the support of the mixed strategies for ω_1 and ω_2 can have at most two common elements. Moreover, as we show in Lemma 1.3 in the appendix, if both types ω_1 and ω_2 prefer one message-action pair (m, a) to another pair (m', a'), then any type between ω_1 and ω_2 have the same preference over the two pairs as well. This implies that if two different messages m_1 and m_2 are in the support of the strategies for both types ω_1 and ω_2, then any type beween ω_1 and ω_2 must send either m_1 or m_2 in equilibrium. In this sense, the degree of randomization is quite small in the perturbed model. Intuitively, the fixed response of the naive receiver creates a differentiated "signaling cost" for the different types of senders so that indifference conditions are not as easily satisfied as in the unperturbed game. Mixed strategy equilibria like the "full randomization" equilibria where different types of sender randomize over a large set of messages in the C-S model are not robust to the introduction of behavioral types, no matter how small the perturbation is.

1.6.2 Non-increasing Message Strategy

In many signaling models, the monotonicity of the signaling strategy arises endogenously in equilibrium due to the assumption of the single-crossing property. However, in our model, $-\frac{\partial u_d^S(\omega,m,a,b)/\partial m}{\partial u_d^S(\omega,m,a,b)/\partial a} = -\frac{2\lambda(m-\omega-b)}{2(1-\lambda)(a-\omega-b)}$ is not monotonic in ω and the single-crossing property is violated, as the indifference curves in figure 1.3 show. This opens the door for non-increasing message strategies. For certain parameter values, we cannot rule out equilibria where the dishonest sender employs a reporting strategy that is not increasing in ω. For example, when λ and θ are small enough, we can construct an equilibrium in which the dishonest sender's message strategy is decreasing in ω, i.e., when he observes a high state, he reports a low message and vice versa.

The counter-intuitive predictions of the non-monotonic equilibria are in sharp constrast with those provided by the monotonic equilibria. In our model, the possibility that the sender is honest affects the inferences that the sophisticated receiver draws from the messages she receives and potentially lends more credibility to the claims that the sender makes. Moreover, the existence of naive receivers, whose choice of action can be easily manipulated by the sender's report, provides an opportunity for exploitation. In a monotonic equilibrium, the dishonest sender takes full advantage of this by distorting the messages in the direction of his bias. By contrast, in an equilibrium where the dishonest sender's message strategy is not increasing in ω, he may be "lying to his disadvantage." A dishonest sender who reports message m, as the equilibrium strategy prescribes, would be better off if the receiver were able to distinguish him from the honest type and identify him as a strategic player with observation $m^{-1}(m)$.[19] Oddly enough, in this equilibrium, the dishonest sender has an incentive to convince the receiver that he is not telling the truth. Furthermore, because the literal meaning of the message that the dishonest sender sends in equilibrium is far from, if not the opposite of, what he wants the receiver to believe, the naive receiver's literal-minded responses make the dishonest sender worse off as well.

1.7 Conclusion

In this paper, we enrich the strategic information transmission model introduced by Crawford and Sobel (1982) by incorporating two behavioral types – honest senders and naive receivers – into the original game. By modifying the C-S model in this simple and empirically plausible way, we present a more realistic picture of how people communicate through messages that have pre-existing, commonly understood meanings (e.g., a natural language).

We find that the existence of the behavioral types profoundly changes the way the game is played. Indeed, the prominent problem of multiple equilibria in the C-S

[19] It is familiar in cheap talk games that certain types of sender would be better off if they could distinguish themselves from the other types but cannot credibly do so in equilibrium. In the C-S model, this is true for types that are at the top of each subinterval in the equilibrium partition, if the sender has an upward bias.

and other cheap talk games, doesn't arise in our model when we consider the class of monotonic equilibria. Only the most informative equilibrium in the C-S model is robust to the perturbation of the behavioral types.

The C-S paper is celebrated for its elegant result about how much information can be transmitted between strategic players. It shows that full separation cannot happen in equilibrium when the interests of the sender and the receiver are not perfectly aligned, but some imprecise information can be conveyed by the sender's intentionally ambiguous messages if the two parties's interests do not differ too much. However, we cannot meaningfully discuss other important concepts such as "distortion" in the C-S model because the messages are completely extrinsic to the game. This paper provides a natural framework for such discussion. Uncertainty about the opponent's payoff/behavior, which seems to be a realistic characterization of many situations of communication, provides an explanation for commonly observed phenonmena such as exaggeration and the clustering of messages at one end of the message space. A somewhat puzzling aspect of communication, that sometimes people react to more aggressive claims with more conservative actions, is also accounted for in a simple way in our model.

1.8 Appendix

The following lemma is used in the proof of Theorem 1.1.

Lemma 1.3. *Consider sender types $t_1 = (\omega_1, dishonest)$ and $t_2 = (\omega_2, dishonest)$ where $\omega_1 < \omega_2$. If both types t_1 and t_2 prefer message/action pair (m_1, a_1) to (m_2, a_2), i.e., if $u_d^S(\omega_1, m_1, a_1, b) \geq u_d^S(\omega_2, m_2, a_2, b)$ for both t_1, t_2, then type $t_3 = (\omega_3, dishonest)$ where $\omega_1 \leq \omega_3 \leq \omega_2$ prefers (m_1, a_1) to (m_2, a_2) as well, i.e., $u_d^S(\omega_3, m_1, a_1, b) \geq u_d^S(\omega_3, m_2, a_2, b)$.*

Simple algebra shows that Lemma 1.3 is true.

Proof of Theorem 1.1. We prove Theorem 1.1 in two steps.

Step 1. Given $M_n = \{m_i\}_{i=0,1,\ldots,n}$, there are $(n+1)$ potential messages. A non-decreasing reporting strategy for the dishonest sender, $m(\omega) : \Omega \to M_n$, is a step function. Define $\sum^S = \{\mathbf{x} \in [0,1]^{n+2}, x_0 = 0, x_1 \leq x_2 \leq, \ldots, \leq x_n, x_{n+1} = 1\}$. Following Athey (2001), $m(\omega)$ can be represented by a vector $\mathbf{x} \in \sum^S$ according to an algorithm (detailed definition can be found in Athey (2001)) where each component of $\mathbf{x}$ is a "jump point" of $m(\omega)$. That is, x_i represents the value of ω at which the dishonest sender jumps from one message to the next higher message.

Now we define the dishonest sender's "restricted" best response correspondence. Think of the constrained maximization problem like this. Given the sophisticated receiver's action strategy $a(m)$, sender type $(1, dishonest)$ chooses his best response from the whole set of M_n. For the dishonest senders who have observed $0 \leq \omega < 1$, they have to choose a best response from M_n that is at least as low as the messages that are chosen by those who have observed higher signals. In the event of indifference, the highest message is chosen.

Formally, define the dishonest sender's "restricted" best response functions like this:

$$B_f(\omega|a(m)) = \begin{cases} \max\{\arg\max_{m\in M_n} u_d^S(\omega, m, a, b)\} & \text{for } \omega = 1 \\ \max\{\arg\max_{m\in M_n, m\leq B_f(\omega'|a(m)), \forall\omega'>\omega} u_d^S(\omega, m, a, b)\} & \text{for } 0 \leq \omega < 1 \end{cases}$$

Given the finiteness of M_n, a "restricted" best response always exists for any $\omega \in \Omega$. The way we treat indifference also guarantees uniqueness.

For the receiver's part, define $\sum^R = \{\mathbf{x} \in [0,1]^{n+1}, \mathbf{x} = (x_0, x_1, ...x_n)\}$. Note that the action strategy for the sophisticated receiver, $a(m) : M_n \to A$, can be represented by a vector $\mathbf{x} \in \sum^R$ where $x_i = a(m_i)$ for $i = 0, 1, ..., n$. The naive receiver's best response is defined in the obvious way. Because of the existence of the honest type of senders (given the utility function, their reporting strategy is to send the message in M_n that is the closest to ω), every $m \in M_n$ is sent on equilibrium path. The sophisticated receiver's posterior can be computed by Bayes' rule. Due to the strict concavity of her utility function, the best response is unique.

Define $\sum = (\sum^S, \sum^R)$. The best response correspondence as described above can be represented by $\sum \to \sum$.

Since $\sum$ is a compact, convex subset of $\mathbb{R}^{n+2+n+1}$, we can apply Kakutani's fixed point theorem.

Following Athey (2001) , we can verify that the best response correspondence is non-empty, convex and has a closed graph and therefore has at least one fixed point.

Step 2. Now we need to check that the fixed point(s) we found in the first step is an equilibrium of the game Γ^n provided that the distance between adjacent messages is less than b.

Denote the strategic players' strategies that are consistent with the fixed point found in step 1 by $(m^*(\omega), a^*(m))$. All we need to show is that $m^*(\omega)$ is a best response to $a^*(m)$.

Given the honest sender's utility function $u_h^S(m,\omega) = -(m-\omega)^2$, an honest sender with observation ω sends a message $m \in M_n$ that minimizes $|m-\omega|$. WLOG, assume that if the honest sender with observation ω is indifferent between sending m_{i-1} and m_i, then he sends m_i. Let $m^h(\omega)$ be the honest sender's message strategy and $a^h(m)$ be the sophisticated receiver's optimal response given that she believes the message m was sent by the honest sender with probability 1. So, $a^h(m) = \frac{\omega+\omega'}{2}$ where $\omega = \inf\{\omega : m^h(\omega) = m\}$ and $\omega' = \sup\{\omega : m^h(\omega) = m\}$.

We'll use the following properties later in the proof. First, $\min_{m\in M_n} |m - a^h(m')| = m', \forall m' \in M_n$. Second, given that $\max\{m_i - m_{i-1}\}_{i=1,\dots,n} \leq b$, it follows that $\max\{a^h(m_i) - a^h(m_{i-1})\}_{i=1,\dots,n} \leq b$ and $\max\{m_i - a^h(m_{i-1})\}_{i=1,\dots,n} \leq b$.

Given $a^*(m)$, a dishonest sender who has observed ω does not want to deviate and send a message $m < m^*(\omega)$ by definition of $m^*(\omega)$. Hence, we only need to show that $\forall\omega \in [0,1]$, the dishonest sender doesn't have an incentive to deviate and send

$m > m^*(\omega)$.

First, note that $m^*(1) = 1$. Suppose not, then $m^*(1) < 1$ and $a^*(1) = a^h(1) < a^*(m^*(1))$. Now suppose $\omega' = \inf\{\omega | m^*(\omega) = m^*(1)\}$ and the dishonest sender with observation ω' is indifferent between sending $m^*(\omega) = m^*(1)$ and sending $m' (< m^*(\omega))$. Since $a^*(m^*(1))$ is a weighted average of $1, \omega'$ and $a^h(m^*(1))$, $1 - \frac{b}{4} < a^h(1) < a^*(m^*(1))$ implies that $\omega' > 1 - \frac{b}{4}$. Since the dishonest sender with observation ω' is indifferent between sending $m^*(\omega)$ and m' where $|m' - \omega' - b| > |m^*(\omega') - \omega' - b|$, it has to be the case that $a^*(m') > a^*(m^*(\omega'))$, but this is not possible.

If $\forall m \in M_n \cap [m^*(0), 1]$, there exists an $\omega \in \Omega$ s.t. $m^*(\omega) = m$, then no type $\omega \in \Omega$ would want to deviate and send $m > m^*(\omega)$. We show this by contradiction. Suppose $m' \in \arg\max_{m \in M_n} \{u_d^S(\omega_0, m, a^*(m))\}$ and $m' > m^*(\omega_0)$. Define $\omega' = \sup\{\omega | m^*(\omega) = m'\}$. Then, it is true for $\omega = \omega_0$ and $\omega = \omega'$ that $\forall m (\in M_n) \leq m'$, $u_d^S(\omega, m', a^*(m')) \geq u_d^S(\omega, m, a^*(m))$. By Lemma 1.3, the inequality holds also for any $\omega \in [\omega_0, \omega']$. It follows from the definition of $m^*(\omega)$ that $m^*(\omega) = m'$ for any $\omega \in [\omega_0, \omega']$, but this contradicts the assumption that $m' > m^*(\omega_0)$.

What remains to be shown is that $\forall m \in M_n \cap [m^*(0), 1]$, there exists an $\omega \in \Omega$ s.t. $m^*(\omega) = m$.

Again, we show this by contradiction. Suppose not. Then, there exists an ω_1 s.t. $u_d^S(\omega_1, m_0', a^*(m_0')) = u_d^S(\omega_1, m_1', a^*(m_1'))$ (i.e., the dishonest sender with

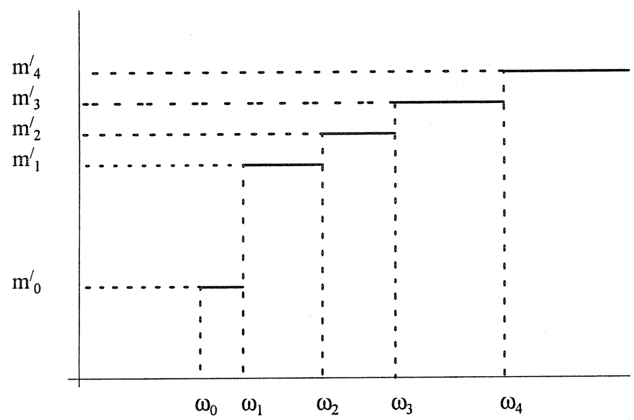

Figure 1.4: m_0' and m_1' are not adjacent

observation ω_1 is indifferent between sending m_0' and m_1') where $m_1' = m^*(\omega_1)$, $m_0', m_1' \in [m^*(0), 1] \cap M_n$, $m_0' < m_1'$ and m_0' and m_1' are not adjacent to each other in M_n. WLOG, we can assume that for any $m \in [m_1', 1] \cap M_n$, there $\exists\ \omega \in \Omega$ s.t. $m^*(\omega) = m$ and the dishonest sender with observation ω_i is indifferent between sending adjacent messages m_{i-1}' and m_i' that are higher than m_1'. See figure 1.4.

Next, we show that $\forall \tilde{m} \in M_n \cap (m_0', m_1')$, $\tilde{m} < \omega_1$. First, note that since given $m^*(\omega)$, $\tilde{m}$ is sent only by the honest sender, $a^*(\tilde{m}) = a^h(\tilde{m})$. It follows that $a^h(\tilde{m}) < \omega_1 + b$ because otherwise there exists $\omega = a^h(\tilde{m}) - b \geq \omega_1$ such that the dishonest sender with observation ω achieves a higher payoff by sending $\tilde{m}$ instead of $m^*(\omega)$, a contradiction.

Since $a^h(\tilde{m}) < \omega_1 + b$ and $\arg\min_m \left|m - a^h(\tilde{m})\right| = \tilde{m}$, it follows that $|\tilde{m} - \omega_1 - b| < |m_0' - \omega_1 - b|$. Since $u_d^S(\omega_1, m_0', a^*(m_0')) \geq u_d^S(\omega_1, \tilde{m}, a^*(\tilde{m}))$, it must be the case that $|a^*(m_0') - \omega_1 - b| < \left|a^h(\tilde{m}) - \omega_1 - b\right|$. It follows that $a^h(\tilde{m}) < \omega_1$. This is because $a^*(m_0')$ is a weighted average of ω_0, ω_1 and $a^h(m_0')$. If $\omega_1 \leq a^h(\tilde{m})$, then

$a^*(m'_0) < a^h(\tilde{m}) < \omega_1 + b$ which contradicts $|a^*(m'_0) - \omega_1 - b| < \left|a^h(\tilde{m}) - \omega_1 - b\right|$.

Now that we have $a^h(\tilde{m}) < \omega_1$, by using the properties of $a^h(m)$, we also have $m'_1 < \omega_1 + b$ and $a^h(m'_1) < \omega_1 + b$.

Since $u_d^S(\omega_1, m'_0, a^*(m'_0)) = u_d^S(\omega_1, m'_1, a^*(m'_1))$ where $|m'_0 - \omega_1 - b| > |m'_1 - \omega_1 - b|$, it must be the case that $|a^*(m'_0) - \omega_1 - b| < |a^*(\omega_1) - \omega_1 - b|$. Since $a^*(m'_0) < \omega_1$, it follows that either $a^*(m'_1) < a^*(m'_0)$ or $a^*(m'_1) > \omega_1 + b$.

The first case cannot happen. To see this, suppose $a^*(m'_1) < a^*(m'_0) < \omega_1$. Since $m'_0 < m'_1 < \omega_1 + b$ and $u_d^S(\omega_1, m'_0, a^*(m'_0)) = u_d^S(\omega_1, m'_1, a^*(m'_1))$, there exists $\omega \in (\omega_1, \omega_2)$ (where dishonest sender with ω_2 is indifferent between sending m'_1 and sending m'_2) such that $u_d^S(\omega, m'_0, a^*(m'_0)) > u_d^S(\omega, m'_1, a^*(m'_1))$, a contradiction.

