Confidence and Competence in Expertise*

Kohei Kawamura†
Nuffield College, University of Oxford

4 May 2007

Abstract

This paper studies communication between an uninformed principal and an expert who have asymmetric beliefs ("confidence") on the expert's ability ("competence") to observe the true state of nature. Based on the standard cheap talk model of Crawford and Sobel (1982) we show that overconfidence on the expert's side may enhance information transmission when he is severely biased, while this is not the case for underconfidence. Moreover, given the level of "competence" a slight degree of overconfidence always improves communication when the expert is biased. Both overconfidence and underconfidence hurt communication in the absence of such bias. Overconfidence gives rise to incentive to "exaggerate" the message away from the prior expectation, while underconfidence leads to incentive to "moderate" the message towards the prior.

1 Introduction

Many economic, political, and personal decisions are made on the basis of advice from those who have more information. However such expert advice is often subject to the expert's incentive to mislead the decision maker and may not reflect all the information the expert actually has. He may wish to induce the decision in the expert's favour rather than the decision maker's.

There are different types of informational distortion in expert advice. First, an expert may wish a certain decision to be taken regardless the information he has. For example, a policy consultant might have his own political position and his advice may be biased in a particular direction. Another potential source of informational distortion often discussed is

*Preliminary draft.

†kohei.kawamura@nuffield.ox.ac.uk

overconfidence of experts, whereby they overestimate the quality of information they have. Interestingly experts are seldom accused of being underconfident, although experimental studies have found that people tend to be underconfident under certain circumstances.

This paper studies how an expert's over- and underconfidence and his intrinsic bias in a particular direction affect the nature and efficiency of information transmission, by incorporating asymmetric beliefs (confidence) on an expert's competence into the standard cheap talk model of Crawford and Sobel (1982). In the absence of such intrinsic bias overconfidence leads to the expert's incentive to report "exaggerated" messages while underconfidence gives rise to incentive to report "moderated" messages. We also show that, while both overconfidence and underconfidence reduce the quality of communication when the expert is not intrinsically biased, slight overconfidence on the part of the expert enhances information transmission whenever he is moderately biased. Moreover overconfidence may extend the possibility of communication when he is severely biased. In contrast, underconfidence reduces this possibility.

Overconfidence has been attracting much attention from psychologists and economists. The literature on judgement under uncertainty has found that people tend to be overconfident about the information they have, in that their subjective probability distributions on relevant events are too tight (Kahneman et al., 1982). Overconfidence has been found in various professions such as lawyers (Wagenaar and Karen, 1982), policy experts (Tetlock, 1999), and security analysts (Chen and Jiang, 2006). The implications of overconfidence for economic choices and especially for financial markets have been studied recently by numerous researchers (e.g. Kyle and Wang, 1997; Gervais and Odean 1998; Daniel et al., 1998; Scheinkman and Xiong, 2003).

Although underconfidence is much less pronounced in the literature, this does not mean underconfidence does not exist in reality. For example, Hoelzl and Rustichini (2005) have found that people tend to be underconfident particularly with unfamiliar tasks. Underconfidence can potentially be an important issue especially in medical contexts. Friedman et al. (2004) have found that physicians are more likely to be underconfident about their diagnoses than to be overconfident. However, to our knowledge there has been no formal analysis of underconfidence in cheap talk games and little is known about the characteristics of communication with an underconfident expert.

In order to model the level of confidence, we allow for asymmetric beliefs on the probability that the expert observes the true state of nature. In other words, the decision maker and the expert may not share the same "confidence" on the expert's "competence" though they are fully aware of the difference in beliefs. Since we do not have to specify which player a priori has the correct belief, our framework can be applied to communication with "over-reliance" or "under-reliance" on expert advice, for which an expert is often supposed to hold the correct belief and his audience does not, as well as overconfidence and underconfidence of experts, for which an audience is often presumed to hold the correct belief.

Recent papers that involve asymmetric beliefs (non-common priors) include Admati and Pfleiderer (2004), Fang and Moscarini (2005), Van den Steen (2005), and Grubb (2006) among others. Fang and Moscarini (2005) consider the effect of workers' overconfidence in their skills on wage policies, and Van den Steen (2005) examines worker incentives when a worker may disagree with the manager regarding the best course of action. Grubb (2006) develops a model of screening with consumers who overestimate the precision of their demand forecasts.

The closest to our paper is Admati and Pfleiderer (2004) who study an information transmission game where the expert can be overconfident in his ability to observe the true state. They argue that in communication with an overconfident expert extreme messages or ratings are less credible (less precise) because if the expert reports the truth honestly, from his viewpoint the decision maker's reaction is too weak and the expert has incentive to "exaggerate" his report. Admati and Pfleiderer (2004) focus on overconfidence and do not consider underconfidence. Also they assume that the expert is otherwise unbiased: if the expert and the decision maker agree on the expert's ability the parties' interests are perfectly aligned. Thus they are unable to address the question how the intrinsic bias of an expert, which as been a focus of attention in the information transmission literature (and an important concern in practice), interacts with his confidence in his competence. In the present paper we analyze underconfidence as well as overconfidence in a systematic way. Moreover we explicitly illustrates the interaction between confidence and the intrinsic bias. We are able to offer an analysis of two important aspects, confidence and intrinsic bias, of communication with experts. Also, while Admati and Pfleiderer (2004) restrict the message space in such a way that the expert chooses one of an exogenously given finite set of messages, we do not impose such a restriction on the message space. In our model the structure of informative communication arises endogenously. Hence we are able to study more naturally how much information an expert can possibly communicate credibly in the presence of over- or underconfidence and intrinsic bias.

Although our model is a simple extension of the canonical model of Crawford and Sobel (1982) we cannot simply adopt their equilibrium characterization because the expert's incentive to misreport may not point in the same direction. Recently Gordon (2006) has provided the general characterization of a class of cheap talk equilibria where the expert's bias can depend on his information. Since he is mainly concerned with equilibrium characterization itself, his model says little about how various features in commutation, such as confidence, competence and intrinsic bias, relate to each other. However, we can use his general results in characterizing informative equilibria of our model. Our focus in the present paper is on how parameters regarding confidence and intrinsic bias alter the nature of communication.

We observe that overconfidence and underconfidence lead to different structures of informative equilibria. The informative equilibria under overconfidence is closely related to communication with multiple agents who observe independent signals, and communication with randomization. On the other hand, informative equilibria in communication with an underconfident expert shares some important characteristics with "reputational cheap talk".

In communication with multiple agents (Alonso, Dessein and Matouschek, 2006; Kawamura, 2007) and in communication with randomization (Krishna and Morgan, 2004; Blume, Board and Kawamura, 2007) the decision maker puts less weight on a message than in one-to-one communication without noise (i.e. the standard CS model), because the decision maker takes into account message from other agents (in the case of multiple agents) or because the received message may not be informative about the state (in the case of noisy communication). In the absence of intrinsic bias, weak response to a message leads to the expert's incentive to "exaggerate" his message relative to the prior expected state. This type of informational distortion appears in communication with an overconfident expert too, as we have already introduced above.[1]

In communication with an underconfident expert, the decision maker puts more weight on the message in updating her belief than the expert wishes her to. This (again in the absence of intrinsic bias) gives rises to the expert's incentive to "moderate" his message towards the prior expectation. The incentive to "moderate" is also observed in models of cheap talk with reputational concerns, where the expert attempts to look more able through messages when his ability (i.e. the precision of the signal he has observed) is unknown to the decision maker but known to the expert himself (Ottaviani and Sørensen, 2006a,b).

The rest of this paper is organized as follows. The next section describes the model, and Section 3 derives the equilibrium partition. Welfare properties are discussed in Section

[1] By incentive to "exaggerate", we mean an expert's incentive to misreport in such a way that, if his message refers directly to the value of a signal and is believed by the decision maker, the expert "overstates" ("understates") the signal when it is high (low).

Similarly, incentive to "moderate" is an expert's incentive to misreport in such a way that, if his message refers directly to the value of a signal and is believed by the decision maker, the expert "understates" ("overstates") the signal when it is high (low).

In cheap talk games messages used are completely arbitrary and do not have to be taken literally. What matters for the equilibrium outcome is the correspondence between the signal an expert has observed and the decision maker's action, so what word (or language) is used to induce a particular action is irrelevant.

