Honesty vs. White Lies[*]

KyungMin Kim and Jonathan Pogach [†]

May 2009

Abstract

We study and compare the behavioral consequences of honesty and white lie in communication. An honest agent always recommends the ex post optimal policy to the principal, while a white liar may manipulate information but only for the sake of the principal. We identify the effects of honesty and white lie on communication and show that, even though a white lie is intended to be beneficial to the principal, the principal is often better off with a possibly honest agent than with a potential white liar. This result provides a fundamental rationale on why honesty is thought to be an important virtue in many contexts.

JEL Classification Numbers: C72, D82, D83.

Keywords : Cheap talk; honesty; white lie.

1 Introduction

Is honesty necessarily the best policy, or might a white lie be preferable when the truth hurts? The question of honesty versus white lies arises in many contexts. Communication between a husband and wife, between the general public and the monetary authority (or government), between an investor and a stock analyst (or manager), and between an employer and an employee are only a few examples. The purpose of this paper is to shed some light on this fundamental question, honesty versus white lies.

A discrepancy between honesty and white lies arises when a listener is uncertain over the motive of the speaker. When the speaker may have a different objective than the listener, then the latter will necessarily discount the credibility of communication. Then, the listener may value the flexibility of a white liar, who is not honest in communication only when it benefits the listener. However, as is true in other contexts, a listener may value commitment by the honest speaker to telling the truth.

---

[*] First draft: March 2008. We are indebted to George Mailath and Andrew Postlewaite for their constant support. We are also grateful to Navin Kartik, Philipp Kircher, Steven Matthews, and seminar participants in Penn Economic Theory club for many helpful comments.

[†] Department of Economics, University of Pennsylvania, McNeil Bldg, 3718 Locust Walk, Philadelphia, PA 19104. Kim : teddy.kyungmin.kim@gmail.com, Pogach: pogachj@econ.upenn.edu.

We model communication as a cheap talk game between a principal and an agent. In a cheap talk game, the agent receives private information and strategically transmits the information to the principal. In turn, the principal takes an action that affects utilities of both players. Crawford and Sobel (1982) (CS hereafter) show that, even though the interests of the agent and the principal are not perfectly aligned, information can be transmitted through cheap-talk (costless and non-binding) communication. However, the communication is typically *coarse* in the sense that only a limited amount of information can be delivered due to the strategic considerations of players.

We study the following two setups. Both introduce another layer of incomplete information in the cheap talk game, which can be naturally interpreted as the principal's uncertainty over the agent's motive or the agent's having an imperfect reputation. In the first setup (honesty model), the agent is honest with positive probability. An honest agent is behavioral and always recommends the ex post optimal policy to the principal. With the complementary probability, the agent is of the type in CS, that is, he is a strategic player and does not have perfectly aligned preferences with those of the principal. In the second setup (white lie model), the agent is always strategic but the principal is uncertain about the agent's bias. The agent may or may not have perfectly aligned preferences with those of the principal. The white liar is the agent with no bias. He knows that the principal is uncertain about his own motive and that the principal will discount the credibility of his report. Due to this consideration, he may lie but this is only for the sake of the principal.

Our first contribution is to identify the effects of honesty and white lie on communication. The effects are highlighted by equilibrium conditions that are generated by the possibilities of honesty and white lies. Commitment by the honest agent to telling the truth introduces an intrinsic meaning to messages.[1] Each message might be sent by the honest agent, and thus the principal must interpret each message differently. This imposes a lower bound on the amount of information transmission and generates a new equilibrium condition (Mass Balance condition). The condition in the white lie model (No Arbitrage condition for White liar) is similar to the one in the classic cheap talk results (No Arbitrage condition for Biased agent). This condition holds in equilibrium precisely because the white liar is a rational player and thus always transmit a message that is optimal conditional on his private information.

Our second contribution is to compare the welfare consequences of honesty and white lies and demonstrate that the principal is often better off with a possibly honest agent than with a potential white liar. To be more precise, let μ be the probability that the agent is honest or a white liar in each model. We show that when μ is sufficiently large, the principal is strictly better off in the honesty model than in the white lie model. We also show that when μ is sufficiently small, for at least 79.1% of bias values, a possibly honest agent is preferable to a potential white liar. Lastly, we explain by some numerical examples that the same conclusion would hold for intermediate cases.

There are two main driving forces for the welfare result. Both of them stem from commitment by the honest agent to telling the truth, but they highlight different aspects of commitment.

---

[1] In the standard cheap talk game, messages themselves do not carry any information. What matters is not which particular message arrives, but how the principal perceives each message.