Now consider the second case where $a^*(m'_1) > \omega_1 + b$. Since $a^*(m'_1)$ is a weighted average of $a^h(m'_1), \omega_1, \omega_2$, $a^*(m'_1) > \omega_1 + b$ and $a^h(m'_1) < \omega_1 + b$ imply that $a^h(m'_1) < a^*(m'_1) < \omega_2$. Also, $m'_1 < \omega_1 + b < a^*(m'_1) < \omega_2$. It follows that $m'_2 < m'_1 + b < \omega_2 + b$ and $a^h(m'_2) < \omega_2 + b$. Since $u_d^S(\omega_2, m'_1, a^*(m'_1)) = u_d^S(\omega_1, m'_2, a^*(m'_2))$ where $|m'_0 - \omega_1 - b| > |m'_1 - \omega_1 - b|$, by the same argument as above, we have $a^*(m'_2) > \omega_2 + b$, $\omega_3 > a^*(m'_2) > m'_2$, $\omega_3 + b > m'_3$, $\omega_3 + b > a^h(m_3)$.

Continuing this argument, if the dishonest sender with observation ω_i is indifferent between sending m'_{i-1} and $m'_i (i = 1, 2, 3, ..., I$, where $m'_I = 1)$, we have $m'_{i-1} < \omega_i$, $m'_i < \omega_i + b, a^*\left(m'_{i-1}\right) < \omega_i$ and $a^*(m'_i) > \omega_i + b$. Hence, we have $\omega_I + b > m'_I = 1$ and $\omega_I + b < a^*(m'_I) = a^*(1)$. These imply that $a^*(1) > 1$, which is impossible.

The constradiction implies that for every $m \in [m^*(0), 1] \cap M_n$, there exists $\omega \in \Omega$

s.t. $m^*(\omega) = m$.

To summarize, we have shown that $m^*(\omega)$ is an unrestricted best response to $a^*(m)$. $a^*(m)$ is a best response to $m^*(\omega)$ by definition. Since $m^*(\omega)$ is weakly increasing in ω, we have established existence of monotonic equilibrium in game Γ^n with discrete message space where the adjacent messages are close to each other. □

To prove Proposition 1.1, we need the following lemmas and claims.

Lemma 1.4. *$m(\omega)$ is continuous.*

Proof. Suppose not. Then there exists an $\omega \in [0,1]$ s.t. $\lim_{\omega \to \omega'} m(\omega) \neq m(\omega')$. This implies that $\exists m' \geq m(0)$ s.t. $pr(h = honest|m = m') = 1$, i.e., only the honest sender sends m'. It follows that $a(m') = m'$. If $m' \geq b$, then by sending m', the dishonest sender with $\omega = m' - b$ gets his highest possible payoff 0, which must be higher than the payoff he gets by sending $m(\omega)$. This contradicts the fact that $m(\omega)$ and $a(m)$ are equilibrium strategies. If $m' < b$, then the dishonest sender with $\omega = 0$ is better off by sending m' instead of $m(0)$, again a contradiction. Hence, the equilibrium strategy $m(\omega)$ has to be continuous. □

Lemma 1.5. *Suppose $m(\omega)$ is strictly increasing, $g(m)$ is continuously differentiable and $a(m)$ is continuous on an open interval I contained in $[0,1]$. Then, for any ω' s.t. $m(\omega') \in I$, $a(m)$ is differentiable at $m(\omega')$ and $\frac{da(m(\omega'))}{dm} = -\frac{\lambda(m(\omega')-\omega'-b)}{(1-\lambda)(a(m(\omega'))-\omega'-b)}$.*

Proof. Fix $b > 0$ and let $v(\omega, m, a) = u_d^S(\omega, m, a, b) = -\lambda(m-\omega-b)^2-(1-\lambda)(a-\omega-b)^2$.

Define, for a fixed ω',

$$h(x_1, x_2, x_3) = v(x_1, x_2, x_3) - v(x_1, m(\omega'), a(m(\omega')))$$

Define, for $\omega'' \neq \omega'$, the following:

$$\begin{aligned} x_1(\alpha) &= \omega', x_2(\alpha) = \alpha m(\omega'') + (1-\alpha) m(\omega'), \\ x_3(\alpha) &= \alpha a(m(\omega'')) + (1-\alpha) a(m(\omega')) \\ \tilde{x}_1(\alpha) &= \alpha\omega'' + (1-\alpha)\omega', \tilde{x}_2(\alpha) = m(\omega''), \tilde{x}_3(\alpha) = a(m(\omega'')) \end{aligned}$$

If $m(\omega)$ is incentive compatiable, then

$$\begin{aligned} h(\omega'', m(\omega''), a(m(\omega''))) &\geq 0 \\ h(\omega', m(\omega''), a(m(\omega''))) &\leq 0 \end{aligned}$$

Expanding $h(\omega'', m(\omega''), a(m(\omega'')))$ and then $h_1(\omega', m(\omega''), a(m(\omega'')))$ in a Taylor series yields, for some $\alpha, \beta \in [0, 1]$,

$$\begin{aligned} & h(\omega'', m(\omega''), a(m(\omega'')) \\ = \; & h(\omega', m(\omega''), a(m(\omega''))) + h_1(\omega', m(\omega''), a(m(\omega''))) (\omega'' - \omega') \\ & + \frac{1}{2} h_{11}(\tilde{x}(\beta)) (\omega'' - \omega')^2 \end{aligned}$$

$$
\begin{aligned}
&= h(\omega', m(\omega''), a(m(\omega''))) + h_1(\omega', m(\omega'), a(m(\omega')))(\omega'' - \omega') \\
&+h_{12}(x(\alpha))(m(\omega'') - m(\omega'))(\omega'' - \omega') \\
&+h_{13}(x(\alpha))(a(m(\omega'') - a(m(\omega'))))(\omega'' - \omega') \\
&+\frac{1}{2}h_{11}(\tilde{x}(\beta))(\omega'' - \omega')^2 \geq 0
\end{aligned}
$$

Because $h_1(\omega', m(\omega'), a(m(\omega'))) = 0$,

$$
\begin{aligned}
0 &\geq h(\omega', m(\omega''), a(m(\omega''))) \\
&\geq -h_{12}(x(\alpha))(m(\omega'') - m(\omega'))(\omega'' - \omega') \\
&\quad -h_{13}(x(\alpha))(a(m(\omega'') - a(m(\omega'))))(\omega'' - \omega') \\
&\quad -\frac{1}{2}h_{11}(\tilde{x}(\beta))(\omega'' - \omega')^2
\end{aligned}
$$

Expanding $h(\omega', m(\omega''), a(m(\omega'')))$ around $(\omega', m(\omega'), a(m(\omega')))$ in the above inequality, we have, for some $\gamma \in [0, 1]$,

$$
\begin{aligned}
0 &\geq h(\omega', m(\omega''), a(m(\omega''))) \\
&= h_2(\omega', m(\omega'), a(m(\omega')))(m(\omega'') - m(\omega')) \\
&+h_3(\omega', m(\omega'), a(m(\omega')))(a(m(\omega'')) - a(m(\omega'))) \\
&+\frac{1}{2}h_{33}(x(\gamma))(a(m(\omega'')) - a(m(\omega')))^2 + \frac{1}{2}h_{22}(x(\gamma))(m(\omega'') - m(\omega'))^2 \\
&+h_{23}(x(\gamma))(a(m(\omega'')) - a(m(\omega')))(m(\omega'') - m(\omega'))
\end{aligned}
$$

$$\geq -h_{12}(x(\alpha))(m(\omega'') - m(\omega'))(\omega'' - \omega')$$

$$-h_{13}(x(\alpha))(a(m(\omega'') - a(m(\omega'))))(\omega'' - \omega')$$

$$-\frac{1}{2}h_{11}(\tilde{x}(\beta))(\omega'' - \omega')^2$$

Dividing by $(\omega'' - \omega')$, we have the following inequalities (here we assume $\omega'' > \omega'$, otherwise the inequalities are reversed but the argument does not change):

$$0 \geq \left(h_2(\omega', m(\omega'), a(m(\omega'))) + \frac{1}{2}h_{22}(x(\gamma))(m(\omega'') - m(\omega'))\right)\frac{(m(\omega'') - m(\omega'))}{\omega'' - \omega'}$$

$$+\left(h_3(\omega', m(\omega'), a(m(\omega'))) + \frac{1}{2}h_{33}(x(\gamma))(a(m(\omega'')) - a(m(\omega')))\right.$$

$$\left.+h_{23}(x(\gamma))(m(\omega'') - m(\omega'))\right)\frac{a(m(\omega'')) - a(m(\omega'))}{\omega'' - \omega'}$$

$$\geq -h_{12}(x(\alpha))(m(\omega'') - m(\omega')) - h_{13}(x(\alpha))(a(m(\omega'') - a(m(\omega'))))$$

$$-\frac{1}{2}h_{11}(\tilde{x}(\beta))(\omega'' - \omega')$$

Since we assume that $m(\cdot)$ and $a(\cdot)$ are continuous, taking the limit as $\omega''-\omega' \to 0$, we have $m(\omega'') - m(\omega') \to 0$ and $a(m(\omega'')) - a(m(\omega')) \to 0$. Therefore,

$$0 \geq h_2(\omega', m(\omega'), a(m(\omega'))) \lim_{\omega'' \to \omega'} \frac{(m(\omega'') - m(\omega'))}{\omega'' - \omega'}$$

$$+h_3(\omega', m(\omega'), a(m(\omega'))) \lim_{\omega'' \to \omega'} \frac{a(m(\omega'')) - a(m(\omega'))}{\omega'' - \omega'} \geq 0$$

As we assume that $m(\cdot)$ is strictly increasing and differentiable, $\lim_{\omega'' \to \omega'} \frac{(m(\omega'')-m(\omega'))}{\omega''-\omega'}$ exists and is strictly greater than 0.

Therefore , we have

$$h_2\left(\omega', m\left(\omega'\right), a\left(m\left(\omega'\right)\right)\right)$$
$$+h_3\left(\omega', m\left(\omega'\right), a\left(m\left(\omega'\right)\right)\right) \lim_{m(\omega'')\to m(\omega')} \frac{a\left(m\left(\omega''\right)\right)-a\left(m\left(\omega'\right)\right)}{m\left(\omega''\right)-m\left(\omega'\right)}=0$$

Since $h_2\left(\omega', m\left(\omega'\right), a\left(m\left(\omega'\right)\right)\right)=-2\lambda\left(m\left(\omega'\right)-\omega'-b\right)$ and

$h_3\left(\omega', m\left(\omega'\right), a\left(m\left(\omega'\right)\right)\right)=-2\lambda\left(a(m\left(\omega'\right))-\omega'-b\right)$, for an ω' such that $a\left(m\left(\omega'\right)\right)-\omega'-b\neq 0$, $a\left(m\right)$ is differentiable at $m=m\left(\omega'\right)$ and

$$\lambda\left(m\left(\omega'\right)-\omega'-b\right)+\left(1-\lambda\right)\left(a\left(m\left(\omega'\right)\right)-\omega'-b\right)\frac{da\left(m\left(\omega'\right)\right)}{dm}=0 \qquad (1.1)$$

Now all we need to show is that for any $m\in I$, $a\left(m\right)-g\left(m\right)-b\neq 0$.

First we show that there doesn't exist E, an open subinterval of I, such that $a\left(m\right)-g\left(m\right)-b=0$ on E. Suppose the contrary. Since we assume that $g\left(m\right)$ is differentiable, $a'\left(m\right)=g'\left(m\right)$ on E. Then, equation (1.1) implies that $m-g(m)-b=0$ as well. But when $m=g\left(m\right)+b$, $a\left(m\right)=\frac{\theta m+(1-\theta)g'(m)g(m)}{\theta+(1-\theta)g'(m)}\neq m$ for $\theta\neq 1$, a contradiction.

Now suppose there exists $m'\in I$ such that $a\left(m'\right)-g\left(m'\right)-b=0$.

First consider the case where $m'-g\left(m'\right)-b\neq 0$. Then, due to the continuity of $a\left(m\right)$ and $g\left(m\right)$ and the assumption $a\left(m'\right)-g\left(m'\right)-b=0$, there exists $\varepsilon>0$, s.t. $0<\left|m''-m'\right|<\varepsilon$ and $a\left(m''\right)-g\left(m''\right)-b\neq 0 \Rightarrow \left|\frac{da(m'')}{dm}\right|>\left|\frac{dg(m'')}{dm}\right|$ and $\left(m''-g\left(m''\right)-b\right)$ has the same sign as $\left(m'-g\left(m'\right)-b\right)$.

Define $f(m) = a(m) - g(m) - b$. Then $f(m') = 0$ and $f'(m'')$ has the same sign as $a'(m'')$ for $a(m'') - g(m'') - b \neq 0$ and $0 < |m'' - m'| < \varepsilon$.

For $0 < |m'' - m'| < \varepsilon$ and $a(m'') - g(m'') - b \neq 0$, equation (1.1) implies $f(m'') > 0 \Leftrightarrow f'(m'')(m'' - g(m'') - b) < 0$.

Suppose for $0 < |m'' - m'| < \varepsilon$, $m'' - g(m'') - b < 0$. Since there does not exist an open subinterval of I on which $a(m) - g(m) - b = 0$, there exists m'' s.t. $m' - \varepsilon < m'' < m'$ and $a(m'') - g(m'') - b \neq 0$. If $f(m'') > 0$, then because of the uniform continuity of f on I, $f'(m) \geq 0$ for all $m'' \leq m < m'$ and therefore $f(m') \geq f(m'') > 0$, a contradiction. If $f(m'') < 0$, then $f'(m'') \leq 0$ for all $m'' \leq m < m'$ and therefore $f(m') \leq f(m'') < 0$, again a contradiction.

Suppose for $0 < |m'' - m'| < \varepsilon$, $m'' - g(m'') - b > 0$. Then choose m'' s.t. $m' < m'' < m' + \varepsilon$ and $a(m'') - g(m'') - b \neq 0$. A similar argument as above can be used with obvious changes.

Finally, let's consider the case where $m' = g(m') + b$. Since we also assume that $a(m') = g(m') + b$, this implies $g'(m') = 0$. Because of the continuity of g, there exists $\varepsilon > 0$, such that for $m' < m'' < m' + \varepsilon$, $a(m'') - g(m'') - b \neq 0$, we have $m'' > g(m'') + b$ and $f'(m'')$ has the same sign as $a'(m'')$. We can see that the argument in the previous paragraph follows through. □

Lemma 1.6. *Suppose $m(\omega)$ is strictly increasing, $g(m)$ is differentiable, and $a(m)$ is continuous. Then we have either (i) $m = a(m)$ or (ii) $m > a(m) > g(m)$ and $a(m) < g(m) + b$.*

Proof. Lemma 1.5 shows that under certain assumptions, $a'(m) = -\frac{\lambda(m-g(m)-b)}{(1-\lambda)(a(m)-g(m)-b)}$. So $a''(m)$ is well defined as well.

Note that $-\lambda(m-g(m)-b)-(1-\lambda)(a(m)-g(m)-b)\,a'(m)=0$ is the F.O.C. for the (local) optimality of $m(\omega)$, given the sophisticated receiver's strategy $g(m)$. The S.O.C. for $m(\omega)$ to be optimal is

$$-\lambda-(1-\lambda)\left((a'(m))^2+(a(m)-g(m)-b)\,a''(m)\right)\leq 0.$$

After substituting, we can simplify the S.O.C. as $\frac{g'(m)(m-a(m))}{(a(m)-g(m)-b)}\leq 0$. Since $g'(m)\geq 0$, the S.O.C. can be satisfied under the following three cases: (i) $m=a(m)$, (ii) $m>a(m)$ and $a(m)<g(m)+b$ or (iii) $m<a(m)$ and $a>g(m)+b$. Since $a(m)$ is a weighted average of $g(m)$ and m, the two inequalities in (iii) cannot hold simultaneously. Also, because $g'(m)$ is bounded when $m(\omega)$ is strictly increasing, $m>a(m)$ implies that $a(m)>g(m)$. □

Lemmas 1.5 and 1.6 provide the equilibrium conditions under a number of monotonicity, continuity and differentiability constraints. The following lemmas show that for $m\in[m(0),1)$, $m(\omega)$ is strictly increasing, $g(m)$ is differentiable and $a(m)$ is continuous so that lemmas 1.5 and 1.6 apply.

Lemma 1.7. *When $m(\omega)$ is strictly increasing and $a(m)$ is continuous, $g(m)$ is differentiable.*

Proof. We can prove the lemma by contradiction.

If $g(m)$ is differentiable, we have

$$a(m) = E(\omega|m) = \frac{\theta m + (1-\theta)g'(m)g(m)}{\theta + (1-\theta)g'(m)}$$

We know that for an increasing function, there can be at most a countable number of points at which the function is not differentiable. Suppose $g(m)$ is not differentiable at $\hat{m}$. Then there exist δ_1, $\delta_2 > 0$ s.t. $g(m)$ is differentiable on $(\hat{m} - \delta_1, \hat{m})$ and $(\hat{m}, \hat{m} + \delta_2)$. But $\lim_{m \to \hat{m}^-} g'(m) \neq \lim_{m \to \hat{m}^+} g'(m)$. Then we have

$$\begin{aligned} \lim_{m \to \hat{m}^-} a(m) &= \lim_{m \to \hat{m}^-} \frac{\theta m + (1-\theta)g'(m)g(m)}{\theta + (1-\theta)g'(m)} \\ &\neq \lim_{m \to \hat{m}^+} \frac{\theta m + (1-\theta)g'(m)g(m)}{\theta + (1-\theta)g'(m)} = \lim_{m \to \hat{m}^+} a(m), \end{aligned}$$

contradicting the assumption that $a(m)$ is continuous. □

Lemma 1.8. *$u_d^S(\omega', m(\omega'), a(m(\omega'')), b) \to u_d^S(\omega', m(\omega'), a(m(\omega')), b)$ as $\omega'' \to \omega'$.*

Proof. For any $\varepsilon > 0$, there exists δ_1 such that $|\omega'' - \omega'| < \delta_1 \Rightarrow |v(\omega', m(\omega'), a) - v(\omega', m(\omega''), a)| < \varepsilon$.