4 and Section 5 concludes.

2 Model

Let us introduce the "uniform-quadratic" model of Crawford and Sobel (1982, henceforth CS). There are an uninformed decision maker and an expert who observes a private signal on the state of nature. Both parties' payoffs depend on the decision maker's action $y \in \mathbb{R}$ and the uniformly distributed state $\theta \in [0,1]$. The expert's utility function is $U^{EX} = -(y-\theta-b)^2$ and that of the decision maker is $U^{DM} = -(y-\theta)^2$, where b represents the expert's bias. We assume $b \geq 0$ without loss of generality.

Before the decision maker chooses her action, the expert observes a private signal $\sigma \in [0,1]$ on the state θ. After observing the signal the expert sends a costless message $m \in M$ to the decision maker, where M is a rich enough set to refer to every possible signal. Following convention we often refer to σ as the expert's "type".

CS assume that the expert observes the true state with probability 1, which is common knowledge. Now we introduce asymmetric beliefs ("confidence") on the quality of the signal (the expert's "competence") as follows: the decision maker believes that $\sigma = \theta$ with probability p, and σ has the identical (i.e. uniform on $[0,1]$) but independent distribution with $1-p$. On the other hand the expert believes that $\sigma = \theta$ with probability c, and σ has the identical but independent distribution with $1-c$. In other words the parties may have different beliefs on the probability that the signal observed by the expert coincides with the true state. Moreover they fully recognize the difference in their beliefs. That is, c and p are common knowledge.

In terms of the level of confidence and intrinsic bias our model is more general than the formulation by Admati and Pfleiderer (2004), who also adopt a similar uniform-quadratic setting, in that we introduce the expert's intrinsic bias b, and we also allow for the possibility that $c < p$ as well as $c \geq p$. In our notation Crawford and Sobel (1982) assume $p = c = 1$ and Admati and Pfleiderer (2004) assume $b = 0$, $c = 1$ and $p \leq 1$.[2]

For expositional convenience we look at the level of the expert's confidence from the decision maker's perspective. If $c > p$ the expert is said to be *overconfident* since the expert believes that his signal is more accurate (i.e. he is more competent) than the decision maker believes. Likewise if $c < p$ he is *underconfident*. However, since we are interested in how the difference between c and p affects communication, we are not concerned with which party has the correct belief on the expert's ability. It could be that neither party does, but this is not important since we calculate the players' expected utilities according to their own

[2] Admati and Pfleiderer (2004) allow for general distributions of the state/signal and focus on studying how the distributions affect information transmission.

subjective beliefs. If we look at the model from the expert's viewpoint and assume that he holds the correct belief, $c > p$ and $c < p$ could be respectively called "under-reliance" and "over-reliance" on expert.

In later sections we conduct welfare analysis according to the expected utilities based on subjective beliefs. Before doing so we derive communication equilibrium and examine how the nature of information transmission (i.e. messages communicated in equilibrium) changes depending on the values of b, c, and p. In what follows we look for equilibrium strategies where each party maximizes expected utility according to his/her subjective beliefs on the expert's competence. Given the message the decision maker updates her (subjective) belief using Bayes' rule.

From the decision maker's viewpoint a message from the expert is informative about the state of nature with probability p. Otherwise the message is completely uninformative. Since neither the decision maker nor the expert knows whether the expert has observed the true state of nature, the decision maker's maximization problem after she has received the message is given by

$$\begin{aligned}\max_y E[-(y-\theta)^2 \mid m] &= p\underbrace{E[-(y-\sigma)^2 \mid m]}_{\theta=\sigma} + (1-p)\underbrace{\int_0^1 -(y-\theta)^2 d\theta}_{\theta\sim U[0,1]} \\ &= p\left[-(y-E[\sigma \mid m])^2 - var(\sigma \mid m)\right] + (1-p)\int_0^1 -(y-\theta)^2 d\theta,\end{aligned}$$

where $E[\sigma \mid m]$ and $var(\sigma \mid m)$ are the expectation and variance of the signal conditional on the message, respectively. Since the utility function is quadratic, what matters for the decision maker's choice is the posterior expectation of the signal (and the state). Hence the first order condition for her best response conditional on the message yields

$$y^*(m) = pE[\sigma \mid m] + (1-p)\frac{1}{2}. \tag{1}$$

$E[\sigma \mid m]$ is weighted at p because the decision maker believes that σ is independent of the true state with probability $1-p$.

Before deriving the expert's best response, let us consider his desired action. Given the signal, from the expert's viewpoint $\theta = \sigma$ with probability c and otherwise θ is uniform on $[0,1]$. Thus his desired action is given by maximizing expected utility

$$\max_y c\underbrace{(-(y-\sigma-b)^2)}_{\theta=\sigma} + (1-c)\underbrace{\int_0^1 -(y-\theta-b)^2 d\theta}_{\theta\sim U[0,1]}.$$

The first order condition for the expert gives his desired action given σ

$$y^{EX}(\sigma) = c\sigma + (1-c)\frac{1}{2} + b. \tag{2}$$

Hypothetically, suppose that the expert fully reveals his signal. Then, since $E[\sigma \mid m] = \sigma$, (1) implies that the decision maker's desired action is

$$y^{DM}(\sigma) = p\sigma + (1-p)\frac{1}{2}. \tag{3}$$

Let us define $\hat{\sigma}$ to be the signal such that the parties' desired actions coincide: $y^{EX}(\hat{\sigma}) = y^{DM}(\hat{\sigma})$. From (2) and (3) we obtain

$$\hat{\sigma} = \frac{1}{2} - \frac{b}{c-p}.$$

We may have $\hat{\sigma} \in [0,1]$ only when $c \neq p$ and b is not too large. In the CS model ($c = p$) both parties' desired actions never coincide for $b > 0$: we have $y^{EX}(\sigma) > y^{DM}(\sigma)$ for all σ.

An important feature when we have $\hat{\sigma} \in [0,1]$ is that the difference between the parties' desired actions may not be consistently negative or positive. As a result if the expert is overconfident ($c > p$) we have $y^{EX}(\sigma) > y^{DM}(\sigma)$ for $\sigma > \hat{\sigma}$ and $y^{EX}(\sigma) < y^{DM}(\sigma)$ for $\sigma < \hat{\sigma}$: if the decision maker takes a reports literally and naively believes it the expert has incentive to "overstate" his signal for $\sigma > \hat{\sigma}$ and incentive to "understate" it for $\sigma < \hat{\sigma}$.[3] In the previous section we have referred to them as incentive to "exaggerate" since when σ is higher (lower) than $\hat{\sigma}$ the expert wants to convince the decision maker that σ is even higher (lower).

In contrast, if the expert is underconfident ($c < p$) we have $y^{EX}(\sigma) < y^{DM}(\sigma)$ for $\sigma > \hat{\sigma}$ and $y^{EX}(\sigma) > y^{DM}(\sigma)$ for $\sigma < \hat{\sigma}$. Now the expert has incentive to "understate" his signal for $\sigma > \hat{\sigma}$ and incentive to "overstate" it for $\sigma < \hat{\sigma}$. We have referred to them together as incentive to "moderate" since when σ is higher (lower) than $\hat{\sigma}$ the expert wants to convince the decision maker that σ is lower (higher) and closer to $\hat{\sigma}$ than it actually is.

3 Communication Equilibrium

Let us consider equilibria of this game. We have seen that the possibility of $c \neq p$ substantially alters the structure of the expert's incentive to misreport. If $c = p$ and $b > 0$, we have $y^{EX}(\sigma) > y^{DM}(\sigma)$ for all σ. For this case CS have shown that under general assumptions on preferences and distribution every perfect Bayesian equilibrium of the cheap talk game is partitional, in that the type (signal) space is divided into a finite number of intervals and all types in an interval induce the same action. The equilibrium characterization is generalized further by Gordon (2006) to cases where, as in our model, the parties' desired actions may coincide for a certain type ($\hat{\sigma} \in [0,1]$ in our model). The decision maker's best

[3] As we have noted in an earlier footnote, in this game the messages are arbitrary and do not necessarily have to refer literally to the values of signals.

response (1) and the expert's desired action (2) imply that our model constitutes a case of the *admissible problem* formulated and analyzed by Gordon (2006), and we can apply his equilibrium characterization to our model. Therefore following CS and Gordon (2006) we are able to characterize equilibria using an "arbitrage" condition, which says that an equilibrium partition is determined in such a way that any boundary type is indifferent between the actions induced by the types in the interval on left hand side and those in the interval on the right hand side.