If $m(\omega)$ is incentive compatible, then

$v(\omega', m(\omega'), a(m(\omega'))) \geq v(\omega', m(\omega''), a(m(\omega''))) > v(\omega', m(\omega'), a(m(\omega''))) - \varepsilon$.

There also exists δ_2 such that $|\omega'' - \omega'| < \delta_2 \Rightarrow |v(\omega', m(\omega'), a) - v(\omega'', m(\omega''), a)| < \frac{\varepsilon}{2}$ and $|v(\omega'', m(\omega'), a) - v(\omega', m(\omega'), a)| < \frac{\varepsilon}{2}$.

Therefore, $v(\omega', m(\omega'), a(m(\omega''))) \geq v(\omega'', m(\omega''), a(m(\omega''))) - \frac{\varepsilon}{2}$

$\geq v(\omega'', m(\omega'), a(m(\omega'))) - \frac{\varepsilon}{2} \geq v(\omega', m(\omega'), a(m(\omega'))) - \varepsilon.$

Combining yields

$|\omega'' - \omega'| < \min\{\delta_1, \delta_2\} \Rightarrow |v(\omega', m(\omega'), a(m(\omega'))) - v(\omega', m(\omega'), a(m(\omega'')))| <$ ε. □

Lemma 1.9. *If $g(m)$ is strictly increasing on an open interval $I \subset [0,1]$, then $a(m)$ is continuous on I.*

Proof. Suppose $a(m)$ is discontinuous at $m_1 \in I$. Then there exists an open interval $I_1 \subset I$ such that $m_1 \in I_1$ and the only discontinuity of $a(m)$ on I_1 is m_1. Suppose $\{m'_n\}$, $\{m''_n\} \subset I_1$ are two convergent sequences such that $m'_n \uparrow m_1$ and $m''_n \downarrow m_1$. Since $u^S_d(g(m_1), m_1, a(m), b) \to u^S_d(g(m_1), m_1, a(m_1), b)$ as $m \to m_1$, we have $(g(m_1) + b - a(m))^2 \to (g(m_1) + b - a(m_1))^2$ as $m \to m_1$. Since $a(m)$ is discontinuous at m_1, either $\lim_{n\to\infty} a(m'_n) > g(m_1) + b$ or $\lim_{n\to\infty} a(m''_n) > g(m_1) + b$, a contradiction of Lemma 1.6. Therefore, $a(m)$ is continuous on I. □

Lemma 1.10. *$m(\omega)$ is strictly increasing on $[0, \overline{\omega})$ where $\overline{\omega} = \inf\{\omega : m(\omega) = 1\}$.*

Proof. First, observe that $\overline{\omega} > 0$. Suppose not. Then, $m(\omega) = 1$ for $\omega \in [0,1]$ and $a(m) = m$ if $m < 1$ and $a(1) = \frac{1}{2}$. Clearly, there is a profitable deviation for the dishonest sender with high observations of ω, say, $\omega = 1$.

We can show by contradiction that $m(\omega)$ is strictly increasing on $[0, \overline{\omega})$. Suppose not. Then there exist ω_1, ω_2 where $0 \leq \omega_1 < \omega_2 < \overline{\omega}$ and $\varepsilon_1 > 0$ such that for

$\omega \in [\omega_1, \omega_2]$, $m(\omega) = m_1$ and for $\omega \in (\omega_2, \omega_2 + \varepsilon_1)$, $m(\omega)$ is strictly increasing and differentiable. Since $a(m_1) = \frac{\omega_1+\omega_2}{2}$, by continuity of $a(m)$, $\lim_{m \to m_1} a(m) = a(m_1) = \frac{\omega_1+\omega_2}{2} < \omega_2$. Since $a(m)$ is a weighted average between m and $g(m)$ for $\omega \in (\omega_2, \omega_2 + \varepsilon_1)$, this implies that $m_1 < \omega_2$ and there exists $\varepsilon_2 > 0$ s.t. for $\omega \in (\omega_2, \omega_2 + \varepsilon_2)$, $m < a(m)$, which contradicts Lemma 1.6.

Hence, there can't be pooling at any message below 1. Therefore, $m(\omega)$ is strictly increasing on $[0, \overline{\omega})$. □

With the above lemmas, we can make further claims on the conditions that equilibrium strategies $m(\omega)$ and $a(m)$ have to satisfy.

Claim 1.1. $a(m(0)) = m(0)$.

Since $a(m) = m$ for $m < m(0)$, this claim, together with Lemma 1.9 and Lemma 1.10, implies that $a(m)$ is continuous on $[0, 1)$.

Proof. $a(m(0)) = E(\omega | m = m(0))$.

If $m(0) = 0$, then $a(m(0)) = m(0) = 0$.

If $m(0) > 0$, then $a(m(0)) \leq m(0)$.

Suppose $m(0) > 0$ and $a(m(0)) < m(0)$. Since for $m < m(0)$, we have $a(m) = m$, if $a(m(0)) < m(0)$, we can find an $\varepsilon > 0$ such that $a(m(0) - \varepsilon) = m(0) - \varepsilon > a(m(0))$ and

$$
\begin{aligned}
& -\lambda (m(0) - \varepsilon - b)^2 - (1 - \lambda)(a(m(0) - \varepsilon) - b)^2 \\
> \; & -\lambda (m(0) - b)^2 - (1 - \lambda)(a(m(0)) - b)^2
\end{aligned}
$$

Then, instead of sending $m(0)$, the rational sender would deviate and send $m(0) - \varepsilon$ when $\omega = 0$, thus violating the equilibrium condition. Hence, when $m(0) > 0$, we must also have $a(m(0)) = m(0)$. □

Claim 1.2. $0 < m(0) < b$.

Proof. Suppose $m(0) = 0$. Then by continuity there exists $\varepsilon \in (0, b)$ s.t. $0 < a(\varepsilon) < b$ and the dishonest sender with $\omega = 0$ would be better off sending $m = \varepsilon$ instead of $m = 0$, a contradiction.

Suppose $m(0) > b$. Then $m(0) - b > 0$. Because $pr(h = honest|m < m(0)) = 1$, we have $a(m) = m$ for $m < m(0)$. Then, instead of sending $m = m(\omega)$, the dishonest sender with $\omega \in [0, m(0) - b)$ would be better off sending $m = \omega + b < m(0)$, inducing both the naive and sophisticated receivers to respond with $a = \omega + b$ and achieving his highest possible payoff (a contradiction).

Suppose instead $m(0) = b$. Then $a(m(0)) = m(0) = b$ and therefore $a(m(0)) = g(m(0)) + b$. We can use the argument in the proof of Lemma 1.5 to show that $a(m(0)) \neq g(m(0)) + b$. Therefore, $m(0) \neq b$. □

Claim 1.3. *For $m \in (m(0), 1)$, $m > a(m) > g(m)$ and $a(m) < g(m) + b$.*

Proof. Consider the condition found in Lemma 1.5: $a'(m) = -\frac{\lambda(m-g(m)-b)}{(1-\lambda)(a(m)-g(m)-b)}$. Suppose $a(m) < g(m) + b$. Then the following must be true. If $m - g(m) - b > 0$, then $a'(m) > 0$; if $m - g(m) - b < 0$, then $a'(m) < 0$; if $m = g(m) + b$, $a'(m) = 0$.

Since $a(m(0)) = m(0) < b$, the right derivative of $a(m)$ at $m = m(0)$ is negative. Therefore, in equilibrium, for $m(0) < m < 1$, either $m < g(m) + b$ and $a'(m) < 0$

or $m \geq g(m) + b$ and $a(m) < g(m) + b$. In both cases, $m > a(m) > g(m)$ and $a(m) < g(m) + b$. □

Proposition 1.1 follows directly from the above lemmas and claims.

Proof of Proposition 1.2. We need to show that when the conditions in the Proposition are satisfied, there is no profitable deviation for the dishonest sender.

We can show this by contradiction. Suppose $(m(\omega), a(m))$ satisfy all the conditions in the Proposition and there exists an $\omega_0 \in [0,1]$ such that $m(\omega_0)$ is not a global maximum for the dishonest sender with observation $\omega = \omega_0$. Instead, there exists $m_1 \neq m(\omega_0)$ s.t. $m_1 \in \arg\max\{-\lambda(m - \omega_0 - b)^2 - (1-\lambda)(a(m) - \omega_0 - b)^2\}$. Note that $m_1 \geq m(0)$.

Suppose $m(0) < m_1 < 1$, then there exists ω_1 s.t. $m(\omega_1) = m_1$ and $\lambda(m_1 - \omega_0 - b) + (1-\lambda)(a(m_1) - \omega_0 - b) = 0$ and $\lambda(m_1 - \omega_1 - b) + (1-\lambda)(a(m_1) - \omega_1 - b) = 0$. This implies that $\frac{m_1 - \omega_0 - b}{a(m_1) - \omega_0 - b} = \frac{m_1 - \omega_1 - b}{a(m_1) - \omega_1 - b}$. Since $m_1 > a(m_1)$ and $\frac{m_1 - \omega - b}{a(m_1) - \omega - b}$ is strictly increasing in ω when $m_1 > a(m_1)$, this is not possible.

Suppose $m_1 = m(0)$ and $\omega_0 \neq 0$. Then there exists $\varepsilon > 0$ s.t. $\omega_0 > \varepsilon$. Consider $0 < \hat{\omega} < \varepsilon$, $\hat{m} = m(\hat{\omega})$, $\hat{a} = a(m(\hat{\omega}))$. We have

$$
\begin{aligned}
\frac{du_d^S(\omega_0, \hat{m}, \hat{a}, b)}{dm} &= -2\lambda(\hat{m} - \omega_0 - b) - 2(1-\lambda)(\hat{a} - \omega_0 - b)a'(\hat{m}) \\
&= -2\lambda(\hat{m} - \omega_0 - b) - 2(1-\lambda)(\hat{a} - \omega_0 - b)\left(\frac{-\lambda}{1-\lambda}\right)\left(\frac{\hat{m} - \hat{\omega} - b}{\hat{a} - \hat{\omega} - b}\right) \\
&= -2\lambda(\hat{a} - \omega_0 - b)\left(\frac{\hat{m} - \omega_0 - b}{\hat{a} - \omega_0 - b} - \frac{\hat{m} - \hat{\omega} - b}{\hat{a} - \hat{\omega} - b}\right).
\end{aligned}
$$

Since $\frac{\hat{m}-\omega-b}{\hat{a}-\omega-b}$ is increasing in ω and $\hat{a}-\omega_0-b<0$, we have $\frac{du_d^S(\omega_0,\hat{m},\hat{a},b)}{dm}>0$ for all $0<\hat{\omega}<\varepsilon$. Therefore, it's not possible to have $m_1=m(0)$ as a global maximum for ω_0.

Suppose $m_1=1$ and $a(m)$ is continuous at $m=1$. Then

$$
\begin{gathered}
-2\lambda(1-\omega_0-b)-2(1-\lambda)(a(1)-\omega_0-b))a'(1)\geq 0 \\
\Rightarrow \lambda(1-\omega_0-b)+(1-\lambda)(a(1)-\omega_0-b)\left(-\frac{\lambda(1-g(1)-b)}{(1-\lambda)(a(1)-g(1)-b)}\right)\leq 0 \\
\Rightarrow (a(1)-\omega_0-b)\left(\frac{(1-\omega_0-b)}{(a(1)-\omega_0-b)}-\frac{(1-g(1)-b)}{(a(1)-g(1)-b)}\right)\leq 0
\end{gathered}
$$

Since $g(1)>\omega_0$, the above inequality implies that $a(1)-\omega_0-b\geq 0$, but since $a(m(\omega_0))-\omega_0-b<0$ and $a(m)$ is continuous, this means that $m=1$ cannot be the optimal choice of message when $\omega=\omega_0$.

Suppose $m_1=1$ and $a(m)$ is not continuous at $m=1$. Let $a^l(1)=\lim_{m\to 1}a(m)$. Since

$$
\begin{aligned}
& -\left(a^l(1)-\omega_0-b\right)^2-\left(-\left(a(1)-\omega_0-b\right)^2\right) \\
= & \left(a(1)-a^l(1)\right)\left(a(1)+a^l(1)-2\omega_0-2b\right) \\
= & \left(a(1)-a^l(1)\right)\left(2g(1)-2\omega_0\right)>0,
\end{aligned}
$$

again $m=1$ cannot be the optimal choice when $\omega=\omega_0$.

Therefore, there is no profitable deviation for the dishonest sender. □

Proof of Theorem 1.3. Step 1: Suppose $(\hat{a}(m),\hat{g}(m))$ and $(\tilde{a}(m),\tilde{g}(m))$ are solutions to 3. (a) and 3. (b) in Proposition 1.3, with initial conditions $(\hat{a}(\hat{m}_0) = \hat{m}_0, \hat{g}(\hat{m}_0) = 0)$ and $(\tilde{a}(\tilde{m}_0) = \tilde{m}_0, \tilde{g}(\tilde{m}_0) = 0)$, respectively. Suppose also that $\hat{m}_0 > \tilde{m}_0$, then we can show that $\hat{a}(m) > \tilde{a}(m)$ and $\hat{g}(m) > \tilde{g}(m)$ for $1 > m > \hat{m}_0$.

Note that at $\hat{m}_0, \hat{a}(\hat{m}_0) = \hat{m}_0 > \tilde{a}(\hat{m}_0)$, $\hat{g}(\hat{m}_0) = 0 < \tilde{g}(\hat{m}_0)$. First, we want to show that if $\hat{a}(m) > \tilde{a}(m)$ for $m > \hat{m}_0$, then $\hat{g}(m) < \tilde{g}(m)$ for $m > \hat{m}_0$. To see this, consider equation $a(m) = \frac{\theta m + (1-\theta) g'(m) g(m)}{\theta + (1-\theta) g'(m)}$. Because $g(m) \leq m$, it follows that if $\hat{a}(m) > \tilde{a}(m)$ and $\hat{g}(m) < \tilde{g}(m)$, then $\hat{g}'(m) < \tilde{g}'(m)$. Together with the condition $\hat{g}(\hat{m}_0) < \tilde{g}(\hat{m}_0)$, $\hat{a}(m) > \tilde{a}(m)$ for $m > \hat{m}_0$ implies $\hat{g}(m) < \tilde{g}(m)$ for $m > \hat{m}_0$.

We need to show that $\hat{a}(m) > \tilde{a}(m)$ for $m > \hat{m}_0$. Suppose not. Since $\hat{a}(\hat{m}_0) > \tilde{a}(\hat{m}_0)$ and $\hat{a}(\cdot), \tilde{a}(\cdot)$ are continuous, $\hat{a}(m)$ and $\tilde{a}(m)$ have to intersect. Suppose m^* is the lowest point at which $\hat{a}(m)$ and $\tilde{a}(m)$ intersect. Then, for $\hat{m}_0 \leq m < m^*$, $\hat{a}(m) > \tilde{a}(m)$ and by the argument made in the previous paragraph, $\hat{g}(m) < \tilde{g}(m)$ for $\hat{m}_0 \leq m \leq m^*$

Consider equation $a'(m) = -\frac{\lambda(m - g(m) - b)}{(1-\lambda)(a(m) - g(m) - b)}$. We can show that at m^*, $\hat{g}(m^*) < \tilde{g}(m^*)$ and $\hat{a}(m^*) = \tilde{a}(m^*)$ imply that $\hat{a}'(m^*) > \tilde{a}'(m^*)$. By continuity, $\exists \varepsilon > 0$ s.t. for m where $|m - m^*| < \varepsilon$, $\hat{a}'(m) > \tilde{a}'(m)$. But since $\hat{a}(m) > \tilde{a}(m)$ for $\hat{m}_0 \leq m < m^*$, it is not possible that $\hat{a}(m^*) = \tilde{a}(m^*)$, a contradiction. Hence, $\hat{a}(m)$ and $\tilde{a}(m)$ never intersect. Since $\hat{a}(\hat{m}_0) > \tilde{a}(\hat{m}_0)$, we can conclude that $\hat{a}(m) > \tilde{a}(m)$ and $\hat{g}(m) < \tilde{g}(m)$ for $1 > m > \hat{m}_0$. From our proof we can also see that $\lim_{m \to 1} \hat{a}(m) > \lim_{m \to 1} \tilde{a}(m)$ and $\lim_{m \to 1} \hat{g}(m) < \lim_{m \to 1} \tilde{g}(m)$.

Step 2: Suppose sequence $\{\hat{\omega}_0, \hat{\omega}_1, ..., \hat{\omega}_{\hat{I}} = 1\}$ is a solution to (A) and satisfies $-(\hat{a}(1) - \hat{\omega}_0 - b)^2 = -\left(\frac{\hat{\omega}_0+\hat{\omega}_1}{2} - \hat{\omega}_0 - b\right)^2$ where $\hat{\omega}_0 < \hat{a}(1) < \hat{\omega}_0 + b$, sequence $\{\tilde{\omega}_0, \tilde{\omega}_1, ..., \tilde{\omega}_{\tilde{I}} = 1\}$ is a solution to (A) and satisfies $-(\tilde{a}(1) - \tilde{\omega}_0 - b)^2 = -\left(\frac{\tilde{\omega}_0+\tilde{\omega}_1}{2} - \tilde{\omega}_0 - b\right)^2$ where $\tilde{\omega}_0 < \tilde{a}(1) < \tilde{\omega}_0 + b$.

If $\hat{I} = \tilde{I}$, then it straightforward to show that if $\hat{\omega}_0 < \tilde{\omega}_0$, then $\hat{a}(1) < \tilde{a}(1)$.

If $\hat{I} \neq \tilde{I}$, without loss of generality we can assume $\hat{I} > \tilde{I}$, we have $-(\tilde{a}(1) - \tilde{\omega}_0 - b)^2 = -\left(\frac{\tilde{\omega}_0+\tilde{\omega}_1}{2} - \tilde{\omega}_0 - b\right)^2$ and $-\left(\frac{\hat{\omega}_{\hat{I}-\tilde{I}-1}+\hat{\omega}_{\hat{I}-\hat{I}}}{2} - \hat{\omega}_{\hat{I}-\hat{I}} - b\right)^2 = -\left(\frac{\hat{\omega}_{\hat{I}-\tilde{I}}+\hat{\omega}_{\hat{I}-\tilde{I}+1}}{2} - \hat{\omega}_{\hat{I}-\tilde{I}} - b\right)^2$. Since $\tilde{\omega}_0 < \tilde{a}(1) < \tilde{\omega}_0 + b$ and $\frac{\hat{\omega}_{\hat{I}-\tilde{I}-1}+\hat{\omega}_{\hat{I}-\hat{I}}}{2} < \hat{\omega}_{\hat{I}-\hat{I}}$,it must be the case that $\hat{\omega}_{\tilde{I}-\hat{I}} < \tilde{\omega}_0$(therefore $\hat{\omega}_0 < \tilde{\omega}_0$) and $\hat{a}(1) < \hat{\omega}_{\tilde{I}-\hat{I}} < \tilde{\omega}_0 < \tilde{a}(1)$.

Combining the two steps, we see that for a set of fixed parameters $\lambda, \theta, b > 0$ there can be no more than one solution to the differential equations system $3.(a) - 3.(e)$. Since the solution to $3.(a) - 3.(e)$, $(a(m), g(m), \{\omega_0, \omega_1, ..., \omega_I\})$, determines the equilibrium outcome defined on $T \times M \times A$, the monotonic equilibrium outcome in the cheap-talk extension game $\Gamma(A)$ is unique. □

Proof of Proposition 1.4. Consider an equilibrium in Γ_{C-S}. Suppose in this equilibrium, there are N subintervals in the partition of Ω. Then, the equilibrium payoff for the sender with observation $\omega = 0$ is equal to $-\left(\frac{\frac{1}{N}+2b(1-N)}{2} - 0 - b\right)^2$.[20]

Since $N(b)$ is the smallest integer that is greater than or equal to $-\frac{1}{2}+\frac{1}{2}\left(1+\frac{2}{b}\right)^{\frac{1}{2}}$, if $N \neq N(b)$, then $N < -\frac{1}{2} + \frac{1}{2}\left(1+\frac{2}{b}\right)^{\frac{1}{2}}$ and $-\left(\frac{\frac{1}{N}+2b(1-N)}{2} - 0 - b\right)^2 < -b^2$; if $N = N(b)$, $N \geq -\frac{1}{2} + \frac{1}{2}\left(1+\frac{2}{b}\right)^{\frac{1}{2}}$ and $-\left(\frac{\frac{1}{N}+2b(1-N)}{2} - 0 - b\right)^2 \geq -b^2$.

[20] For a step-by-step derivation, see Crawford and Sobel (1982), p. 1441.

Recall that in Claim 1.1 and Claim 1.2, we show that in any $\Gamma(A)$ with $\lambda > 0$ and $\theta > 0$, $0 < a(m(0)) = m(0) < b$ in equilibrium. Therefore, the equilibrium payoff for the rational sender with observation $\omega = 0$ is $u_d^S(0, m(0), a(m(0))) = -\lambda(m(0) - 0 - b)^2 - (1-\lambda)(a(m(0)) - 0 - b)^2 > -b^2$. Now consider the converging sequence of games $\{\Gamma_{\lambda_i,\theta_j}\}$. Since the inequality holds for all $\lambda_i > 0, \theta_j > 0$, in the limit equilibrium as λ_i and θ_j approach 0, the payoff for the dishonest sender with observation $\omega = 0$ must be greater than or equal to $-b^2$. From the previous paragraph, we know that only in the most informative equilibrium in Γ_{C-S} does this inequality hold. Hence the result. □

Proof of Proposition 1.5. Let $\hat{E}^n$ be a limit monotonic equilibrium of $\{\Gamma^n_{\lambda_i,\theta_j}\}$ as λ_i and θ_j approach 0. Suppose $\{[\omega_{k-1}, \omega_k]\}_{k=1,\dots,K}$ $(1 \leq K \leq N(b))$ is the partition of Ω in $\hat{E}^n$.

First, we show that $m(\omega) = 1$ for $\omega \in (\omega_{K-1}, \omega_K)$ in $\hat{E}^n$. Suppose not. Then, $m(\omega) < 1$ for $\omega \in (\omega_{K-1}, \omega_K)$ and $a(m(1)) = \frac{\omega_{K-1}+\omega_K}{2}$ in E^n. Also, $a(1)$ is a weighted average of $\omega_K = 1$ and $a^h(1)$ where $a^h(1) > 1 - \frac{b}{4} > \frac{\omega_{K-1}+\omega_K}{2}$ given our assumption that $\max\{m_i - m_{i-1}\} < b$. Therefore, $a(1) > \frac{\omega_{K-1}+\omega_K}{2}$ and the sender with observation $\omega = 1$ would be better off sending $m = 1$ instead of sending $m(1) < 1$, a contradiction.

Next, we show that if $K \geq 2$, then $m(\omega) = m_{n-1}$ for $\omega \in (\omega_{K-2}, \omega_{K-1})$ in E^n. Again, we can show this by contradiction. Suppose not, then $m(\omega) < m_{n-1}$ for $\omega \in (\omega_{K-2}, \omega_{K-1})$. Since we have established in the previous paragraph that

$m(\omega) = 1$ for $\omega \in (\omega_{K-1}, \omega_K)$ in E^n, this implies that $a(m_{n-1})$ is a weighted average of ω_{K-1} and $a^h(m_{n-1})$[21] $(> \omega_{K-1})$, which in turn implies that $a(m_{n-1}) > \omega_{K-1}$. It follows that $a(m_{n-1}) = a(m_n) = \frac{\omega_{K-1}+\omega_K}{2} > \omega_{K-1} + b$.(If $a(m_{n-1}) \neq a(m_n)$, then there always exists $\omega \in (\omega_{K-1}, \omega_K)$ s.t. it is a profitable deviation for the sender with observation ω to send $a(m_{n-1})$ instead of $a(m_n)$.) Now, given that $\hat{E}^n$ is a limit monotonic equilibrium of $\{\Gamma^n_{\lambda_i,\theta_j}\}$ as λ_i and θ_j approach 0, for any $\varepsilon_1 > 0$, there exists $\delta > 0$ s.t. if $0 < \lambda_i,\ \theta_j < \delta$, there exists ω satisfying $|\omega - \omega_{K-1}| < \varepsilon_1$ s.t. in a monotonic equilibrium of $\Gamma^n_{\lambda_i,\theta_j}$, the sender with observation ω is indifferent between sending m_n and sending m_{n-1}. We can always find ε_1 s.t. $m_n > m_{n-1} > \omega + b$ and $a(m_{n-1}) = a(m_n) > \omega + b$ in a monotonic equilibrium of $\Gamma^n_{\lambda_i,\theta_j}$. But then there exists $\varepsilon_2 > 0$ s.t. for $0 < \omega' - \omega < \varepsilon$, the dishonest sender with observation ω' prefers sending m_{n-1} to sending m_n, a contradiction. Therefore, in the limit equilibrium $\hat{E}^n, m(\omega) = m_{n-1}$ for $\omega \in (\omega_{K-2}, \omega_{K-1})$.

Continuing this argument, we can show that in $\hat{E}^n$, $m(\omega) = m_{n-k}$ for $\omega \in (\omega_{K-k-1}, \omega_{K-k})$.

Finally, we need to show that $K = N(b)$. From the previous paragraphs, we know that in $\hat{E}^n$, $m(\omega) > m_0 = 0$ for $\omega \in (0,1]$. Therefore, $a(0)$ is a weighted average of 0 and $a^h(0)$ and $0 \leq a(0) \leq \frac{b}{4}$. It follows that in $\hat{E}^n$, the equilibrium payoff of the sender with observation $\omega = 0$ must be greater than or equal to $-|a(m_0) - 0 - b|^2\ (\geq -b^2)$. In light of the proof for proposition 4, we know that the limit equilibrium $\hat{E}^n$ must

[21]The definition of $a^h(m_i)$ is the same as in the proof of Theorem 1.1.

correspond to a most informative equilibrium in the C-S model, i.e., $K = N(b)$. Hence the result. □

Chapter 2

Partially Informed Decision Makers in Communication Games

2.1 Introduction

Standard sender-receiver games[1] usually assume that the sender has all the relevant information for making a good decision and the receiver, while having the decision making power, is completely ignorant of the state of the world and has to rely on the sender for useful information. While this assumption simplifies analysis, it also fails to capture something quite important in real life communication – that the decision maker is usually partially informed as well. For example, the decision maker may have her own expertise. Imagine a CEO trying to decide which project is most profitable to take. She may need the division managers' input on the specifics of certain projects,

[1] For example, Crawford and Sobel's (1982) classic model of strategic information transmission.

but she has her own knowledge about investments as well. Alternatively, there may exist other sources of information. Consider an individual trying to decide what brand/model of a product she should purchase by consulting a salesperson in a store. she may already know something about the candidate products because she has read ratings on their characteristics and qualities in a magazine or her friends have told her about them.

The interesting situations to study involve players whose interests are not perfectly aligned and the expert may have an incentive to lie. How does the decision maker's partial information affects the expert's communication incentives? Is the expert more likely to lie to someone who knows practically nothing or someone who is well informed? What could the decision maker do to elicit more information from the expert? And how is the welfare of the players affected? To address these questions, we analyze three extensive form games (which correspond to three different communication environments) with partially informed decision makers.

In all three games, the players – the expert and the decision maker (DM) – both privately observe (conditionally independent) signals about the state of the world and communicate (through costless messages[2]) before the DM chooses an action that affects both players' payoffs. The games differ in the way the communication is structured. In game 1, the DM reveals her information truthfully to the expert before the expert reports to her. In game 2, the DM keeps her information private while

[2] The messages are "cheap talk" except in game 1 where the DM commits to telling the truth.

the expert communicates to her. In game 3, the DM cannot commit to revealing her information truthfully to the expert, but before the expert reports to her, she has an opportunity to communicate strategically to the expert.

Several unexpected results arise. The decision maker's expected equilibrium payoff is not monotonically increasing in the accuracy of her own information. This is because an enhancement in the quality of the decision maker's information may discourage the expert from revealing his information. In the games that we consider, whether the expert truthfully reveals his information in equilibrium depends on the expected impact of his messages on the DM's choice of action. When the impact is sufficiently large/small, the expert can/cannot credibly reveal his information to the DM in equilibrium. For example, suppose the expert has an upward bias in the binary model we study. So the expert's ideal point (i.e., the DM's action that maximizes his payoff) is always higher than the DM's ideal point, no matter what the state of the world is. The expert with the high signal always wants to convince the DM that his signal is indeed high and the question is whether the expert with the low signal can credibly convey his information to the DM. If the DM's responses to different messages are far apart, then the low type expert would not want to send a high message because the action induced would be too high for himself, in spite of his upward bias. However, if the DM's responses to different messages are close, the expert with either the high or the low signal would prefer the higher action and therefore the low type expert cannot credibly communicate his observation to the DM. If the DM's own information is highly accurate, then the expert's messages do not change her beliefs very much, which implies that the impact of the expert's messages on the DM's choice of action is small. In this case, as we explained above, the expert cannot credibly convey his information in equilibrium. Therefore, the DM's expected payoff may fall as a result of an increase in the quality of her own information if the gain does not adequately compensate for the loss of the expert's information.

Another somewhat surprising result is that allowing one extra round of communication in which the DM strategically communicates to the expert does not help her extract more information from the expert than if she keeps her signal private. It is obvious that any equilibrium outcome in game 2 is also an equilibrium outcome in game 3 because there always exists an equilibrium in game 3 in which the DM babbles in the first stage and in effect keeps her information private. Therefore, in game 3 the DM can do at least as well as in game 2. Why can't she do strictly better in game 3? Observe that the DM wants to elicit as much information as possible from the expert and the expert realizes this. If sending a particular message to the expert induces him to reveal more information, then the DM would always want to send this message, irrespective of the realization of her signal. But this undermines the credibility of her messages. Therefore, without committing to revealing her signal truthfully, the DM cannot effectively communicate with the the expert and the extra round of strategic communication does not help her extract more information.

What would the DM do if she could choose between the different communication environments, i.e., what game to play? Our analysis shows that whether she prefers to reveal her information or to keep it private depends on the parameters of the model. Fixing the other parameters and varying the expert's bias, we find that the DM is indifferent between playing game 1 and game 2 when the expert's bias is extreme, but prefers one to the other when the bias is in the intermediate range. When the bias is very small, the expert reveals his information truthfully in both games and the DM is indifferent. As the bias increases from a small value to a moderate value, the DM may extract more information from the expert by keeping her signal private than by revealing her signal truthfully to the expert. However, as the bias gets even larger, the DM's preference switches – revealing her information to the expert generates a higher expected payoff. When the bias is sufficiently large, the expert babbles in both games and the DM is indifferent again. A similar result is obtained by varying the accuracy of the DM's information while keeping other parameters fixed.

Only a few papers in the literature have explicitly modeled informed receivers in communication games. Olszewski (2004) analyzes a model in which the receiver, as well as the sender, has private information on the state of the world and the sender wants to be perceived as honest. The paper provides conditions on the information structure with which the unique equilibrium is full information revelation if the sender's reputational concerns are strong enough. Harris and Raviv (2004) consider the problem of delegation versus communication when both a CEO and a division manager have private information on the profitability of different investment projects.

They find that if the division manager's information is sufficiently important relative to the CEO's, then it is optimal for the CEO to delegate the investment decision to the division manager instead of making the decision herself with a report from the manager. Seidmann (1990) gives examples which illustrate that the receiver's partial information may induce the sender to tell the truth in equilibrium even if the sender always prefers higher actions independent of his private information. The examples show that the uncertainty created by the receiver's private information helps separation in equilibrium.

Austen-Smith (1993) does not explicitly model informed receivers in communication games. However, since his paper considers multiple referrals under open rule and the receiver gets information from two different sources, some of his results are similar to what our model generates.

2.2 The Model

Suppose there are two players, the expert (player 1) and the decision maker (player 2).

The state of the world, ω, is a random variable which takes on two values, 0 and 1. The common prior on ω is $prob\{\omega = 1\} = p$ and $prob\{\omega = 0\} = 1 - p$ where $p \in (0, 1)$. The expert has private information about the state of the world. Specifically, he privately observes a signal s_1 where $prob\{s_1 = x|\omega = x\} = q_1$ and $prob\{s_1 = 1 - x|\omega = x\} = 1 - q_1$ for $x = 0, 1$. The accuracy of s_1 is parameterized by

q_1 and we assume that $q_1 \in (\frac{1}{2}, 1]$. When $q_1 = 1$, the expert is perfectly informed. Like most sender-receiver games in the literature, we consider games in which the expert strategically and costlessly communicates to the DM about his private information and the DM then takes an action a which affects both players's payoffs. We depart from the standard models by assuming that the DM may also have private information on ω. Specifically, we assume that she privately observes a signal s_2 with accuracy $q_2 \in [\frac{1}{2}, 1)$ (i.e., $prob\{s_2 = x|\omega = x\} = q_2$ and $prob\{s_2 = 1 - x|\omega = x\} = 1 - q_2$ for $x = 0, 1$). Obviously, when $q_2 = \frac{1}{2}$, the DM is not informed as in the standard models.[3]

The von Neumann-Morgenstern utility functions of the players are assumed to take the following "quadratic loss" functional form:

$$\begin{aligned} u^E(a, \omega, b) &= -(a - \omega - b)^2 \\ u^{DM}(a, \omega) &= -(a - \omega)^2 \end{aligned}$$

where $a \in A = \mathbb{R}$ is the action that the decision maker takes and b, short for bias, measures the divergence of interest between the two players. WLOG, we assume that $b \geq 0$.

Now that the DM is partially informed, the strategic incentives of the players are changed. How they are changed depends how the communication between the expert

[3]It may be natural to assume that $q_1 > q_2$, i.e., the expert is better informed than the DM. However, making this assumption does not generate extra insight in our model. For generality, we will not impose it.

and the DM is structured.