Let $\underline{a}$ and $\overline{a}$ be two points in $[0,1]$ such that $\underline{a} < \overline{a}$. Suppose that the expert observes $\sigma \in [\underline{a}, \overline{a})$. Define $\bar{y}(\underline{a}, \overline{a})$ to be the decision maker's best response given her belief $\sigma \in [\underline{a}, \overline{a})$. Since σ is uniformly distributed (1) implies

$$\begin{aligned}
&\bar{y}(\underline{a}, \overline{a}) \\
&= pE[\sigma \mid \sigma \in [\underline{a}, \overline{a})] + (1-p)\frac{1}{2} \\
&= p\frac{\underline{a} + \overline{a}}{2} + (1-p)\frac{1}{2}.
\end{aligned} \tag{4}$$

In an equilibrium partition with N intervals each cut-off point a_i must satisfy an "arbitrage" condition, which says that the expert with $\sigma = a_i$ must be indifferent between inducing $\bar{y}(a_{i-1}, a_i)$ and $\bar{y}(a_i, a_{i+1})$:

$$\begin{aligned}
& -c(\bar{y}(a_{i-1}, a_i) - a_i - b)^2 + (1-c)\int_0^1 -(\bar{y}(a_{i-1}, a_i) - \theta - b)^2 d\theta \\
= \; & -c(\bar{y}(a_i, a_{i+1}) - a_i - b)^2 + (1-c)\int_0^1 -(\bar{y}(a_i, a_{i+1}) - \theta - b)^2 d\theta.
\end{aligned} \tag{5}$$

By using (4), (5) can be written

$$pa_{i+1} - (4c - 2p)a_i + pa_{i-1} = 4b - 2(c-p). \tag{6}$$

This second-order difference equation describes the (unique) equilibrium partition for given c, p, and b, where

$$a_1 = 0, a_N = 1 \tag{7}$$

and $a_i < a_{i+1}$ for all $i \in \{0, 1, 2, ..., N-1\}$. In what follows we refer to an equilibrium with at least two non-degenerate intervals as an *informative equilibrium*, where at least two different actions are induced with positive probability.

We formally describe the set of all equilibria in the following proposition. Let the expert's strategy specify the probability of sending message $m \in M$ conditional on observing signal σ and we denote it by $q(m \mid \sigma)$. The decision maker's best response is given by $y^*(m)$ in (1).

Proposition 1 (CS and Gordon, 2006)

Suppose that $\hat{\sigma} \notin [0,1]$ or that $\hat{\sigma} \in [0,1]$ and $c \leq p$. Then there exists an integer $\bar{N}$ such that, for every integer N with $1 \leq N \leq \bar{N}$, there exists at least one equilibrium $(y^*(m), q(m \mid \sigma))$, where $q(m \mid \sigma)$ is uniform, supported on $[a_i, a_{i+1}]$ if $\sigma \in (a_i, a_{i+1})$; $y^*(m) = \bar{y}(a_i, a_{i+1})$ for all $m \in (a_i, a_{i+1})$; and (7) and (6) hold.

Suppose that $c > p$ and $\hat{\sigma} \in [0,1]$. Then for every positive integer N there exists at least one equilibrium $(y^*(m), q(m \mid \sigma))$, where $q(m \mid \sigma)$ is uniform, supported on $[a_i, a_{i+1}]$ if $\sigma \in (a_i, a_{i+1})$; $y^*(m) = \bar{y}(a_i, a_{i+1})$ for all $m \in (a_i, a_{i+1})$; and (7) and (6) hold.

Moreover, for given b, c, and p, any other equilibria have the relationships between σ and y that are the same as those in the class described above, for some value of N. Therefore, they are economically equivalent.

Proof. The first part follows from Theorem 1 of CS. The second part is implied by Theorem 2 and Theorem 4 of Gordon (2006). ■

Proposition 1 says that there may be an infinite number of intervals in the type space when the expert is overconfident ($c > p$). In the following we derive equilibrium partitions for various parameter values.

3.1 Common Confidence ($c = p$)

If both parties have the same level of confidence about the expert's competence (probability that the expert observes the true state) we have $c = p$. Substituting this into (6) the equilibrium partitions in this case are given by

$$a_{i+1} - 2a_i + a_{i-1} = \frac{4b}{p}. \tag{8}$$

Under common confidence the only difference from CS is that we may have $p < 1$ while they assume $p = 1$. If $b = 0$, we have $y^{DM} = y^{EX}$ for any σ and hence full revelation is possible in equilibrium because both parties' interests are perfectly aligned. Rewriting (8) we have $(a_{i+1} - a_i) = (a_i - a_{i-1}) + \frac{4b}{p}$. In other words, the equilibrium partitions are such that an interval becomes longer as σ becomes larger. A message is less informative about the signal (or state) when the expert observes higher σ. This reflects the assumption that the expert is positively biased ($b > 0$). As in other cheap talk models there are multiple equilibria, including the "babbling" equilibrium where $a_0 = 0$ and $a_1 = 1$ so that no information is transmitted. We will discuss the issue of multiple equilibria later.

Let us define $\bar{b}$ to be the level of intrinsic bias such that an informative equilibrium exists for any $b \in [0, \bar{b})$. This can be obtained by checking at what value of b the equilibrium with two non-degenerate intervals can be supported. Substituting $a_0 = 0$ and $a_2 = 1$ into

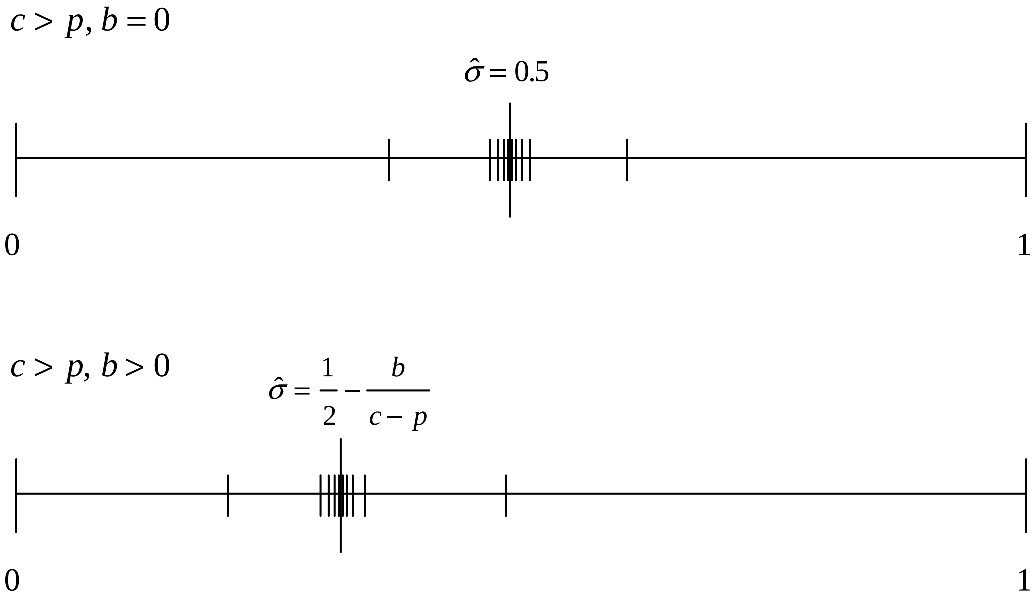

Figure 1: Infinite equilibrium partitions with an overconfident expert

(8) we obtain $a_1 = \frac{1}{2} - \frac{2b}{p}$. Hence $a_1 > 0$ implies that an informative equilibrium exists if $b < \bar{b}_{c=p} \equiv \frac{p}{4}$. That is, the less competent the expert is, the less bias is allowed for the message to be informative. When p is lower (4) implies that the message has less impact on the decision maker's action. This makes the relative size of intrinsic bias b with respect to the influence of a message (and signal) larger. As a result given the level of b, lower p leads to more limited information transmission.