There are at least three simple communication environments (which correspond to three different extensive form games) we can consider.

1. The DM reveals her signal s_2 truthfully to the expert before the expert reports to her. Call this Γ_1.

2. The DM keeps her signal s_2 private while the expert reports to her. Call this Γ_2.

3. The DM cannot commit to revealing s_2 truthfully to the expert. But before the expert reports to the DM, the DM strategically sends a message to the expert about her signal s_2. Call this Γ_3.

In Γ_1, the DM has the commitment power to reveal her signal truthfully to the expert. This could happen, for example, if the DM's information comes from a neutral third party and the DM allows this third party to reveal the information to the expert.

So far we have referred to s_2 as the DM's private signal. However, there is an alternative and perhaps more natural interpretation of s_2 in Γ_1 and Γ_2. Think of s_2 as a public signal that both players observe and it arrives after the expert observes s_1. Then, game Γ_1 describes the strategic situation if the expert reports to the DM after the arrival of s_2. On the other hand, if the expert reports to the DM before the arrival of s_2 and the DM chooses her action after hearing the expert's report and learning about the realization of s_2, we are in Γ_2. In this interpretation , it is the *timing* of communication that makes the difference.

Throughout the analysis of these three games, we are going to use m to denote the message that the expert sends to the DM and l to denote the message that the DM sends to the expert (if she does that in the game). Since we have a binary state space, we assume, for simplicity, that the message spaces are also binary, i.e., $L = M = \{0, 1\}$.

Let σ_i denote the expert's mixed reporting strategy and s_i denote his pure reporting strategy in game Γ_i. Due to the strict concavity of the DM's payoff function, she never plays a mixed strategy in equilibrium and let's use a_i to denote her pure action strategy in game Γ_i.

Since the DM truthfully reveals s_2 to the expert in Γ_1, the expert's mixed strategy is $\sigma_1 : S_1 \times S_2 \times M \to [0, 1]$, where $\sigma_1 (s_1, s_2, m)$ stands for the probability that the expert with observation s_1 sends m if the DM reveals her signal to be s_2. The expert's pure strategy m_1 is a mapping from $S_1 \times S_2$ to M. In Γ_1, the DM's choice of action depends on her signal s_2 as well as the message sent by the expert. Therefore, the DM's strategy is $a_1 : M \times S_2 \to A$.

In Γ_2, the expert does not know what s_2 is when sending a message to the DM. Therefore, his mixed strategy is $\sigma_2 : S_1 \times M \to [0, 1]$ where $\sigma_2 (s_1, m)$ stands for the probability that the expert with observation s_1 sends message m. His pure strategy is $m_2 : S_2 \to M$. Also, the DM's action strategy is $a_2 : M \times S_2 \to A$.

In Γ_3, the DM sends a message to the expert in the first round of communication. Let $\rho : S_2 \times L \to [0, 1]$ denote her mixed strategy where $\rho (s_2, l)$ stands for the probability that the DM with observation s_2 sends a message l to the expert. In the second round of communicaiton, the expert's choice of message may depend on s_1 as well as the DM's cheap-talk message l. Therefore, his mixed strategy is $\sigma_3 : S_1 \times L \times M \to [0, 1]$ where $\sigma_3 (s_1, l, m)$ stands for the probability that the expert with observation s_1 sends message m if he receives a message l sent by the DM. His pure strategy is $m_3 : S_1 \times L \to M$. And the DM's action strategy is $a_3 : L \times M \times S_2 \to A$.

The solution concept we use is Perfect Bayesian Equilibrium (PBE). As is typical in cheap talk models, the problem of multiple equilibria arises in all three games. We will address this problem later. For now, we focus on finding the conditions under which separation (truth telling) on the expert's part can happen in equilibrium.

2.3 Baseline case : $q_2 = \frac{1}{2}$

Let's first consider the baseline case where $q_2 = \frac{1}{2}$, i.e., the DM is not informed. Since we assume that the DM has no useful information, letting her communicate (either truthfully or strategically) to the expert does not add anything. Therefore, we will analyze game Γ_2 in which there is only one round of communication from the expert to the DM. Much of the intuition we develop in analyzing this simple case can be generalized easily and will help us understand the results in more complicated settings. For this reason, we will discuss the derivation of the results in details here.

Define a truth-telling equilibrium as an equilibrium in which the expert adopts a truthful message strategy, i.e., $m(s_1)^4 = s_1$ where $s_1 \in \{0,1\}$.[5]

To find the conditions for truth telling to happen in equilibrium, we analyze the incentive compatibility (IC) constraints for the players. Suppose there exists a truth-telling equilibrium. What is the DM's action strategy, $a(m)$, in the truth-telling equilibrium? Given her payoff functions, the DM's best response is the conditional expectation of ω, i.e., $a(0) = E(\omega|s_1 = 0)$ and $a(1) = E(\omega|s_1 = 1)$. By sending message m, the expert with observation $s_1 = 0$ (call him type 0 expert) has an expected payoff equal to $E(-(a(m) - \omega - b)^2|s_1 = 0)$. Similarly, the expert with observation $s_1 = 1$ (call him type 1 expert) has an expected payoff equal to $E(-(a(m) - \omega - b)^2|s_1 = 1)$. It's useful to note that given the quadratic loss functional form, $E(-(a(m) - \omega - b)^2|s_1) = -(a(m) - E(\omega|s_1) - b)^2 - Var(\omega|s_1)$. Since the variance terms do not depend on m, they usually cancel out when we consider the IC constraints.

Since we assume that the expert has an upward bias, type 1 expert would like to convince the DM that his signal is indeed equal to 1. That is what the following lemma says.

Lemma 2.1. *Suppose $b \geq 0$ and $q_2 = \frac{1}{2}$. In a truth-telling equilibrium in Γ_2,* $E(U^E(a(1), \omega, b)|s_1 = 1) > E(U^E(a(0), \omega, b)|s_1 = 1)$.

---

[4] To make notation cleaner, we suppress the subscript that indicates that we are considering Γ_2. We will continue doing this in this chapter whenver it is clear what game is under consideration and the suppression does not cause confusion.

[5] In general, a truth-telling equilibrium in Γ_i is defined as a PBE in which $m_i(s_1, \cdot) = s_1$ for $s_1 \in \{0,1\}$.

Proof. All we need to show is that

$$\begin{aligned} & -\left(a\left(1\right) - E\left(\omega|s_1 = 1\right) - b\right)^2 - Var\left(\omega|s_1 = 1\right) \\ > & -\left(a\left(0\right) - E\left(\omega|s_1 = 1\right) - b\right)^2 - Var\left(\omega|s_1 = 1\right). \end{aligned}$$

Since $a(1) = E(\omega|s_1 = 1) > a(0) = E(\omega|s_1 = 0)$, the above inequality is satisfied when $b \geq 0$. □

Therefore, the IC constraint for type 1 expert is never binding and we just need to consider one IC constraint for the truth-telling equilibrium, that is, type 0 expert does not have an incentive to deviate and report to the DM that his signal is 1. By analyzing this IC constraint, we have the following proposition.

Proposition 2.1. *Suppose $q_2 = \frac{1}{2}$ and $b \geq 0$. A truth-telling equilibrium exists if and only if $b \leq \frac{1}{2}\left(\frac{p(1-p)(2q_1-1)}{pq_1+(1-p)(1-q_1)(p(1-q_1)+(1-p)q_1)}\right)$.*

Proof. If the IC constraint for the type 0 expert is satisfied, then there exists a truth-telling equilibrium.

The constraint requires that

$$E\left(-\left(a\left(0\right) - \omega - b\right)^2 |s_1 = 0\right) \geq E\left(-\left(a\left(1\right) - \omega - b\right)^2 |s_1 = 0\right).$$

This implies that

$$-\left(a\left(0\right) - E\left(\omega|s_1 = 0\right) - b\right)^2 - Var\left(\omega|s_1 = 0\right)$$

$$\begin{aligned} &\geq -\left(a\left(1\right) - E\left(\omega|s_1 = 0\right) - b\right)^2 - Var\left(\omega|s_1 = 0\right) \Rightarrow \\ -b^2 &\geq -\left(\frac{pq_1}{pq_1 + (1-p)(1-q_1)} - \frac{p(1-q_1)}{p(1-q_1) + (1-p)q_1} - b\right)^2 \end{aligned}$$

Since we assume $b \geq 0$, the above inequality is equivalent to

$$b \leq \frac{1}{2}\left(\frac{p(1-p)(2q_1-1)}{pq_1 + (1-p)(1-q_1)(p(1-q_1) + (1-p)q_1)}\right).$$

If $b \geq \frac{1}{2}\left(\frac{p(1-p)(2q_1-1)}{pq_1+(1-p)(1-q_1)(p(1-q_1)+(1-p)q_1)}\right)$, then the IC constraint for type 0 expert is violated and truth telling cannot happen in equilibrium. □

Quite intuitively, a truth-telling equilibrium exists when the expert's bias is sufficiently small. In fact, if a truth-telling equilibrium exists for a certain (positive) value of b, then it must exist for any smaller (positive) value of b.

From the proof of Proposition 2.1, we see that whether a truth-telling equilibrium exists depends on the difference between $a(1)$ and $a(0)$, relative to the size of b. When $a(1)$ and $a(0)$ are far apart, type 0 expert has no incentive to deviate from telling the truth: although $a(0)$ is below his expected ideal point, $a(1)$ is just too high to be profitable for him. On the other hand, when the distance between $a(1)$ and $a(0)$ is small relative to b, type 0 expert has an incentive to lie and induce the DM to choose $a(1)$ which is higher than $a(0)$, but not too high for the expert since he has an upward bias. In that case, the IC constraint for type 0 expert is violated and truth telling cannot happen in equilibrium.

How does the accuracy of the expert's information affect his incentives for telling the truth? Let $\overline{b} = \frac{1}{2}\left(\frac{p(1-p)(2q_1-1)}{pq_1+(1-p)(1-q_1)(p(1-q_1)+(1-p)q_1)}\right)$. From Proposition 2.1, we know that when b lies in the interval $[0, \overline{b}]$, a truth-telling equilibrium exists. We have the following result.

Remark 2.1. *The range of b for which a truth-telling equilibrium exists is increasing in the accuracy of the expert's signal, q_1.*

See appendix for proof.

The more accurate the expert's information is, the larger the range of b for which truth telling can happen in equilibrium. The intuition behind this result should be quite clear. According to previous discussion, whether a truth-telling equilibrium exists depends on the distance between $a(0)$ and $a(1)$. Suppose the expert's signal is not very informative. Then, even if he completely reveals his information to the DM, his messages cannot have a large impact on the DM's choice of actions, i.e., the difference between $a(0)$ and $a(1)$ is small. In this case, truth telling cannot be sustained in equilibrium because type 0 expert would have an incentive to lie and induce the DM to choose $a(1)$ instead of $a(0)$. In constrast, if the expert observes a highly accurate signal, then he expects the DM's choice of action to be very sensitive to the messages that he sends, provided that the DM believes that he is telling the truth. In this case, type 0 expert would like to be separated from type 1 expert because $a(1)$ is too high to be profitable. Hence truth telling is sustainable in equilibrium.

When the expert's bias is higher than the threshold value $\overline{b}$, the IC constraint for type 0 expert is violated and a truth-telling equilibrium fails to exist. But does there exist an equilibrium in which some information is transmitted from the expert to the DM, that is, an equilibrium other than babbling? The discussion below shows that the answer is no.

To see this, fix the parameters p, q_1 and $b > \bar{b}$ and suppose there exists a non-babbling equilibrium. Since the expert's messages contain some information, the DM responds to the messages 0 and 1 with different actions in equilibrium, i.e., $a(m=1) \neq a(m=0)$. This immediately implies that type 1 expert does not play a mixed strategy in this equilibrium since he strictly prefers the higher action. WLOG, let's assume that type 1 expert sends message 1 with probability 1, i.e., $\sigma(1,1) = 1$. In this semi-separating equilibrium, type 0 expert must send both messages 0 and 1 with positive probability, i.e., $\sigma(0,1), \sigma(0,0) \in (0,1)$. Given these reporting strategies, if the DM receives message 0, she infers that it is sent by type 0 expert with probability 1 and if she receives message 1, she infers that it could be sent by either type of the expert with positive probability. Therefore, $a(m=0) = E(\omega|s_1 = 0)$ and $E(\omega|s_1 = 0) < a(m=1) < E(\omega|s_1 = 1)$. In this mixed strategy non-babbling equilibrium, type 0 expert is indifferent between the two actions $a(m=0)$ and $a(m=1)$ as defined above. It follows that the action $E(\omega|s_1 = 1)$ is too high to be profitable for type 0 expert and he strictly prefers the action $E(\omega|s_1 = 0)$ to the action $E(\omega|s_1 = 1)$. So type 0 expert's IC constraint for truth telling is satisfied, which contradicts our assumption that $b > \bar{b}$. Therefore, we have the following proposition.

Proposition 2.2. *A mixed strategy non-babbling equilibrium exists only when a truth-*

telling equilibrium exists. In this mixed strategy equilibrium, type 0 *expert is indifferent between sending message* 0 *and message* 1 *and sends both with positive probability.*

Hence, given the parameters p, q_1 and b, the most informative equilibrium involves either truth telling or babbling.

What about the players' preferences over different equilibria? Since the DM wants to elicit as much information as possible in order to make better decisions, she has the highest expected payoff in the most informative equilibrium. Given the assumption that the players' payoff functions take on the quadratic-loss form, it is not surprising that the expert also has a higher expected payoff (unconditional on his type) in a truth-telling equilibrium than in a babbling equilibrium. Moreover, if a mixed strategy non-babbling equilibrium coexists with a truth-telling equilibrium, then both types of the expert have (weakly) higher payoffs in the truth-telling equilibrium. To see why, recall that in a mixed strategy non-babbling equilibrium, type 0 expert is indifferent between sending message 0 (and being identified as type 0) and sending message 1 (and partially pooling with type 1). Since he is identified as type 0 in both equilibria, type 0 expert's expected payoff is the same in the truth-telling equilibrium and the semi-separating equilibrium. As to type 1 expert, he induces a lower action in the semi-separating equilibrium than in the truth-telling equilibrium. With an upward bias, type 1 expert has a strictly higher expected payoff in the truth-telling equilibrium. Hence, we have the following proposition.

Proposition 2.3. *The truth-telling equilibrium ex ante Pareto dominates the babbling* *equilibrium. It also Pareto dominates the mixed strategy non-babbling equilibrium, even in the interim sense.*

See appendix for proof.

2.4 Partially Informed DM: $\frac{1}{2} < q_2 < 1$

Now suppose the decision maker also observes an informative signal s_2, with its accuracy being $q_2 \in \left(\frac{1}{2}, 1\right)$.

We are going to consider the three games described before one by one.

2.4.1 Game Γ_1: Commitment – The DM truthfully reveals s_2

Suppose the DM reveals her signal to the expert before the expert sends his message. (Alternatively, suppose the expert observes the public signal s_2 as well as his private signal s_1 before communication.)

In this game, the expert's message can depend on s_2 as well as on s_1. Again, type 1 expert would like to convince the DM that $s_1 = 1$, independent of the realization of s_2. So the IC constraints for type 1 expert are never binding. To find the conditions for truth telling in this game, we just need to consider the IC constraints for type 0 expert. In Γ_1, there are two of them, one when $s_2 = 0$, and the other when $s_2 = 1$. In each case, we need type 0 expert not to have an incentive to lie to the DM. Let's use IC_0 and IC_1 to refer to the IC constraint for type 0 expert when $s_2 = 0$ and when

$s_2 = 1$, respectively.

Let

$$\begin{aligned} b_0 &= \frac{1}{2}\left(E\left(\omega|s_1=1,s_2=0\right)-E\left(\omega|s_1=0,s_2=0\right)\right) \\ &= \frac{1}{2}\left(\frac{pq_1\left(1-q_2\right)}{pq_1\left(1-q_2\right)+\left(1-p\right)\left(1-q_1\right)q_2}-\frac{p\left(1-q_1\right)\left(1-q_2\right)}{p\left(1-q_1\right)\left(1-q_2\right)+\left(1-p\right)q_1q_2}\right) \\ &= \frac{1}{2}\left(\frac{p\left(1-p\right)\left(1-q_2\right)q_2\left(2q_1-1\right)}{\left(pq_1\left(1-q_2\right)+\left(1-p\right)\left(1-q_1\right)q_2\right)\left(p\left(1-q_1\right)\left(1-q_2\right)+\left(1-p\right)q_1q_2\right)}\right)\end{aligned}$$

and

$$\begin{aligned} b_1 &= \frac{1}{2}\left(E\left(\omega|s_1=1,s_2=1\right)-E\left(\omega|s_1=0,s_2=1\right)\right) \\ &= \frac{1}{2}\left(\frac{pq_1q_2}{pq_1q_2+\left(1-p\right)\left(1-q_1\right)\left(1-q_2\right)}-\frac{p\left(1-q_1\right)q_2}{p\left(1-q_1\right)q_2+\left(1-p\right)q_1\left(1-q_2\right)}\right) \\ &= \frac{1}{2}\left(\frac{p\left(1-p\right)\left(1-q_2\right)q_2\left(2q_1-1\right)}{\left(pq_1q_2+\left(1-p\right)\left(1-q_1\right)\left(1-q_2\right)\right)\left(p\left(1-q_1\right)q_2+\left(1-p\right)q_1\left(1-q_2\right)\right)}\right).\end{aligned}$$

We have the following proposition.

Proposition 2.4. *Suppose $b \geq 0$. In Γ_1, there exists an equilibrium in which the expert reports s_1 truthfully when $s_2 = 0$ (i.e., $m(s_1, 0) = s_1$ for $s_1 \in \{0,1\}$) if and only if $b \leq b_0$ and there exists an equilibrium in which the expert reports s_1 truthfully when $s_2 = 1$ (i.e., $m(s_1, 1) = s_1$ for $s_1 \in \{0,1\}$) if and only if $b \leq b_1$.*

See appendix for proof.

Again, whether truth telling can happen in equilibrium depends on the difference in beliefs and therefore actions that the expert's messages induce, relative to the size of the expert's bias. As to the comparative statics, we have the following remarks.

(Proofs of the remarks can be found in the appendix.)

Remark 2.2. *$b_0(p, q_1, q_2) \geq b_1(p, q_1, q_2)$ if and only if $p \geq \frac{1}{2}$.*

This is quite intuitive. If the prior is such that $\omega = 1$ with a higher probability, then the difference in posterior beliefs that the expert's messages induce (if DM believes that the expert is telling the truth) is larger when the DM's signal is against the prior, i.e., when $s_1 = 0$. Therefore, the range of bias for which a truth-telling equilibrium exists is larger when $s_2 = 0$ than when $s_2 = 1$ if $p \geq \frac{1}{2}$. The result is the mirror image of the above if $p \leq \frac{1}{2}$.

Remark 2.3. *The signs of $\frac{db_0(p,q_1,q_2)}{dq_1}$ and $\frac{db_1(p,q_1,q_2)}{dq_1}$ are both positive.*

The intuition is similar to the baseline case where $q_2 = \frac{1}{2}$. Here, since s_1 and s_2 are conditionally independent, even with the knowledge of s_2, an expert with a highly accurate private signal expects that his messages will have a large impact on the DM's choice of action if they are believed by the DM to be truthful. Therefore, the ranges of bias for which a truth-telling equilibrium exists is increasing in the accuracy of the expert's signal, no matter what the realization of s_2 is.

Remark 2.4. *The sign of $\frac{db_0(p,q_1,q_2)}{dq_2}$ is the same as the sign of $(p - q_2)$ and the sign of $\frac{db_1(p,q_1,q_2)}{dq_2}$ is the same as the sign of $(1 - p - q_2)$.*

Therefore, when the prior is symmetric $\left(p = \frac{1}{2}\right)$, both $b_0(p, q_1, q_2)$ and $b_1(p, q_1, q_2)$ are decreasing in q_2, the accuracy of the DM's signal. When the prior is not symmetric $\left(p \neq \frac{1}{2}\right)$, at least one of b_0 and b_1 is decreasing in q_2. This means that an increase in the accuracy of the DM's signal may discourage the expert from telling the truth. This happens when an increase in q_2 lowers the impact of the expert's information (which is transmitted to the DM in a truth-telling equilibrium) on the DM's choice of action.

We have shown before that in the baseline case where $q_2 = \frac{1}{2}$, a mixed strategy non-babbling equilibrium exists only when there exists a truth-telling equilibrium. The same argument applies in Γ_1. Consider the subgames in Γ_1 after s_2 is revealed to the players. Since the DM has no private information in the subgames, each of the subgames is formally equivalent to a baseline game where $q_2 = \frac{1}{2}$, with an appropriately chosen parameter p. No matter s_2 is revealed to be 0 or 1, an equilibrium where the expert's private information is partially revealed to the DM exists only when there exists another equilibrium where the expert completely reveals his private information. Again, depending on the paramters, the most informative equilibrium involves either truth telling or babbling. And the Pareto ranking of the equilibria in each of the subgames is the same as in Proposition 2.3.

2.4.2 Game Γ_2: No Commitment – The DM keeps s_2 private

Let's turn to the case where the DM does not reveal her signal to the expert but keeps it private throughout the communication game.

Because the DM keeps her information private, the expert does not know for sure what action the DM will choose in response to his messages in equilibrium, even though the DM always plays a pure strategy. From the expert's point of view, his messages induce probability distributions of actions by the DM.

Again, to find the conditions for a truth-telling equilibrium, we need to consider the IC constraint for type 0 expert. Since the DM keeps s_2 private in Γ_2, there is only one IC constraint to consider. Call this constraint $IC_{private}$. Let $b_{private}$ $= \frac{1}{2}\frac{p(1-p)(2q_1-1)q_2(1-q_2)}{pq_1+(1-p)(1-q_1)}\Big(\frac{pq_1q_2+(1-p)(1-q_1)(1-q_2)}{(pq_1(1-q_2)+(1-p)(1-q_1)q_2)(p(1-q_1)(1-q_2)+(1-p)q_1q_2)}$ $+\frac{pq_1(1-q_2)+(1-p)(1-q_1)q_2}{(pq_1q_2+(1-p)(1-q_1)(1-q_2))(p(1-q_1)q_2+(1-p)q_1(1-q_2))}\Big)$. By analyzing $IC_{private}$, we establish the following proposition.

Proposition 2.5. *Suppose $b \geq 0$. In Γ_2, there exists a truth-telling equilibrium if and only if $b \leq b_{private}$.*

See appendix for proof.

The constraint $IC_{private}$ in Γ_2 is a convex combination of the constraints IC_0 and IC_1 in Γ_1. Accordingly, the threshold value $b_{private}$ is a convex combination of b_0 and b_1, which are derived from IC_0 and IC_1. In the following section, we will compare these constraints and discuss the economic implications.

Is the expert more likely to lie to someone who is well informed, or someone poorly informed? One may expect the expert to be more reluctant to lie when facing a DM who has accurate information of her own. However, the following result shows that the contrary is true in Γ_2.

Remark 2.5. $\frac{db_{private}(p,q_1,q_2)}{dq_2} < 0$.

See appendix for proof.

The more accurate the DM's signal is, the smaller the range of bias for which a truth-telling equilibrium exists when the DM keeps s_2 private.

As we explained before, whether or not a truth-telling equilibrium exists depends on the difference in the actions that the DM chooses in response to messages 0 and 1. When this difference in actions is sufficiently large relative to the expert's bias, type 0 expert does not have an incentive to lie and induce the DM to choose the higher action and therefore truth telling can happen in equilibrium. However, when the difference in actions is sufficiently small, the IC constraint for type 0 expert is violated and a truth-telling equilibrium fails to exist. When it is common knowledge that the DM has relatively accurate private information, the expert expects that his messages do not weigh very much in the DM's choice of action. This means that the DM's (expected) responses to different messages are not very far apart and henceforth, type 0 expert has an incentive to deviate from revealing his signal honestly and truth telling fails to be an equilibrium.

One interesting and perhaps surprising implication is that having a more accurate signal does not necessarily benefit the DM. It is possible that the loss of information from the expert more than offsets the gain from the increase in the accuracy of the DM's own signal. This will be illustrated in the example we provide in the next section.

Using a similar argument as in section 2.3, we can show that in Γ_2, if there exists a mixed strategy equilibrium in which the expert's information is partially transmitted to the DM, then it must be the case that $b < b_{private}$ and a truth-telling equilibrium exists as well. Therefore, given the parameters, the most informative equilibrium involves either full revelation or babbling on the expert's part.

2.4.3 Comparision of equilibria in Γ_1 and Γ_2

Both Γ_1 and Γ_2 are commonly observed in real life. We would imagine that at least sometimes the DM may have some discretion in choosing the communication environment. For example, if s_2 comes from the (truthful) report of a neutral third party, the DM may decide whether or not to allow the expert to be present when the third party makes the report. Alternatively, if s_2 is a public signal, then it may be possible for the DM to choose the timing of the communication. That is, she could decide whether to ask the expert to report his private observation of s_1 before or after the arrival of the public signal s_2. In yet another interpretation, s_2 could be the DM's private information that can be verified by the expert at no cost once revealed. Then, the choice between playing Γ_1 and Γ_2 is the same as the decision of whether or not to reveal s_2 to the expert[6].

If the DM could choose between Γ_1 and Γ_2, what would she do? With the analysis of Γ_1 and Γ_2 in previous sections, we can compare the two games. In particular, we

[6]In this interpretation, we assume that the DM's decision of whether or not to reveal s_2 is made before she learns the realization of s_2 so that the choice of "not revealing" does not signal anything to the expert.

can compare their effectiveness as mechanisms to facilitate information transmission from the expert to the DM. Note that the DM's expected utility is monotonically increasing in the amount of information she extracts from the expert and therefore she favors the environment that is most conducive to information transmission.

We find that whether the DM extracts more information from the expert (in the most informative equilibrium) in Γ_1 or Γ_2 depends on the parameters. Neither Γ_1 nor Γ_2 dominates the other in terms of facilitating information transmission.

First, let's fix p, q_1, q_2 and vary the expert's bias b. Suppose $p \geq \frac{1}{2}$, which implies that $b_1 \leq b_{private} \leq b_0$.

If $b \leq b_1$, then both IC_0 and IC_1 are satisfied. As a convex combination of IC_0 and IC_1, $IC_{private}$ is also satisfied. This means that a truth-telling equilibrium exists in both Γ_1 and Γ_2 and the DM extracts the maximal amount of information in both games.

If $b_1 < b \leq b_{private}$, then IC_0 and $IC_{private}$ are satisfied while IC_1 is violated. This implies that if the DM reveals s_2 to the expert before he reports, then in the most informative equilibrium, the expert truthfully reveals his signal if $s_2 = 0$ but babbles if $s_2 = 1$. In constrast, if the DM does not reveal s_2 to the expert, then there exists an equilibrium in which the expert truthfully reveals his signal. Therefore, the DM can extract more information from the expert by keeping her signal private when b lies in this range.

If $b_{private} < b \leq b_0$, then only IC_0 is satisfied. In Γ_1, the expert still truthfully reveals his signal if $s_2 = 0$. However, only babbling can happen in equilibrium in Γ_2 since $IC_{private}$ is violated. Therefore, when b lies in this range, by commiting to revealing s_2 to the expert, the DM can extract useful information from him in the event that $s_2 = 0$ (which is the less likely event *a priori*) while she can extract no information if she keeps s_2 private.

Finally, if $b_0 < b$, then all three incentive constraints are violated and the expert babbles in both Γ_1 and Γ_2.

We just provided a comparison between Γ_1 and Γ_2 by varying the size of the expert's bias. Suppose the DM can choose between Γ_1 and Γ_2, then we have the following result. She is indifferent between the two games if the bias of the expert is extreme (either very small or very large) but strictly prefers one over the other if the bias is in the intermediate range. Specifically, as the bias increases from a small value to a moderate value, the DM chooses keeping her signal private over revealing it to the expert. However, as the bias gets even larger, her preference switches from keeping the signal private to revealing it.

It is also interesting to compare the two games by fixing p, q_1, b and varying q_2, the accuracy of the DM's signal. The analysis is parallel to the above since there is a one-to-one mapping between q_2 and each of the threshold values b_0, b_1 and $b_{private}$. Specifically, when q_2 is either very low or very high, the DM is indifferent between the two games. However, when q_2 is in the intermediate range, the DM prefers one mechanism over the other.

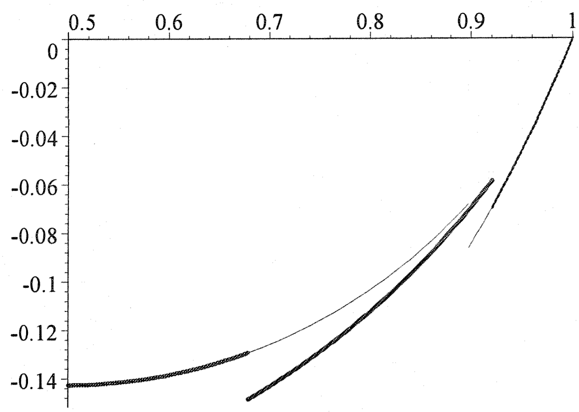

Figure 2.1: the DM's expected equilibrium payoff

To illustrate, consider the following example.

Example 2.1. *Suppose $p = 0.7$, $q_1 = 0.8$ and $b = 0.2$.*

It is useful to note that when $q_2 = 0.67839$, $b_1 = 0.2$, when $q_2 = 0.91990$, $b_0 = 0.2$ and when $q_2 = 0.89711$, $b_{private} = 0.2$.

Figure 2.1 shows the DM's expected payoff as a function in q_2 in the most informative equilibirum in Γ_1 and Γ_2. The thick plot is for Γ_1 and the thin one is for Γ_2. The two plots coincide for extreme values of q_2, but they are different for intermediate values of q_2.

Let's consider Γ_1 first. The DM's payoff depends on how much information she has when making a decision. The increase in q_2 has two effects on her payoff. On the one hand, it benefits the DM since the signal she directly observes is more accurate. On the other hand, it could be disadvantageous because the increase in q_2 may discourage *the expert from revealing his information, as we have shown earlier. Combining these two effects, we see that the DM's expected payoff is NOT monotonically increasing in q_2. Specifically, the thick curve has two discontinuities at $q_2 = 0.67839$ and $q_2 = 0.91990$. These are the two points at which the information content of the expert's communication changes. They divide the interval $[0.5, 1]$ into three ranges. Within each of the three ranges, the DM's payoff is increasing in q_2.*

The analysis is quite similar in Γ_2. The thin curve has only one discontinuity at $q_2 = 0.89711$. If $q_2 \in [0.5, 0.89711]$, a truth-telling equilibrium exists in Γ_2; if $q_2 \in (0.89711, 1]$, only babbling equilibrium exists.

As we can see from the figure, when $q_2 \in (0.67839, 0.89711]$, the DM extracts more information in Γ_2 by keeping s_2 private and therefore enjoys a higher expected utility. However, when $q_2 \in (0.89711, 0.91990]$, she could extract more information in Γ_1 by revealing s_2 to the expert.

2.4.4 Game Γ_3: No commitment – The DM strategically communicates to the expert

Now suppose we add an extra round of communication in which the DM strategically and costlessly communicate her observation of s_2 to the expert before the expert reports to her. How does this extra round of communication affect the DM's ability to extract information from the expert?

Define equilibrium outcomes in all three games on the payoff relevant space $S_1 \times$

$S_2 \times A$. Denote by $EO(\Gamma_i)$ the set of equilibrium outcomes for game Γ_i, $i = 1, 2, 3$. Clearly, the set depends on the parameter values p, q_1, q_2 and b.

Lemma 2.2. *Fix p, q_1, q_2 and b. $EO(\Gamma_2) \subseteq EO(\Gamma_3)$.*

The intuition behind this lemma is simple. Consider any equilibrium in Γ_2. Call it $E(\Gamma_2)$. Since the messages used by the DM in the first round of communication in Γ_3 are "cheap talk," there always exists an equilibrium in Γ_3 in which the DM babbles in the first round of communication and in effect keeps her signal s_2 private and the players follow the strategies prescribed in $E(\Gamma_2)$ in the continuation of Γ_3. Therefore, when given an opportunity to costlessly communicate to the expert first, the DM can do at least as well as when she keeps her information private.

What is more surprising is the following result.

Proposition 2.6. *Fix p, q_1, q_2 and b. The DM cannot extract more information from the expert in Γ_3 than in Γ_2.*

See appendix for proof.

This proposition says that the extra round of strategic communication does not help the DM. Her expected payoff in the most informative equilibrium in Γ_3 is the same as that in Γ_2.

What is the intuition behind this result? In Γ_3, the sole purpose of the first round of (the DM's) communication is to elicit information from the expert in the second round[7]. If one particular message does strictly better than the other for this purpose, the other message cannot be sent by the DM in equilibrium and pooling happens in the first round of communication. In this case, the set of equilibrium outcomes in Γ_3 is the same as that in Γ_2. If both messages 0 and 1 are sent in equilibrium in the first round of communication in Γ_3, then it must be the case that the expert responds to messages 0 and 1 in the same way. That is, the expert reveals the same amount of information in the most informative equilibrium in the continuation of Γ_3, no matter what message he receives from the DM. It follows that the expert will reveal the same amount of information even if the DM keeps her signal private. Again, in this case, the DM does no better in Γ_3 than in Γ_2. Put it in a nutshell, since the DM's preference for information (at the final stage of the game, when she makes decisions) is monotone (the more information the better), she cannot effectively communicate her private information to the expert without the commitment to revealing s_2 truthfully and is therefore unable to extract more information from the expert.

2.5 Equilibrium Selection

So far we have focused on finding the conditions for the existence of truth-telling equilibria. In our comparative statics analysis and comparison of different games,

[7]This is different from Aumann and Hart's "long cheap talk" (2003). They consider two-person games in which one player is better informed but both players take payoff-relevant actions and they allow the players to talk as long as they wish. In their model, different rounds of cheap talk can help the players agree on compromises as well as reveal substantive information.

we have implicitly focused on the "most informative" equilibrium. But is there any reason why we should select it among the multiple equilibria that may exist? When the parameters are such that a truth-telling equilibrium exists as well as a babbling equilibrium, why should the former be considered more "reasonable"?