3.2 Overconfident Expert ($c > p$)

When an expert is overconfident ($c > p$) and $\hat{\sigma} \in [0, 1]$, from (2) and (3) we have seen that the expert has incentive to "exaggerate" his message relative to $\hat{\sigma}$. In this case there exists an equilibrium with an infinite number of intervals, as depicted in Figure 1, where the horizontal line represents the type space and vertical lines are boundary types of intervals. We can see that the length of intervals becomes larger as it is away from $\hat{\sigma}$. It indicates that the overconfident expert may reveal more accurate information as his type is closer to $\hat{\sigma}$. This feature applies to equilibria with any finite number of intervals derived from (6) and (7). However, if $\hat{\sigma} \notin [0, 1]$ the structure of equilibria is such that any equilibrium partition is finite and the length of an interval becomes larger with σ, as we have seen for equilibrium partitions for $c = p$.

In the absence of intrinsic bias ($b = 0$) overconfidence reduces information transmission by making messages on extreme signals less precise. If $b = 0$ and $c = p$ perfect revelation is possible because both players' desired actions perfectly coincide for any σ. However if $c > p$ equilibrium communication is less than perfect as we can see in the first part of Figure 1.

In contrast, overconfidence may enhance information transmission in the presence of intrinsic bias. Let us consider the case where b is very large. Recall that when the decision maker and the expert share the same confidence ($c = p$) there exists an informative equilibrium for all $b \in \left[0, \frac{p}{4}\right)$. On the other hand in communication with an overconfident expert ($c > p$), substituting $a_0 = 0$ and $a_2 = 1$ into (8) the equilibrium with two intervals is given by

$$\left\{\left[0, \frac{1}{2} - \frac{2b}{2c-p}\right), \left[\frac{1}{2} - \frac{2b}{2c-p}, 1\right]\right\}.$$

Hence this equilibrium is informative for

$$a_1 = \frac{1}{2} - \frac{2b}{2c-p} > 0$$

or

$$b < \frac{2c-p}{4} = \bar{b}_{c>p}.$$

We have $\bar{b}_{c>p} > \bar{b}_{c=p}$. Therefore given the level of p, a more overconfident expert allows for more bias in informative communication. The intuition behind this result is as follows. Comparing (1) and (2) we can see that when $c > p$ the expert's desired action is more weighted towards the signal σ and away from the prior $1/2$. Suppose that the expert's type is low. He has two opposing incentives. One is to "overstate" his type due to bias b, and the other is to "understate" due to overconfidence. These two kinds of informational distortion partly offset each other, so that he may have more incentive to reveal (partially). Therefore overconfidence may mitigate the level of conflict and encourage information transmission. We will see later that slight overconfidence improves communication also when b is positive but small.

3.3 Underconfident Expert ($c < p$)

Let us consider the case where $c < p$, which implies the expert believes that his competence is lower than the decision maker believes. From (2) and (3) we have $y^{EX}(\sigma) > y^{DM}(\sigma)$ for $\sigma \in [0, \hat{\sigma})$ and $y^{EX}(\sigma) < y^{DM}(\sigma)$ for $\sigma \in (\hat{\sigma}, 1]$. Now the decision maker's reaction to a message is stronger than the expert wants it to be. For σ lower than $\hat{\sigma}$ the expert's desired action is higher than that of the decision maker and for σ higher than $\hat{\sigma}$ the expert's desired action is lower. In other words, contrary to the case of an overconfident expert, an underconfident expert has incentive to "moderate" his message.

Proposition 1 says that the number of intervals in equilibrium must be finite even if there exists $\hat{\sigma} \in [0, 1]$ for which both parties desired actions do coincide. Note that in

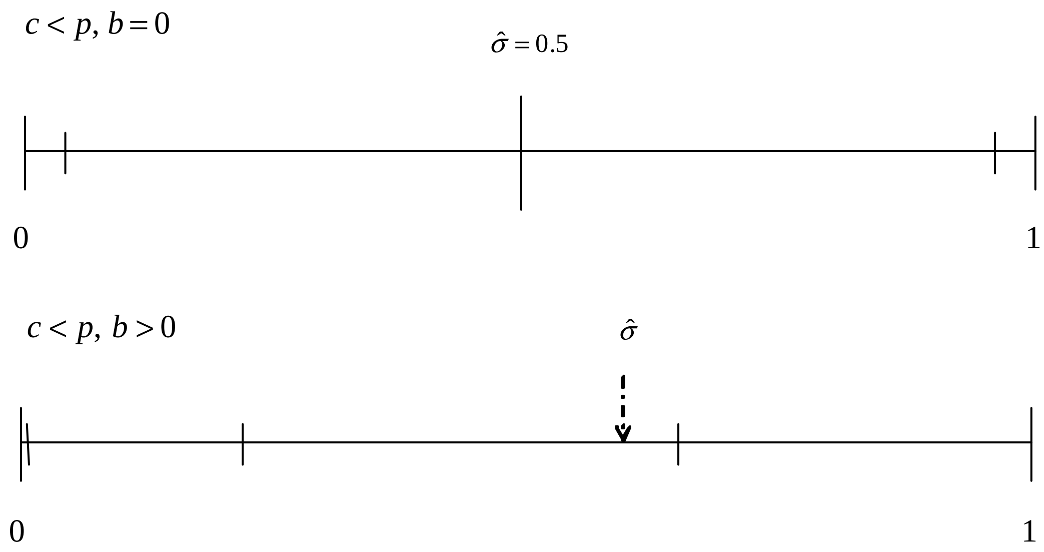

Figure 2: Informative equilibrium partitions with an underconfident expert

the overconfidence case there can be infinitely many intervals in the neighbourhood of $\hat{\sigma}$. The intuition behind finiteness is simple. In order for an equilibrium to have an infinite number of intervals it must be that the expert with $\sigma = \hat{\sigma}$ induces a unique action and the neighbouring types induce different actions as we see in Figure 1. However, when $c < p$ the expert whose type is very close to $\hat{\sigma}$ mimics $\sigma = \hat{\sigma}$ because of incentive to "moderate". Thus in equilibrium $\hat{\sigma}$ and its neighbouring types induce the same action.

Typical informative equilibrium partitions with an underconfident expert ($c < p$) are shown in Figure 2. In the absence of intrinsic bias ($b = 0$) the partition is symmetric with respect to the average (0.5) and intervals are shorter as they are closer to the extremes (0 and 1). When $b > 0$ the intervals are also shorter when they are closer to extremes although the partition is asymmetric. In both cases equilibrium partitions take into account the incentive for the expert to "moderate" his report towards $\hat{\sigma}$ because when choosing her action the decision maker puts more weight on the report than the expert wishes her to. As a result the message from the expert whose signal is closer to $\hat{\sigma}$ is less precise.

Let us consider to what level of bias an informative equilibrium exists. (6) gives the equilibrium partition with two intervals with $a_1 = \frac{1}{2} - \frac{2b}{2c-p}$. First suppose that $\frac{1}{2}p < c < p$, then $a_1 < \frac{1}{2}$ and the informative equilibrium partition is supported in equilibrium if $a_1 > 0$ or

$$b < \frac{2c - p}{4}.$$

Second, if $0 \leq c < \frac{1}{2}p$, then $a_1 > \frac{1}{2}$ and we must have $a_1 < 1$. This implies

$$b < \frac{p - 2c}{4}.$$

Third, if $c \to 1/2p$ then $a_1 \to \infty$ for any $b > 0$ so that the partition cannot be supported in equilibrium for $b > 0$. In summary, an informative equilibrium exists for

$$\begin{aligned} b &< \bar{b}_{c<p} \equiv \begin{cases} \frac{2c-p}{4} & \text{if } \frac{1}{2}p < c < p \\ \frac{p-2c}{4} & \text{if } \; 0 < c < \frac{1}{2}p. \end{cases} \\ b &= 0 \text{ if } c = \tfrac{1}{2}p \end{aligned} \tag{9}$$

Consider the common confidence case $c = p$ and $b > 0$. In this case lower types can credibly transmit more precise information. Underconfidence weakens this incentive because it makes desired actions of lower types higher through the incentive to "moderate" towards the prior 1/2 (i.e. incentive for lower types to "overstate"). What happens to higher types' incentive to (partially) separate? Underconfidence might lead to incentive to separate through the incentive to "moderate" towards the prior (i.e. incentive for higher types to "understate") but when the expert is moderately underconfident ($\frac{p}{2} < c < p$) the incentive to "understate" is not strong enough for higher types to overcome the positive bias.

On the other hand when the expert is extremely underconfident ($0 < c < \frac{1}{2}p$) higher types may partially separate because the incentive to "moderate" for those types may now offset the positive bias. Consequently, for this range of c the possibility of information transmission increases as the expert becomes more underconfident. It is easy to see that $\bar{b}_{c<p} < \bar{b}_{c=p}$. As we have seen above when $c = p$ there exists an informative equilibrium for $b \in [0, \frac{p}{4})$. Therefore, for given b and p, underconfidence reduces the prospect of informative communication. A higher degree of underconfidence may lead to more incentive to reveal for $c \in [0, \frac{p}{2}]$ (when the expert is extremely underconfident) but in the limit $c = 0$ the expert's ideal action is $1/2 + b$ regardless of the signal and for $b \geq \frac{p}{4}$ the only equilibrium is uninformative.

The effect of underconfidence on communication is in stark contrast to that of overconfidence. As we have seen in Figures 1 and 2, in communication without intrinsic bias an underconfident expert is able to transmit more precise information when his type is further away from the prior, while an overconfident expert's message carries more precise information when his type is closer to the prior. Also, while underconfidence reduces the possibility of information transmission for large bias, overconfidence has been shown to expand this possibility.

3.4 Asymmetric Beliefs and Binary Partition

Note that $a_1 = \frac{1}{2} - \frac{2b}{2c-p}$ in the equilibrium with two intervals implies that if $b = 0$ the equilibrium with partition $\{[0, 1/2), [1/2, 1]\}$ exists regardless of the levels of c and p. As we have seen so far, asymmetric beliefs on the expert's competence severely limit the informativeness of the expert's report, but unlike the intrinsic bias b, this asymmetry in beliefs per se never completely eliminates the possibility of information transmission. The intuition is simple: facing a choice of two messages the expert cannot "exaggerate" or "moderate" his message. In other words a binary choice of message is robust both to the incentive to "exaggerate" and the incentive to "moderate", which are the only sources of informational distortion when $b = 0$.

As we have seen earlier in communication with an overconfident expert and $b = 0$ there always exists an equilibrium with an infinite number of intervals. However, it is easy to check that as the degree of overconfidence $(c - p)$ becomes larger this (and in fact any informative) equilibrium converges to the equilibrium with the binary partition $\{[0, 1/2), [1/2, 1]\}$. Similarly when the degree of underconfidence is large enough $(c \leq \frac{p}{2})$ the only informative equilibrium is the one with the binary partition.

4 Multiple Equilibria and "Welfare"

As in most cheap talk models, there are multiple equilibria in our model. In Proposition 1 we have seen that if there is an equilibrium with $\bar{N}$ intervals there also exists an equilibrium with N intervals for any $1 \leq N \leq \bar{N}$. In CS and other related models the "informativeness" of an equilibrium can be measured by the number of intervals. In the CS model with quadratic utilities and uniform type distribution, both the decision maker's and the expert's expected utilities are higher as the number of intervals is larger. Thus the most efficient (and "informative") equilibrium under given parameter values is the one that has the largest number of intervals.

In our model the calculation of expected utility can be potentially problematic because we have not specified who holds the "correct" belief on the expert's ability (the probability that he observes the correct signal). Rather, neither party's belief (c nor p) may be true for the above equilibrium construction to be valid. In what follows we calculate each player's expected utility according to his/her own belief, without considering the expert's "true" competence.

Proposition 2 If $c \geq p$ then both the decision maker's and the expert's expected utilities are higher in an equilibrium with a larger number of intervals.

Proof. See Appendix ■

When the players share the same confidence ($c = p$) or the expert is overconfident ($c > p$) the informativeness can still be measured by the number of intervals. However, this convenient property does not hold in communication with an underconfident expert ($c < p$). Let us see examples where one party is worse off in an equilibrium with more intervals.

Suppose that $c = 0.55$, $p = 1$ and $b = 0$. Then we have $\bar{N} = 3$: the largest number of intervals supported in equilibrium is three, and the corresponding partition is given by $\{[0, 1/12), [1/12, 11/12), [11/12, 1]\}$ where $EU^{DM} = -0.0483$ and $EU^{EX} = -0.0444$. On the other hand in the equilibrium where the partition has two intervals $\{[0, 1/2), [1/2, 1]\}$, we obtain $EU^{DM} = -0.0208$ and $EU^{EX} = -0.0771$. Thus the decision maker is better off in the equilibrium with two intervals, while the underconfident expert is better off in the equilibrium with three intervals.

	$N = 2$	$N = 3$
EU^{DM}	-0.0208	-0.0483
EU^{EX}	-0.0771	-0.0444

Table 1: $c = 0.55, p = 1$

	$N = 1$	$N = 2$
EU^{DM}	-0.0833	-0.0208
EU^{EX}	-0.0833	-0.0958

Table 2: $c = 0.40, p = 1$

Note that the length of an interval reflects the precision of the information conveyed by messages. The partition $\{[0, 1/12), [1/12, 11/12), [11/12, 1]\}$ implies that a message is very imprecise (the decision maker's posterior is such that $\sigma \in [1/12, 11/12)$) with large probability (0.8333), while only with small probability (0.1777) it is relatively precise (i.e. the posterior is either $\sigma \in [0, 1/12)$ or $\sigma \in [11/12, 1]$). In contrast, when the equilibrium partition is $\{[0, 1/2), [1/2, 1]\}$ the message is always moderately precise (the posterior is either $\sigma \in [0, 1/2)$ or $\sigma \in [1/2, 1]$) compared with the other equilibrium. Overall the decision maker, who wants to minimize the expected variance of σ, prefers to be always moderately informed in this example. The underconfident expert prefers the equilibrium with $N = 3$ because due to underconfidence he prefers an action around the one based on the prior $y^{EX}(1/2) = 1/2$, and indeed the action is likely to be (with probability 0.8333) $y = 1/2$ in this equilibrium.

Suppose also that the expert is even more underconfident $c = 0.4$, $p = 1$ and $b = 0$. We then have $\bar{N} = 2$: one is uninformative and the other has two intervals $\{[0, 1/2), [1/2, 1]\}$. In the uninformative equilibrium $EU^{DM} = EU^{EX} = -0.0833$. In the equilibrium with two intervals we have $EU^{DM} = -0.0208$ and $EU^{EX} = -0.0958$. Therefore while the decision maker prefers the informative equilibrium with two intervals, the expert prefers the uninformative equilibrium. The decision maker prefers $N = 2$ to $N = 1$ because in the former equilibrium she can update her belief on the true state, which never hurts her.

On the other hand from the underconfident expert's viewpoint the decision maker's action varies "too much" according to his message since the decision maker puts more weight on the message than the expert wants her to. This reduces the expert's utility because the ex ante difference between the decision maker's action and the expert's ideal action is larger for $N = 1$ than $N = 2$.

These two examples suggest that in communication with an underconfident expert there is not a simple monotonic relationship between the number of intervals and welfare.

Proposition 3 The decision maker's expected utility is higher in an equilibrium with two or more intervals than in the uninformative equilibrium.

Proof. This follows directly from the fact that the decision maker chooses her action to maximize her expected utility according to her posterior belief. An equilibrium with any non-degenerate intervals enables the decision maker to update her belief on the signal and gives higher expected utility than the uninformative equilibrium where the decision depends only on the prior. ■

The decision maker can put an appropriate weight on the message from the expert according to p in order to maximize expected utility. Therefore when she updates her belief on σ she is better off than in the uninformative equilibrium. However, since his message has a stronger influence on the decision maker's action than optimal, the expert may prefer the uninformative equilibrium to an informative one especially when he is severely underconfident, as we have seen in Table 2.

4.1 Slight Overconfidence Always Helps

We have seen that overconfidence may extend the possibility of information transmission when b is very high, or $1/4 \leq b < 1/2$. In fact we are able to show that at least slight overconfidence is strictly beneficial also for $0 < b < 1/4$.

Proposition 4 Suppose that there is an informative equilibrium for given levels of $b > 0$, and $p = c$. Then for $c = p + \epsilon$ with some small $\epsilon > 0$, there exists an equilibrium where the decision maker's expected utility is higher than the case with $c = p$.