From a welfare perspective, we have shown before that all types of players have higher expected payoffs in a truth-telling equilibrium than in a mixed strategy non-babbling equilibrium. Moreover, the truth-telling equilibrium *ex ante* Pareto dominates the babbling equilibrium. Therefore, if we believe that the players would coordinate on a Pareto dominant equilibrium, then the most informative equilibrium should prevail. However, Pareto dominance (especially in the *ex ante* sense) is not generally accepted as a formal refinement criterion. Below we'll use two different approaches and see whether we can eliminate "unreasonable" equilibria in our model.

The first selection criterion is Farrell's (1993) "neologism-proofness." To understand this concept, let's fix an equilibrium of a cheap talk game. Consider an announcement by the sender "my type is in the set X." This announcement is a neologism if it is not sent in the candidate equilibrium. A neologism is deemed credible (relative to the equilibrium) if it is precisely those types in the set X that receive strictly higher payoffs than their equilibrium payoffs if the neologism is believed by the receiver. An equilibrium is neologism-proof if there does not exist any credible neologism relative to it.

If we apply the neologism-proofness criterion to the games under consideration, we have the following result.

Proposition 2.7. *In game Γ_2 and the subgames of Γ_1 and Γ_3 after the DM sends the expert a message, a truth-telling equilibrium is always neologism-proof for any parameter values; a mixed strategy non-babbling equilibrium is never neologism-proof; a babbling equilibrium is neologism-proof if the following two conditions hold:* (1) *type* 0 *expert (strictly) prefers being perceived as type* 1 *than pooling with type* 1 *and* (2) *type* 0 *expert (weakly) prefers pooling with type* 1 *than being identified as type* 0.

In a truth-telling equilibrium, two announcements "my type is 1" and "my type is 0" are made in equilibrium and the only neologism is "my type is in $\{0, 1\}$." Since type 1 expert would never want to make this announcement, it is not credible and therefore a truth-telling equilibrium is neologism-proof.

As we have shown before, in a mixed strategy non-babbling equilibrium, type 0 expert randomizes between sending message 0 and message 1 and type 1 expert sends only one message. Therefore, "my type is 1" is a neologism. It is also a credible one because only type 1 expert gets a strictly higher payoff than the equilibrium payoff if the neologism is believed.[8] Hence the equilibrium is not neologism-proof.

Now consider a babbling equilibrium. Relative to a babbling equilibrium, there are two neologisms: "my type is 1" and "my type is 0." First consider the neologism

[8] In the candidate equilibrium, type 0 is indifferent between being identified as type 0 and (partially) pooling with type 1. So the action that the DM takes when believing that the expert is type 1 is too high to be profitable. And obviously, with an upward bias, type 1 expert would be better off if identified as type 1.

"my type is 1." Type 1 expert would want to make the announcement since he gets a strictly higher payoff if it is believed. Moreover, if type 0 expert prefers playing the babbling equilibrium than being perceived as type 1, then it is a credible neologism. As to the other neologism "my type is 0," it is clear that type 1 expert would not want to make the announcement and type 0 would if and only being identified as type 0 gives him a (strictly) higher payoff than babbling and in that case, this is a credible neologism. For the babbling equilibrium to be neologism-proof, there cannot exist any credible neologism and hence the result in Proposition 2.7.

One well-recognized problem with the selection criterion of "neologism-proofness" is that in certain games, "neologism-proof" equilibria do not exist. In fact, take any of the three games considered in this paper, there are parameters with which "neologism-proof" equilibria fail to exist. To see this, fix a game and suppose the parameters are such that only babbling on the expert's part can happen in equilibrium. The two neologisms relative to the equilibrium are "my type is 0" and "my type is 1." Are they credible? As we have shown before, when babbling is the unique equilibrium outcome, the IC constraint for type 0 expert does not hold and he prefers being perceived as type 1 to being identified as type 0. It follows that he prefers babbling to being identified as type 0 as well. Hence the neologism "my type is 0" is not credible. However, in light of Proposition 2.7, the other neologism "my type is 1" is credible provided that type 0 prefers babbling to being perceived as type 1. Hence, when the expert's bias is sufficiently large (so that babbling is the unique equilibrium outcome) but not too large (so that "my type is 1" is a credible neologism), no equilibrium passes the "neologism-proofness" test.

The second approach of refinement that we take is similar to what we did in Chapter 1. Imagine there is a small probability that the players do not behave strategically. Specifically, suppose with probability θ, the expert is an "honest" type who always tells the truth about his signal s_1 through the messages m and with probability λ, the DM is a "naive" type who believes (potentially incorrectly) that the messages sent by the expert are truthful and chooses her actions accordingly. Assume that the probability distributions of the players' types (strategic or "behavioral") are independent and they are also independent of the distribution of ω.

When we perturb our games with small probabilities of the behavioral types, we can conduct a robustness test of the multiple equilibria that exist in the unperturbed game. Similar to Chapter 1, we focus on the class of "monotonic" equilibria. However, since we are considering a model where the state/message space is binary instead of continuous, the definition of "monotonicity" is modified. Consider the strategic expert's strategy, $\sigma_i(s_1, m, \cdot)$, in a perturbed version of the game Γ_i. Fix the other variables that $\sigma_i(s_1, m, \cdot)$ may depend on. Monotonicity requires that $\sigma_i(0, 0, \cdot) \geq \sigma_i(1, 0, \cdot)$ (or equivalently, $\sigma_i(1, 1, \cdot) \geq \sigma_i(0, 1, \cdot)$. That is, the probability with which type 0 expert sends message 0 is higher than the probability with which type 1 expert sends message 0.

Consider an unperturbed game Γ (Γ can be Γ_2 or the subgames in Γ_1 and Γ_3 after the DM sends a message to the expert.) Fix an equilibrium $\hat{E}$ in game Γ. We say that $\hat{E}$ is robust to the perturbation if there exist converging sequences $\{\theta_m\}_{m=1}^{\infty}$ and $\{\lambda_n\}_{n=1}^{\infty}$ ($\lim_{m\to\infty}\theta_m = 0$ and $\lim_{n\to\infty}\lambda_n = 0$) such that $\hat{E}$ is the limit monotonic equilibrium of the sequence of games $\{\Gamma_{\theta_m,\lambda_n}\}$. We have the following proposition.

Proposition 2.8. *The equilibrium $\hat{E}$ in game Γ is robust to the perturbation if and only if one of the following conditions is satisfied:*

1. $\hat{E}$ is a truth-telling equilibrium.

2. $\hat{E}$ is a mixed strategy non-babbling equilibrium in which type 1 expert sends $m = 1$ with probability 1 and type 0 expert sends both $m = 1$ and $m = 0$ with strictly positive probability.

3. $\hat{E}$ is a babbling equilibrium in which both types of expert send $m = 1$ with probability 1 and type 0 expert strictly prefers babbling to being identified as type 0.

See appendix for proof.

As we see, the above two approches of equilibrium selection generate different predictions.

The concept of "neologism-proofness" yields stronger predictions for the games under consideration. However, sometimes it is too strong that no equilibrium is neologism-proof.

The "perturbation" approach does not have the non-existence problem. To see this, fix a game and suppose the parameters are such that a truth-telling equilibrium fails to exist (which implies that a mixed strategy non-babbling equilibrium does not exist either), so the unique equilibrium outcome involves babbling. Since type 0 expert's IC constraint for truth telling is violated, type 0 expert prefers being (incorrectly) perceived as type 1 to being identified as type 0 by the DM. This implies that he strictly prefers pooling with type 1 to being identified as type 0.[9] According to Proposition 2.8, the babbling equilibrium where both type of the expert send the high message is robust to the perturbation.

Although the problem of non-existence does not arise, the perturbation approach is not fully satisfactory for the purpose of equilibrium selection either because for certain parameter values, it still admits multiple equilibria.[10] It is interesting to note that if there are multiple equilibria that pass our robustness test, then among them, only the truth-telling equilibrium is neologism-proof. To see this, first note that the mixed strategy non-babbling equilibrium is not neologism-proof. Now consider the babbling equilibrium where both types of the expert send message 1 and type 0 expert prefers babbling (and pooling with type 1) to being identifed as type 0. Since a truth-telling equilibrium exists, type 0 expert must prefer being identifed as type 0 to being

[9]The action that type 0 expert induces the DM to take when pooling with type 1 is between the actions he induces when identified as type 0 and when perceived as type 1. Since type 0's IC constraint for truth telling does not hold, the action that he induces when perceived as type 1 is *not* too high to be profitable. Accordingly, type 0 expert must prefer the higher action that he induces when pooling with type 1 to the action he induces when identified as type 0.

[10]When applied to the Crawford and Sobel (1982) model with a continuous state space, the perturbation approach generates strong results. However, when the state space is binary, the results are weaker. In both cases, we find that in a monotonic equilibrium of a perturbed game, the expected equilibrium payoff for the lowest type of expert (type 0) must be at least as high as the payoff he gets if identified by the DM as the lowest type. This condition implies uniqueness in the case of a continuous state space, but not necessarily so when the state space is binary.

perceived as type 1. It follows that type 0 expert prefers babbling to being perceived as type 1. This implies that relative to the babbling equilibrium, “my type is 1” is a credible neologism and therefore the babbling equilibrium is not neologism-proof.

2.6 Conclusion

In order to make good decisions, people often seek advice and information from others (whom are usually referred to as experts in the literature), who do not necessarily share the same interests.

In a one-shot communication game between an expert and a DM where reputational concerns are absent, the expert's incentives for information transmission are largely determined by the actions that the DM takes in response to his messages, which are in turn affected by the private information that the DM possesses, as we assume that the DM is partially informed as well.

When taking this strategic interaction into consideration, we find that the DM does not necessarily benefit from having more accurate information of her own since it may have the adverse effect of discouraging the expert from revealing his private information in equilibrium.

Even in the simple setting that we consider, the DM's choice between revealing her information to the expert before the expert reports to her or keeping it private is not trivial. Which alternative yields a higher expected payoff to the decision maker depends on how aligned the players' interests are. Furthermore, given that the DM always wants to extract as much information as possible from the expert, even if she has an opportunity to communicate (without committing to telling the truth) to the expert before he reports, the DM cannot extract more information than if she keeps her own signal private.

2.7 Appendix

Proof of Remark 2.1. We need to show that $\frac{d\bar{b}}{dq_1} > 0$.

$\frac{d\bar{b}}{dq_1} = \frac{1}{2}(1-p)\,p\frac{-2q_1+1+8pq_1-2p-8pq_1^2+2q_1^2-8p^2q_1+8p^2q_1^2+2p^2}{(2pq_1+1-q_1-p)^2(-p+2pq_1-q_1)^2}$.

Since $\frac{1}{2}\frac{(1-p)p}{(2pq_1+1-q_1-p)^2(-p+2pq_1-q_1)^2} > 0$, all we need to show is $-2q_1+1+8pq_1-2p-8pq_1^2+2q_1^2-8p^2q_1+8p^2q_1^2+2p^2 > 0$. Note that

$$\begin{aligned} & -2q_1+1+8pq_1-2p-8pq_1^2+2q_1^2-8p^2q_1+8p^2q_1^2+2p^2 \\ = \; & 1+2(2pq_1-q_1+1-p)(2pq_1-q_1-p) \\ = \; & 1-2(pq_1+(1-p)(1-q_1))(p(1-q_1)+(1-p)q_1) > 0, \end{aligned}$$

since $(pq_1+(1-p)(1-q_1))(p(1-q_1)+(1-p)q_1) < \frac{1}{2}$. Hence $\frac{d\bar{b}}{dq_1} > 0$. □

Proof of Proposition 2.3. We have shown in the main text that the DM and both types of the expert has higher payoff in the truth-telling equilibrium than in the mixed strategy non-babbling equilibrium (type 1 expert has strictly higher payoff). We need to compare the players' expected equilibrium payoffs in the truth-telling equilibrium and the babbling equilibrium.

Obviously, the DM's expected payoff is higher in the truth-telling equilibrium since more information is transmitted. Now consider the expert.

In a truth-telling equilibrium, the expert's expected payoff is

$$prob(s_1=1)\left(-b^2-var(\omega|s_1=1)\right)+prob(s_1=0)\left(-b^2-var(\omega|s_1=0)\right)$$

$$= -b^2 - prob\,(s_1 = 1)\, var\,(\omega | s_1 = 1) - prob\,(s_1 = 0)\, var\,(\omega | s_1 = 0)\,.$$

In a babbling equilibrium, however, the expert's expected payoff is equal to

$$\begin{aligned} & p\left(-\left(p-1-b\right)^2\right) + (1-p)\left(-\left(p-0-b\right)^2\right) \\ = \quad & -b^2 - var\,(\omega) \end{aligned}$$

Since $prob\,(s_1 = 1)\, var\,(\omega | s_1 = 1) + prob\,(s_1 = 0)\, var\,(\omega | s_1 = 0) < var\,(\omega)$ when $q_1 > \frac{1}{2}$, the expert has a higher expected payoff in the truth-telling equilibrium. □

Proof of Proposition 2.4. Suppose there exists an equilibrium in which the expert reports s_1 truthfully when $s_2 = 0$. In this equilibrium, the DM's action strategy must satisfy $a\,(m = 0, s_2 = 0) = E\,(\omega | s_1 = 0, s_2 = 0)$ and $a\,(m = 1, s_2 = 0) = E\,(\omega | s_1 = 1, s_2 = 0)$. The constraint IC_0 requires that

$$E(U^E\,(a\,(0,0)\,, \omega, b)\,| s_1 = 0, s_2 = 0) > E(U^E\,(a\,(1,0)\,, \omega, b)\,| s_1 = 0, s_2 = 0),$$

which is equivalent to

$$\begin{aligned} & E(-\left(a\,(0,0) - \omega - b\right)^2 | s_1 = 0, s_2 = 0) \geq E(-\left(a\,(1,0) - \omega - b\right)^2 | s_1 = 0, s_2 = 0) \Rightarrow \\ & -\left(E\,(\omega | s_1 = 0, s_2 = 0) - E\,(\omega | s_1 = 0, s_2 = 0) - b\right)^2 - Var\,(\omega | s_1 = 0, s_2 = 0) \\ & \geq -\left(E\,(\omega | s_1 = 1, s_2 = 0) - E\,(\omega | s_1 = 0, s_2 = 0) - b\right)^2 - Var\,(\omega | s_1 = 0, s_2 = 0) \Rightarrow \end{aligned}$$

$$-b^2 \geq -\left(\frac{pq_1\,(1-q_2)}{pq_1\,(1-q_2) + (1-p)\,(1-q_1)\,q_2} - \frac{p\,(1-q_1)\,(1-q_2)}{p\,(1-q_1)\,(1-q_2) + (1-p)\,q_1 q_2} - b\right)^2$$

When $b \geq 0$, the above inequality can be simplified as $b \leq b_0$.

If $b > b_0$, then IC_0 is violated and truth telling cannot happen in equilibrium when $s_2 = 0$.

Suppose there exists an equilibrium in which the expert reports s_1 truthfully when $s_2 = 1$. In this equilibrium, the DM's action strategy must satisfy $a(m = 0, s_2 = 1) = E(\omega | s_1 = 0, s_2 = 1)$ and $a(m = 1, s_2 = 1) = E(\omega | s_1 = 1, s_2 = 1)$. The constraint IC_1 requires that

$$E(U^E(a(0,1), \omega, b) | s_1 = 0, s_2 = 1) > E(U^E(a(1,1), \omega, b) | s_1 = 0, s_2 = 1),$$

When $b \geq 0$, the above inequality can be simplified as $b \leq b_1$. Also, if $b > b_1$, then IC_1 is violated and truth telling cannot happen in equilibrium when $s_2 = 1$. Hence the result. □

Proof of Remark 2.2. $b_0 = \frac{1}{2} \frac{p(1-p)q_2(1-q_2)(2q_1-1)}{(pq_1(1-q_2)+(1-p)(1-q_1)q_2)(p(1-q_1)(1-q_2)+(1-p)q_1q_2)}$ and

$b_1 = \frac{1}{2} \frac{p(1-p)q_2(1-q_2)(2q_1-1)}{(pq_1q_2+(1-p)(1-q_1)(1-q_2))(p(1-q_1)q_2+(1-p)q_1(1-q_2))}$.