Proof. See Appendix ■

Given the same levels of competence p and intrinsic bias b, the decision maker is better off by communicating with a slightly overconfident expert than an expert with common confidence $c = p$. In order to obtain the intuition, consider the simple case where the equilibrium has two intervals $\{[0, a_1), [a_1, 1]\}$ such that $a_1 = \frac{1}{2} - \frac{2b}{2c-p}$. The most desirable binary partition for the decision maker is $a_1 = 1/2$ for which the expected ex post variance of the signal is lowest. Note that this is the binary equilibrium partition for $b = 0$. For any $b > 0$ we have $a_1 \neq 1/2$ but since a_1 is increasing in c, compared to the common confidence case $c = p$, as c becomes larger a_1 is closer to the "best" binary partition $a_1 = 1/2$. The decision maker clearly favours this change. In other words, overconfidence has the same effect as lower b. While excessive overconfidence reduces informativeness when b is small, slight overconfidence always helps as long as b is positive but not too large.

Clearly the same reasoning does not apply to communication with an underconfident expert. Compared to the common confidence case $c = p$ reducing c moves a_1 away from $1/2$, which increases the expected variance of the signal from the decision maker's viewpoint.

5 Conclusion

This paper offers a first systematic analysis of both over- and underconfidence in communication. We have studied communication between a principal and an expert who have asymmetric beliefs ("confidence") on the expert's ability to observe the state of nature ("competence"). In the absence of intrinsic bias, overconfidence in communication is characterized by the expert's incentive to "exaggerate", and underconfidence entails incentive to "moderate" his report. Consequently, signals closer to the prior expectation are more accurately reported in communication with an overconfident expert, whereas extreme singles are more accurately reported in communication with an underconfident expert. When the expert is intrinsically biased in a particular direction, we have shown that overconfidence on the expert's side may mitigate the intrinsic bias, while underconfidence exacerbates it. This might explain why overconfidence is prevalent in expertise and overconfident experts do survive despite potential incentive to "exaggerate" reports: since they may reveal more information, those seeking expert advice might in fact welcome a certain degree of overconfidence.

6 Appendix

6.1 Preliminaries to the Proof of Proposition 2

Before we prove the Proposition, we provide some useful lemmas and outline how we construct the main proof. Let us call a sequence $(a_0, a_1, ..., a_N)$ that satisfies the arbitrage condition (6) a "*solution*" to (6). The monotonicity condition (M) in CS requires that, for given γ and b, if we have two solutions a^+ and a^{++} with $a_0^+ = a_0^{++}$ and $a_1^+ > a_1^{++}$, then $a_i^+ > a_i^{++}$ for all $i = 2, 3, ...$ In other words, (M) says that starting from a_0, all solutions to (6) must move up or down together. Solving (6) explicitly for $c \geq p$ with $a_0 = 0$ and $a_N = 1$ we obtain

$$a_i = \hat{\sigma} + \frac{1 - \hat{\sigma}(1 - y^N)}{x^N - y^N} x^i + \frac{-1 + \hat{\sigma}(1 - x^N)}{x^N - y^N} y^i \tag{10}$$

where x and y are two distinct roots

$$x = \frac{2c - 2p + 2\sqrt{c(c-p)}}{p} \text{ and } y = \frac{2c - 2p - 2\sqrt{c(c-p)}}{p}.$$

Also, solving (6) in terms of a_1

$$a_i = \hat{\sigma} + \frac{a_1 - \hat{\sigma}(1 - y)}{x - y} x^i - \frac{a_1 - \hat{\sigma}(1 - x)}{x - y} y^i.$$

Since $da_i/da_1 > 0$ for all $i = 2, 3, ..., N$ we can see that the equilibrium partition of our model satisfies (M) for $c \geq p$.

In order to show that the expected utilities of both players are higher in an equilibrium with more intervals, CS deform the partition with N intervals to that with $N + 1$ intervals, continuously increasing the player's expected utility throughout the deformation. We follow this method, but we need to proceed by two step deformation, rather than one, because when $\hat{\sigma} \in (0, 1]$ the deformation takes place towards the opposite directions for the right-hand and left-hand sides of $\hat{\sigma}$. Intuitively, as the number of interval increases, each boundary type on the left hand side of $\hat{\sigma}$ move to the left (except for $a_0 = 0$) while each boundary type of the right hand side of $\hat{\sigma}$ move to the right (except for $a_N = 1$). We need to perform a different comparative statics for each case.

Let $a(N)$ be the equilibrium partition of size N. We show that $a(N)$ can be deformed to $a(N+1)$ by two steps, continuously increasing the players' expected utility in each step. Here we consider the case where $\hat{\sigma} \in (0, 1]$. We omit the case where $\hat{\sigma} \notin (0, 1]$ because the Proposition for this case can be proven similarly, by using the first step only.

Let the sub-partition of $a(N)$ equal or below $\hat{\sigma}$ be $\underline{a}(N) \equiv (a_0(N), a_2(N), ..., a_K(N))$ where $a_0(N) = 0$. Also, suppose that $a_K(N)$ is closer to $\hat{\sigma}$ than $a_{K+1}(N)$ is, in other words, $\hat{\sigma} - a_K(N) < a_{K+1}(N) - \hat{\sigma}$. In the following we proceed in two steps:

1. We fix $a_K(N)$ and make the sub-partition $(a_K(N), a_{K+1}(N), ..., a_N(N))$ deform continuously to $(a_K(N), a_{K+1}(N+1), a_{K+2}(N+1), ..., a_{N+1}(N+1))$, increasing the expected utility.

2. We make the sub-partition $(a_0(N), a_1(N)..., a_K(N))$ deform continuously to $(a_0(N+1), a_2(N+1), ..., a_K(N+1))$, increasing the expected utility.

- If $\hat{\sigma} - a_K(N) \geq a_{K+1}(N) - \hat{\sigma}$ then the first step deforms $(a_0(N), a_1(N), ..., a_K(N), a_{K+1}(N))$ to $(a_0(N+1), a_1(N+1), ..., a_{K+1}(N+1), a_{K+1}(N))$ while fixing $a_{K+1}(N)$, and the second step deforms $(a_{K+1}(N), a_{K+2}(N), ..., a_N(N))$ to $(a_{K+2}(N+1), a_{K+3}(N+1), ..., a_{N+1}(N+1))$. Except for this the same method and result as the case where $\hat{\sigma} - a_K(N) < a_{K+1}(N) - \hat{\sigma}$ apply.

Lemma 1 If $a(N)$ and $a(N+1)$ are two equilibrium partitions for the same values of b and γ, then $a_{i-1}(N) < a_i(N+1) < a_i(N)$.

Proof. See Lemma 3 (p.1446) in CS. The proof follows directly from (M). ■

The first step of deformation is carried out as follows. Let $(a_K^x, a_{K+1}^x, ..., a_i^x, ..., a_{N+1}^x)$ be the sub-partition that satisfies (6) for all $i = K+1, K+2, ..., N$ with $a_K^x = a_K(N)$, $a_N^x = x$ and $a_{N+1}^x = 1$. If $x = a_{N-1}(N)$ then $a_{K+1}^x = a_K^x = a_K(N)$. If $x = a_N(N+1)$ then we have $(a_K(N), a_{K+1}(N+1), ..., a_N(N+1))$, where (6) is satisfied for all $i = K+2, K+3, ..., N$. We are going show that, if $x \in [a_{N-1}(N), a_N(N+1)]$, which is again a non-degenerate interval by Lemma 1, then the sender's expected utility is strictly increasing in x.

In the second step, let $(a_0^z, a_1^z, ..., a_i^z, ..., a_K^z)$ be the sub-partition that satisfies (6) for $i = 1, 2, ..., K-1$, with $a_0^z = 0$ and $a_K^z = z$. If $z = a_K(N)$ then $a_i^z = a_i(N)$ for all $i = 0, 1, ..., K$. If $z = a_K(N+1)$ then $a_i^z = a_i(N+1)$ for all $i = 0, 1, ..., K$.. We will show that when $z \in [a_K(N+1), a_K(N)]$, which is again a non-degenerate interval by Lemma 1, the sender's expected utility is strictly decreasing in z.

Lemma 2 Suppose that $(a_0, a_1, ..., a_i, ..., a_N)$ is a solution to (6). Then for all $i = 1, 2, ..., N-1$ if $a_i > (<)\hat{\sigma}$ then $a_i - a_{i-1} < a_{i+1} - a_i$ $(a_i - a_{i-1} > a_{i+1} - a_i)$. If $a_i = \hat{\sigma}$ then $a_i - a_{i-1} = a_{i+1} - a_i$.

Proof. Rearranging (6) we have

$$(a_{i+1} - a_i) - (a_i - a_{i-1}) = \frac{4ca_i + 4b - 2c}{p} - 4a_i + 2. \tag{11}$$

The left hand side $(a_{i+1} - a_i) - (a_i - a_{i-1}) = 0$ if

$$\begin{aligned} \frac{4a_i + 4b - 2c}{p} - 4a_i + 2 &= 0 \Rightarrow \\ 4a_i(c-p) &= -4b + 2c - 2p \Rightarrow \\ a_i &= \frac{1}{2} - \frac{b}{c-p} \equiv \hat{\sigma}. \end{aligned}$$

Since the right hand side of (11) is increasing in a_i for $c > p$. Thus if $a_i > \hat{\sigma}$ then $(a_{i+1} - a_i) - (a_i - a_{i-1}) > 0$, and if $a_i < \hat{\sigma}$ then $(a_{i+1} - a_i) - (a_i - a_{i-1}) < 0$. ■

The above lemma says that an interval $[a_i, a_{i+1})$ is longer (shorter) than the previous interval $[a_{i-1}, a_i)$ when $a_i > (<)\hat{\sigma}$. The intuition is captured in Figure 1. The following Lemma is similar but cannot be implied by Lemma 2. Since by definition a_K^x and a_{K+1}^z are fixed throughout the respective deformation, (6) is not satisfied at $a_i = a_{K+1}^x$ for $x \in (a_{N-1}(N), a_N(N+1))$ or $a_i = a_K^z$ for $z \in (a_K(N+1), a_K(N))$.

Lemma 3 $a_{K+1}^x - a_K^x < a_{K+2}^x - a_{K+1}^x$ and $a_K^z - a_{K-1}^z > a_{K+1}^z - a_K^z$.

Proof. From Lemma 2 we have $a_{K+1}^x - \tilde{a}_K < a_{K+2}^x - a_{K+1}^x$ where $\tilde{a}_K$ is defined such that $\{a_{i-1} = \tilde{a}_K,\ a_i = a_{K+1}^x, a_{i+1} = a_{K+2}^x\}$ satisfies (6). Since $a_K(N+1) < \tilde{a}_K < a_K(N) = a_K^x$ from Lemma 1, we have $a_{K+1}^x - a_K^x < a_{K+2}^x - a_{K+1}^x$. This proves the first part of the Lemma.

Similarly we have $a_K^z - a_{K-1}^z \geq \check{a}_{K+1} - a_K^z$ where $\check{a}_{K+1}$ is defined such that $\{a_{i-1} = a_{K-1}^z, a_i = a_K^z, a_{i+1} = \check{a}_{K+1}\}$ satisfies (6). Lemma 1 implies $a_{K+1}^z = a_{K+1}(N+1) < \check{a}_{K+1} < a_{K+1}(N)$. Hence we have $a_K^z - a_{K-1}^z > a_{K+1}^z - a_K^z$. ■

6.2 Proof of Proposition 2

The expert's expected utility for the first part of deformation is given by

$$\begin{aligned} EU^{EX} \equiv & -c\left[\sum_{i=1}^{K}\int_{a_{i-1}}^{a_i}\left(p\frac{a_{i-1}+a_i}{2}+\frac{1-p}{2}-b-\theta\right)^2 d\theta \right. \\ & \left. +\sum_{i=K+1}^{N+1}\int_{a_{i-1}^x}^{a_i^x}\left(p\frac{a_{i-1}^x+a_i^x}{2}+\frac{1-p}{2}-b-\theta\right)^2 d\theta\right] \\ & -(1-c)\left[\sum_{i=1}^{K}(a_i-a_{i-1})\int_0^1\left(p\frac{a_{i-1}+a_i}{2}+\frac{1-p}{2}-b-\theta\right)^2 d\theta \right. \\ & \left. +\sum_{i=K+1}^{N+1}(a_i^x-a_{i-1}^x)\int_0^1\left(p\frac{a_{i-1}^x+a_i^x}{2}+\frac{1-p}{2}-b-\theta\right)^2 d\theta\right]. \end{aligned}$$

It follows that

$$\frac{dEU^{EX}}{dx} \equiv \sum_{i=K+1}^{N+1} \frac{da_i^x}{dx} \times$$
$$\Bigg\{ -c \Bigg[\left(p\frac{a_{i-1}^x + a_i^x}{2} + \frac{1-p}{2} - b - a_i^x \right)^2 - \left(p\frac{a_i^x + a_{i+1}^x}{2} + \frac{1-p}{2} - b - a_i^x \right)^2$$
$$+ p \int_{a_{i-1}^x}^{a_i^x} \left(p\frac{a_{i-1}^x + a_i^x}{2} + \frac{1-p}{2} - b - \theta \right) d\theta + p \int_{a_i^x}^{a_{i+1}^x} \left(p\frac{a_i^x + a_{i+1}^x}{2} + \frac{1-p}{2} - b - \theta \right) d\theta \Bigg]$$
$$- (1-c) \Bigg[\int_0^1 \left(p\frac{a_{i-1}^x + a_i^x}{2} + \frac{1-p}{2} - b - \theta \right)^2 d\theta - \int_0^1 \left(p\frac{a_i^x + a_{i+1}^x}{2} + \frac{1-p}{2} - b - \theta \right)^2 d\theta$$
$$+ p(a_i^x - a_{i-1}^x) \int_0^1 \left(p\frac{a_{i-1}^x + a_i^x}{2} + \frac{1-p}{2} - b - \theta \right) d\theta$$
$$+ p(a_{i+1}^x - a_i^x) \int_0^1 \left(p\frac{a_i^x + a_{i+1}^x}{2} + \frac{1-p}{2} - b - \theta \right) d\theta \Bigg] \Bigg\}$$
$$= \sum_{i=K+1}^{N-1} \frac{da_i^x}{dx} \frac{(a_{i+1}^x - a_i^x)(2c-p)p}{4} (a_{i-1}^x - 2a_i^x + a_{i+1}^x) > 0 \qquad (12)$$

The inequality follows since $\frac{da_i^x}{dx} > 0$ and from Lemmas 2 and 3 we have $a_i - a_{i-1} < a_{i+1} - a_i \Rightarrow a_{i-1} - 2a_i - a_{i+1} > 0$ for $i = 1, 2, ..., N-1$.

Let us look at the second part of deformation.

$$\frac{dEU^{EX}}{dz} \equiv \sum_{i=1}^{K} \frac{da_i^z}{dz} \times$$
$$\Bigg\{ -c \Bigg[\left(p\frac{a_{i-1}^z + a_i^z}{2} + \frac{1-p}{2} - b - a_i^z \right)^2 - \left(p\frac{a_i^z + a_{i+1}^z}{2} + \frac{1-p}{2} - b - a_i^z \right)^2$$
$$+ p \int_{a_{i-1}^z}^{a_i^z} \left(p\frac{a_{i-1}^z + a_i^z}{2} + \frac{1-p}{2} - b - \theta \right) d\theta + p \int_{a_i^z}^{a_{i+1}^z} \left(p\frac{a_i^z + a_{i+1}^z}{2} + \frac{1-p}{2} - b - \theta \right) d\theta \Bigg]$$
$$- (1-c) \Bigg[\int_0^1 \left(p\frac{a_{i-1}^z + a_i^z}{2} + \frac{1-p}{2} - b - \theta \right)^2 d\theta - \int_0^1 \left(p\frac{a_i^z + a_{i+1}^z}{2} + \frac{1-p}{2} - b - \theta \right)^2 d\theta$$
$$+ p(a_i^z - a_{i-1}^z) \int_0^1 \left(p\frac{a_{i-1}^z + a_i^z}{2} + \frac{1-p}{2} - b - \theta \right) d\theta$$
$$+ p(a_{i+1}^z - a_i^z) \int_0^1 \left(p\frac{a_i^z + a_{i+1}^z}{2} + \frac{1-p}{2} - b - \theta \right) d\theta \Bigg] \Bigg\}$$
$$= \sum_{i=1}^{K} \frac{da_i^z}{dz} \frac{(a_{i+1}^z - a_i^z)(2c-p)p}{4} (a_{i-1}^z - 2a_i^z + a_{i+1}^z) < 0 \qquad (13)$$

The inequality follows since we have $\frac{da_i^z}{dz} > 0$ and Lemmas 2 and 3 imply $a_i - a_{i-1} >$

$a_{i+1} - a_i \Rightarrow a^z_{i-1} - 2a^z_i + a^z_{i+1} < 0$. Therefore, EU^{EX} is increasing throughout the second part of deformation for which z decreases from $a_K(N)$ to $a_K(N+1)$.