The numerators are the same. If we compare the denominators, we see that

$$\begin{aligned} & (pq_1(1-q_2) + (1-p)(1-q_1)q_2)(p(1-q_1)(1-q_2) + (1-p)q_1q_2) - \\ & (pq_1q_2 + (1-p)(1-q_1)(1-q_2))(p(1-q_1)q_2 + (1-p)q_1(1-q_2)) \\ = \; & q_1(1-q_1)(2q_2-1)(1-2p) \end{aligned}$$

Given that $q_1, q_2 \in \left(\frac{1}{2}, 1\right)$, we have $b_0 \geq b_1$ if and only if $p \geq \frac{1}{2}$. □

Proof of Remark 2.3. Consider b_0 first. Since

$$\frac{d\left(\frac{1}{(pq_1(1-q_2)+(1-p)(1-q_1)q_2)(p(1-q_1)(1-q_2)+(1-p)q_1q_2)}\right)}{dq_1}$$

$$=\frac{(p-q_2)^2(2q_1-1)}{(pq_1(1-q_2)+(1-p)(1-q_1)q_2)^2(p(1-q_1)(1-q_2)+(1-p)q_1q_2)^2},\text{ we have}$$

$$\frac{db_0}{dq_1}=\frac{p(1-p)q_2(1-q_2)}{2}\Big(\frac{(p-q_2)^2(2q_1-1)^2}{(pq_1(1-q_2)+(1-p)(1-q_1)q_2)^2(p(1-q_1)(1-q_2)+(1-p)q_1q_2)^2}$$

$$+\frac{2}{(pq_1(1-q_2)+(1-p)(1-q_1)q_2)(p(1-q_1)(1-q_2)+(1-p)q_1q_2)}\Big)>0.$$

Similarly, since

$$\frac{d\left(\frac{1}{(pq_1q_2+(1-p)(1-q_1)(1-q_2))(p(1-q_1)q_2+(1-p)q_1(1-q_2))}\right)}{dq_1}$$

$$=\frac{(1-p-q_2)^2(2q_1-1)}{(pq_1q_2+(1-p)(1-q_1)(1-q_2))^2(p(1-q_1)q_2+(1-p)q_1(1-q_2))^2},\text{ we have}$$

$$\frac{db_1}{dq_1}=\frac{p(1-p)q_2(1-q_2)}{2}\Big(\frac{(1-p-q_2)^2(2q_1-1)^2}{(pq_1q_2+(1-p)(1-q_1)(1-q_2))^2(p(1-q_1)q_2+(1-p)q_1(1-q_2))^2}$$

$$+\frac{2}{(pq_1q_2+(1-p)(1-q_1)(1-q_2))(p(1-q_1)q_2+(1-p)q_1(1-q_2))}\Big)>0.$$ □

Proof of Remark 2.4. $\frac{db_0}{dq_2}=\frac{q_1(1-q_1)(p(1-q_2)+(1-p)q_2)(p-q_2)}{(pq_1(1-q_2)+(1-p)(1-q_1)q_2)^2(p(1-q_1)(1-q_2)+(1-p)q_1q_2)^2}$. Since all the other terms are positive, the sign of $\frac{db_0(p,q_1,q_2)}{dq_2}$ is the same as the sign of $(p-q_2)$.

$\frac{db_1}{dq_2}=\frac{q_1(1-q_1)(pq_2+(1-p)(1-q_2))(1-p-q_2)}{(pq_1q_2+(1-p)(1-q_1)(1-q_2))^2(p(1-q_1)q_2+(1-p)q_1(1-q_2))^2}$. Similarly, since all the other terms are positive, the sign of $\frac{db_0(p,q_1,q_2)}{dq_2}$ is the same as the sign of $(1-p-q_2)$. □

Proof of Proposition 2.5. Suppose there exists a truth-telling equilibrium in Γ_2. In this equilibrium, the DM's action strategy satisfies $a\,(m=0,s_2=0)$ $=E\,(\omega|s_1=0,s_2=0)$ and $a\,(m=1,s_2=0)=E\,(\omega|s_1=1,s_2=0)$.

The constraint $IC_{private}$ requires that

$$prob\,(s_2=0|s_1=0)\left(E(U^E\,(a\,(0,0)\,,\omega,b)\,|s_1=0,s_2=0)\right)$$

$$+prob\,(s_2=1|s_1=0)\left(E(U^E\,(a\,(0,1)\,,\omega,b)\,|s_1=0,s_2=1)\right)$$

$$\geq \; prob\left(s_2=0|s_1=0\right)\left(E(U^E\left(a\left(1,0\right),\omega,b\right)|s_1=0,s_2=0)\right)$$
$$+prob\left(s_2=1|s_1=0\right)\left(E(U^E\left(a\left(1,1\right),\omega,b\right)|s_1=0,s_2=1)\right).$$

Simplifying this constraint, we have

$$\begin{aligned}-b^2 \;\geq\; & -prob\left(s_2=0|s_1=0\right)\Big(\frac{pq_1\left(1-q_2\right)}{pq_1\left(1-q_2\right)+\left(1-p\right)\left(1-q_1\right)q_2} \\ & -\frac{p\left(1-q_1\right)\left(1-q_2\right)}{p\left(1-q_1\right)\left(1-q_2\right)+\left(1-p\right)q_1q_2}-b\Big)^2 \\ & -prob\left(s_2=1|s_1=0\right)\Big(\frac{pq_1q_2}{pq_1q_2+\left(1-p\right)\left(1-q_1\right)\left(1-q_2\right)} \\ & -\frac{p\left(1-q_1\right)q_2}{p\left(1-q_1\right)q_2+\left(1-p\right)q_1\left(1-q_2\right)}-b\Big)^2,\end{aligned}$$

where $prob\left(s_2=0|s_1=0\right)=\frac{p(1-q_1)(1-q_2)+(1-p)q_1q_2}{p(1-q_1)+(1-p)q_1}$ and $prob\left(s_2=1|s_1=0\right)=\frac{p(1-q_1)q_2+(1-p)q_1(1-q_2)}{p(1-q_1)+(1-p)q_1}$.

Further simplification shows that the constraint is equivalent to $b\leq b_{private}$ if $b\geq 0$. This proves that $b\leq b_{private}$ is a necessary condition for a truth-telling equilibrium to exist in Γ_2. It is also a sufficient condition because type 0 expert does not have an incentive to deviate from telling the truth if $IC_{private}$ is satisfied and type 1 expert's IC constraint is not binding. □

Proof of Proposition 2.6. WLOG, assume $p\geq\frac{1}{2}$, which implies that $b_0\geq b_{private}\geq b_1$. Fix p,q_1,q_2.

If $b\leq b_{private}\left(p,q_1,q_2\right)$, then a truth-telling equilibrium exists in Γ_2. Since the DM extracts the maximal amount of information fromt the expert in Γ_2, he cannot do better in Γ_3.

Now suppose $b > b_{private}(p, q_1, q_2)$. So there is no information transmitted in Γ_2.

Suppose $b > b_0 \geq b_1$. Then clearly the DM cannot extract any information from the expert in Γ_3 either. If $b > b_0$, for any belief that the expert may have over the distribution of s_2, the bias b is too high for the IC constraint to hold. As shown before, when truth telling fails to be an equilibrium, babbling is the unique equilibrium outcome.

Suppose $b_0 \geq b > b_{private} \geq b_1$. We can show by contradiction that no information can be transmitted from the expert to the DM in Γ_3. First, observe that in a PBE in Γ_3, the expert has to reveal the same amount of information in response to any of the messages sent with positive probability by the DM (and the information revealed by the expert on the equilibrium path has to be at least as much the information revealed off the equilibrium path). Also, recall that an equilibrium where partial information is transmitted from the expert to the DM exists only if a truth-telling equilibrium exists. Let's suppose in Γ_3 an equilibrium exists in which the expert truthfully reveals s_1 upon hearing either message 1 or message 0 from the DM. Then, for the IC constraints to hold, type 0 expert's posterior belief on the distribution of s_2 upon hearing any one of the DM's messages must be such that with probability strictly higher than $prob(s_2 = 0|s_1 = 0)$, $s_2 = 0$. Obviously, this can never be true with any admissible (mixed) strategy of the DM. Therefore, the expert can only babble in Γ_3 if $b_0 \geq b > b_{private} \geq b_1$. The DM cannot extract more information by strategically sending a message about her signal than by keeping it private. □

Proof of Proposition 2.8. We will prove the proposition for the case where Γ is the subgame of Γ_1 after s_2 is revealed to the expert to be equal to 0 because this case involves the simplest notation. The same argument applies in other games.

Let $\sigma_{\lambda,\theta}$ and $a_{\lambda,\theta}$ denote the strategic expert and the strategic DM's strategies in the perturbed game $\Gamma_{\lambda,\theta}$.

Suppose $\hat{E}$ is a truth-telling equilibrium in the unperturbed game Γ. That is, in $\hat{E}$, $\sigma(0,0)$[11] $= 1$, $\sigma(1,1) = 1$, $a(0) = E(\omega|s_1 = 0, s_2 = 0)$, $a(1) = E(\omega|s_1 = 1, s_2 = 0)$.

Now consider a perturbed game $\Gamma_{\lambda,\theta}$. It is easy to see that we have an equilibrium in $\Gamma_{\lambda,\theta}$ if the strategic players keep their strategies prescribed in $\hat{E}$ since the strategies of the strategic types and the behavioral types are the same. Hence a truth-telling equilibrium is robust to the perturbation.

Suppose $\hat{E}$ is a mixed strategy non-babbling equilibrium in the unperturbed game Γ. That is, in $\hat{E}$, $\sigma(0,0)$, $\sigma(0,1) \in (0,1)$ and $\sigma(1,1) = 1$. Also, we have $a(0) = E(\omega|s_1 = 0, s_2 = 0) = \frac{p(1-q_1)(1-q_2)}{p(1-q_1)(1-q_2)+(1-p)q_1q_2}$ and $a(1) = \frac{p(1-q_2)(q_1+(1-q_1)\sigma(0,1))}{p(1-q_2)(q_1+(1-q_1)\sigma(0,1))+(1-p)q_2(q_1\sigma(1,1)+(1-q_1))}$. Type 0 expert is indifferent between $a(0)$ and $a(1)$.

Due to continuity, we can show that for any δ_1, δ_2, δ_3, $\delta_4 > 0$, there exist ε_1, ε_2 s.t. for any $\lambda \in (0, \varepsilon_1)$ and $\theta \in (0, \varepsilon_2)$, there exist a strategy profile $(\sigma_{\lambda,\theta}(\cdot), a_{\lambda,\theta}(\cdot))$

[11] We suppress the dependence of the strategies on s_2 because in Γ, s_2 is already revealed to be equal to 0.

that satisfy $|\sigma_{\lambda,\theta}(0,0) - \sigma(0,0)| < \delta_1$, $|\sigma_{\lambda,\theta}(0,1) - \sigma(0,1)| < \delta_2$, $\sigma_{\lambda,\theta}(1,1) = 1$, $|a_{\lambda,\theta}(0) - a(0)| < \delta_3$ and $|a_{\lambda,\theta}(1) - a(1)| < \delta_4$ and $(\sigma_{\lambda,\theta}(\cdot), a_{\lambda,\theta}(\cdot))$ is a monotonic equilibrium in $\Gamma_{\lambda,\theta}$. Hence, a mixed strategy non-babbling equilibrium is robust to the perturbation.

Suppose $\hat{E}$ is a babbling equilibrium in Γ_2 in which both types of expert send $m = 1$ and type 0 expert strictly prefers babbling to being identified as type 0. That is, in $\hat{E}$, $\sigma(0,1) = \sigma(1,1) = 1$ and $a(1) = E(\omega|s_2 = 0) = \frac{p(1-q_2)}{p(1-q_2)+(1-p)q_2}$. Assume that $a(0) = E(\omega|s_1 = 0, s_2 = 0)$. By assumption, type 0 expert prefers $a(1)$ to $a(0)$.

Now consider a perturbed game $\Gamma_{\lambda,\theta}$. Suppose in $\Gamma_{\lambda,\theta}$, the strategic expert follows his strategy in $\hat{E}$, i.e., $\sigma_{\lambda,\theta}(\cdot) = \sigma(\cdot)$. Then, the best response for the strategic DM satisfies $a_{\lambda,\theta}(0) = a(0)$ and $a_{\lambda,\theta}(1) = \frac{p(1-q_2)(\theta q_1+(1-\theta))}{p(1-q_2)(\theta q_1+(1-\theta))+(1-p)q_2(1-\theta)}$ $= \frac{p(1-q_2)\left(\frac{\theta q_1}{\theta q_1+(1-\theta)}+1\right)}{p(1-q_2)\left(\frac{\theta q_1}{\theta q_1+(1-\theta)}+1\right)+(1-p)q_2} > a(1)$. The naive DM responds to $m = 1$ with $a = E(\omega|s_1 = 1, s_2 = 0)$ and to $m = 0$ with $a = E(\omega|s_1 = 0, s_2 = 0)$. Since type 0 expert has a continuous utility function and he strictly prefers $a(1)$ to $a(0)$, there exist ε_1, ε_2 s.t. for any $0 < \lambda < \varepsilon_1$ and $0 < \theta < \varepsilon_2$, type 0 expert prefers inducing $a_{\lambda,\theta}(1)$ with probability $(1-\lambda)$ and $a = E(\omega|s_1 = 1, s_2 = 0)$ with probability λ to inducing $a(0)$ with probability 1. Therefore, for $0 < \lambda < \varepsilon_1$ and $0 < \theta < \varepsilon_2$, $(\sigma_{\lambda,\theta}(\cdot), a_{\lambda,\theta}(\cdot))$ constitutes an equilibrium. Hence, $\hat{E}$ is robust to the perturbation.

To see the "only if" part, note that in a monotonic equilibrium in a perturbed game $\Gamma_{\lambda,\theta}$, we have $a_{\lambda,\theta}(1) > a_{\lambda,\theta}(0)$ and therefore $\sigma_{\lambda,\theta}(1,1) = 1$.

Hence, a separating equilibrium or a semi-separating equilibrium in which $\sigma(1,1) \neq$

1 is not robust to the perturbation. Similarly, a babbling equilibrium in which both types of expert send $m = 0$ with probability 1 is not robust to the perturbation. Neither is a babbling equilibrium where both types of expert send both messages with the same positive probabilities.

Now consider a babbling equilibrium in which both types of expert send $m = 1$ with probability 1 but type 0 expert weakly prefers being identified as type 0 to babbling. This equilibrium is not robust to the perturbation because in $\Gamma_{\lambda,\theta}$, $a_{\lambda,\theta}(0) = E(\omega|s_1 = 0, s_2 = 0)$ and $a_{\lambda,\theta}(1) > E(\omega|s_2 = 0)$ and type 0 expert always has an incentive to deviate and send $m = 0$ instead of $m = 1$. $\square$

Bibliography

Athey, S. (2001): "Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of Incomplete Information." *Econometrica*, Vol. 69, No. 4, 861-889.

Aumann, R. and S. Hart (2003): "Long Cheap Talk." *Econometrica*, Vol. 71, No. 6, 1619-1660.

Austen-Smith, D. (1993): "Interested Experts and Policy Advice: Multiple Referrals under Open Rule." *Games and Economic Behavior*, 5, 3-43.

Benabou, R. and G. Laroque (1992): " Using Priviledged Information to Manipulate Markets: Insiders, Gurus and Credibility." *Quarterly Journal of Economics*, Vol. 107, 921-958.

Crawford, V. (2003): "Lying for Strategic Advantage: Rational and Bounded Rational Misrepresentation of Intentions." *American Economic Review, 93,133-149.*

Crawford, V. and J. Sobel (1982): "Strategic Information Transmission." *Econometrica,* Vol. 50, No. 6, 1431-1451.

Farrell, J (1993): "Meaning and Credibility in Cheap-talk Games." *Games and Economic Behavior*, 5, 514-531.

Forsythe, R., R. Lundholm and T. Rietz (1999): "Cheap Talk, Fraud, and Adverse Selection in Financial Markets: Some Experimental Evidence." *Review of Financial Studies*, Vol. 12, No. 3, 481-518.

Harris, M. and A. Raviv (2004): "Allocation of Decision-Making Authority." mimeo.

Jackson, M., L. Simon, J. Swinkels and W. Zame (2002): "Communication and Equilibrium in Discontinuous Games of Incomplete Information." *Econometrica*, Vol. 70, No. 5, 1711-1740.

Kartik, N. (2003): "Information Transmission with Cheap and Almost-Cheap Talk." mimeo.

Kreps, D and R. Wilson (1982): "Sequential Equilibria." *Econometrica*, Vol. 50, No. 4, 863-894.

Mailath, G. (1987): "Incentive Compatibility in Signaling Games with a Continuum of Types." *Econometrica,* Vol. 55, No. 6, 1349-1365.

Mailath, G. (1992): "Signaling Games," in J. Creedy, J. Borland and J. Eichberger (eds.), *Recent Developments in Game Theory*, Edward Elgar, 65-93.

Manelli, A. (1996): "Cheap Talk and Sequential Equilibria in Signaling Games." *Econometrica*, Vol. 64, No. 4, 917-942.

Matthews, S., M. Okuno-Fujiwara and A. Postlewaite (1991): "Refining Cheap-Talk Equilibria." *Journal of Economic Theory*, 55, 247-273.

Morgan, J. and P. Stocken (2003): "An Analysis of Stock Recommendation." *Rand Journal of Economics*, Spring 2003, 34(1), 183-203.

Morris, S. (2001): "Political Correctness." *Journal of Political Economy*, Vol. 109, No. 2, 231-265.

Olszewski, W. (2004): "Informal Communication." *Journal of Economic Theory*, 117, 180-200.

Ottaviani, M. and F. Squintani (2002): "Non-fully Strategic Information Transmission.", mimeo.

Rabin, M. (1990): "Communication between Rational Agents." *Journal of Economic Theory,* 51, 144-170.

Seidmann, D. (1990): "Effective Cheap Talk with Conflicting Interests." *Journal of Economic Theory*, 50, 445-458.

Sobel, J. (1985): "A Theory of Credibility." *Review of Economic Studies*, 52, 557-573.