Since we have completed the deformation from $a(N)$ to $a(N+1)$ by two steps while increasing the expected utility, we conclude that the sender's expected utility is higher in an equilibrium with more intervals.

6.2.1 Decision Maker

Following the above two-step deformation, the receiver's expected utility for the first part of deformation is given by

$$\begin{aligned} EU^{DM} \equiv & -p\left[\sum_{i=1}^{K}\int_{a_{i-1}}^{a_i}\left(p\frac{a_{i-1}+a_i}{2}+\frac{1-p}{2}-\theta\right)^2 d\theta \right. \\ & \left. +\sum_{i=K+1}^{N+1}\int_{a^x_{i-1}}^{a^x_i}\left(p\frac{a^x_{i-1}+a^x_i}{2}+\frac{1-p}{2}-\theta\right)^2 d\theta\right] \\ & -(1-p)\left[\sum_{i=1}^{K}(a_i-a_{i-1})\int_0^1\left(p\frac{a_{i-1}+a_i}{2}+\frac{1-p}{2}-\theta\right)^2 d\theta \right. \\ & \left. +\sum_{i=K+1}^{N+1}(a^x_i-a^x_{i-1})\int_0^1\left(p\frac{a^x_{i-1}+a^x_i}{2}+\frac{1-p}{2}-\theta\right)^2 d\theta\right]. \end{aligned}$$

Note that the expected utility is identical to that of the expert, except that $b = 0$ and $c = p$. Therefore, in order to show that the decision maker's expected utility is higher in an equilibrium with more intervals, we can directly apply the the argument we have used for the expert's expected utility.

6.3 Proof of Proposition 4

Differentiating (10) with respect to $\hat{\sigma}$ we have

$$\lim_{c\to p}\frac{da_i}{dc} = \frac{\overbrace{i(i-N)}^{-}[-2bN(1-i^2+iN+N^2)+(2i-N)p]}{\underbrace{3Np^2}_{+}} > 0 \text{ for } i = 1,2,...,N-1. \quad (14)$$

For the inequality to follow, the term in the square brackets must be negative

$$\underbrace{-2bN(1-i^2+iN+N^2)}_{-} + (2i-N)p < 0.$$

If this inequality holds for $i = N-1$ (so that the second term on the left hand side is the largest), it is also satisfied for all $i = 1,2,...,N-1$. Thus substituting $i = N-1$ we obtain

$$b > \frac{(N-2)p}{2N^2(N+1)}. \quad (15)$$

The assumption that the equilibrium with N intervals is the most informative equilibrium implies

$$\underbrace{\frac{p}{2N(N-1)}}_{\text{largest bias for eqm with } N \text{ intervals.}} > b > \underbrace{\frac{p}{2N(N+1)}}_{\text{largest bias for eqm with } N+1 \text{ intervals}} \tag{16}$$

and it is easy to check that

$$\frac{p}{2N(N+1)} > \frac{(N-2)p}{2N^2(N+1)}.$$

This implies that in the most informative equilibrium (15) and hence (14) hold.

The decision maker's expected utility is given by

$$\begin{aligned} EU^{DM} &= -p\sum_{i=1}^{N}\int_{a_{i-1}}^{a_i}\left(p\frac{a_{i-1}+a_i}{2}+\frac{1-p}{2}-\theta\right)^2 d\theta \\ &\quad -(1-p)\sum_{i=1}^{N}(a_i-a_{i-1})\int_0^1\left(p\frac{a_{i-1}+a_i}{2}+\frac{1-p}{2}-\theta\right)^2 d\theta \end{aligned}$$

Fix the level of p and let $a(N,\bar{c})$ be the partition with N non-degenerate intervals and $p = c = \bar{c}$. Since $b > 0$ any informative equilibrium has a finite number of intervals (Proposition 1). By continuity we can construct the equilibrium partition with N intervals with $c = \bar{c}+\epsilon$ for small enough ϵ, which we denote by $a(N,\bar{c}+\epsilon)$.

Now we can deform $a(N,\bar{c})$ into $a(N,\bar{c}+\epsilon)$, increasing the receiver's expected utility throughout the deformation, as we have done in 12.

References

[1] Admati A. and P. Pfleiderer (2004), "Broadcasting Opinions with an Overconfident Sender," *International Economic Review*, 45, 467-497.

[2] Blume, A., O. Board and K. Kawamura (2007), "Noisy Talk," working paper.

[3] Crawford V. and J. Sobel (1982), "Strategic Information Transmission," *Econometrica*, 50, 1431-1451.

[4] Chen, Q. and W. Jiang (2006),."Analysts' Weighting of Private and Public Information," *Journal of Financial Studies*, 19, 320-355.

[5] Daniel, K., D. Hirshleifer, A. Subrahmanyam (1998), "Investor Psychology and Security Market Under- and Overreactions." *Journal of Finance*, 53, 1839-1885.

[6] Fang, H. and G. Moscarini (2005), “Morale Hazard,” *Journal of Monetary Economics*, 52, 749–777.

[7] Friedman, C., C. Gatti, T. Franz, G. Murphy, F. Wolf, S. Heckerling, P. Fine, T. Miller, A. Elstein (2004), "Do Physicians Know When Their Diagnoses Are Correct?: Implications for Decision Support and Error Reduction," *Journal of General Internal Medicine*, 20, 334-339.

[8] Gervais, S., T. Odean (1998), "Learning to be Overconfident," *Review of Financial Studies* 14, 1-27.

[9] Gordon, S. (2006). "Informative Cheap Talk Equilibria as Fixed Points," working paper.

[10] Griffin, D., A. Tversky (1992), "The Weighing of Evidence and the Determinants of Confidence," *Cognitive Psychology*, 24, 411-435.

[11] Grubb, M. (2006), "Selling to Overconfident Consumers," working paper.

[12] Hoelzl, E. and A. Rustichini (2005), "Overconfident: Do you put your money on it?" *Economic Journal*, 115, 305-318.

[13] Kahneman, D., P. Slovic, and A. Tversky (eds.) (1982), *Judgment Under Uncertainty: Heuristics and Biases* Cambridge University Press, Cambridge.

[14] Kawamura, K. (2007), "Constrained Communication with Multiple Agents: Anonymity, Equal Treatment and Public Good Provision," working paper.

[15] Kyle, A., A. Wang (1997), "Speculation Duopoly with Agreement to Disagree: Can Overconfidence Survive the Market Test?" *Journal of Finance*, 52, 2073-2090.

[16] Melumad, N. D. and T. Shibano (1991), "Communication in Settings with No Transfers," *Rand Journal of Economics*, 22, 173-198.

[17] Oskamp, S. (1965), "Overconfidence in Case Study Judgments," *Journal of Consulting Psychology*, 29, 261-265.

[18] Ottaviani, M. and P. N. Sørensen (2006a), "Reputational Cheap Talk," *RAND Journal of Economics,* 37, 155-175.

[19] ———, ——— (2006b), "The Strategy of Professional Forecasting," *Journal of Financial Economics*, 81, 441-466.

[20] Scheinkman, J., W. Xiong (2003), "Overconfidence and speculative bubbles," *Journal of Political Economy*, 111, 1183-1219.

[21] Tetlock, P. (1999), "Theory-driven reasoning about plausible pasts and probable futures in world politics: are we prisoners of our preconceptions?," *American Journal of Political Science*, 43, 335-366.

[22] Wagenaar, W., and G. Keren (1986), "Does the expert know? the reliability of predictions and confidence ratings of experts," In Hollnagel, E., G. Maneini and D. Woods (eds.), *Intelligent Decision Support in Process Environment.* Springer, Berlin.

[23] Van den Steen, E. (2005), "Too Motivated?," working paper.