The first reason relates to the fact that commitment simplifies communication protocol and thus reduces the loss due to strategic considerations. To see this point, suppose the principal is certain that the agent is honest. In this case, there is a unique communication outcome. The principal perfectly trusts the agent's report, and thus communication is fully efficient. This implies that when there is a small probability that the agent is biased, any communication outcome in the honesty model is close to that of perfect communication. Now suppose the agent is the white liar for sure. Perfect communication is still possible, but there are lots of other possibilities. For example, there may be no information transmission at all.[2] This multiplicity of equilibrium yields the following consequence in the white lie model: when there is a small probability that the agent is biased, the best communication outcome is far away from the perfect communication outcome. That is, the loss from imperfect communication does not vanish even when the agent is the white liar almost for sure.

The second reason is that commitment has the effect of enriching language used in communication. In the honesty model each message might be sent by the honest agent, and thus the principal interprets each message differently.[3] Therefore, different from CS and the white lie model, all messages are fully used in the honesty model, which allows freer communication between the principal and the *biased* agent. To see how this can improve the principal's welfare, recall the result by Dessein (2002). He shows that whenever communication is possible in equilibrium, both the principal and the biased agent would benefit if the former is committed to following the latter's recommendation. This implies that they have a coordination incentive, which is hindered by their selfishness in communication. The possibility of honesty, by enriching language used in communication, allows them to utilize the coordination incentive in a way that is not possible in the original cheap talk game and in the white lie model.[4]

Our welfare result provides a micro-foundation for the preference for honesty. Honesty is thought to be an important virtue in many contexts. For example, according to an Associated Press-Ipsos poll, "55% of those surveyed consider honesty, integrity and other values of character the most important qualities they look for in a presidential candidate *(USA Today, 3/12/2007).*" There are numerous discussions about honesty (or deception) in several disciplines (for example, philosophy, theology, sociology, psychology, and business ethics). In economics and political science, a number of works consider agents who have an exogenous preference for honesty. Nonetheless, it has not been answered in a formal model why honesty is valued over other characteristics. In particular, why should one prefer an honest agent to one who is strategic and completely altruistic? Our welfare result provides a fundamental rationale for the preference for honesty, focusing on the effect of honesty on communication.

---

[2] If the agent does not provide any meaningful information, the principal has no reason to pay attention to the agent's report, which in turn annihilates the agent's incentive to transmit information.

[3] In the honesty model, even though the principal takes a constant action on a set of messages, her posterior over the set of states is not constant over those messages. Only her conditional expectation is constant.

[4] This effect is similar to the one generated by "noise" in communication. See Blume, Board, and Kawamura (2007).

This paper is related to three branches of literature about strategic information transmission. The first strand includes Sobel (1985), Benabou and Laroque (1992), and Morris (2001). In our terminology, Sobel and Morris study the white lie cases, while Benabou and Laroque consider the honesty case. The focus of these works is the dynamic incentive of the agent to maintain credibility as well as to manipulate information. To highlight the intertemporal perspective, they consider simple stage games in which there are only two possible states. We study a static setting with a continuum of states and concentrate on the effects of honesty and white lie on communication itself.

Another branch is cheap talk games with uncertain bias. Morgan and Stocken (2003) study the white lie model in the context of stock analyst's problem where the stock analyst may or may not have an incentive to produce a favorable report to the firm's investment banking clients. They characterize two classes of equilibria, while we characterize the set of all equilibria. The complete characterization is important in our paper because the new equilibria have a significant welfare implication. Li and Madarasz (2007) consider the case where both agent types have non-zero biases and examine whether requiring the agent to disclose his own bias is necessarily welfare-improving.

The last branch considers the honest type in the cheap talk game. Olszewski (2004) studies a repeated cheap talk game in which there are a finite number of possible states, the principal has uncertainty over the agent's type, and the strategic agent prefers to be perceived as the honest type. He shows that if the agent's honesty concern is sufficiently strong, information can be fully revealed. Chen, Kartik, and Sobel (2008) introduce both the honest agent and the naive principal (who always follows the agent's recommendation) to the standard cheap talk game. They restrict attention to the class of equilibria in which the agent uses a non-decreasing strategy (in the sense that the strategic agent sends weakly higher messages in higher states) and show that only the most informative equilibrium in CS survives as the perturbation vanishes. We characterize a broader class of equilibria in the honesty model and concentrate on the comparison between honesty and white lie.

The remainder of the paper is organized as follows. The next section briefly reviews the standard cheap talk game. Then, we study the honesty model and the white lie model in Sections 3 and 4, respectively. We compare the two models in Section 5 and conclude in Section 6.

2 Review of the Standard Cheap Talk Game

2.1 Cheap talk game

There are two players, the principal (she) and the agent (he). The agent observes a random variable θ and strategically transmits information on θ to the principal. A random variable θ is drawn from a uniform distribution with support on $\Theta = [0, 1]$. The principal takes an action, denoted by y, from the real line that affects utilities of both players. The principal's utility function is $U^P(y, \theta) \equiv -(y - \theta)^2$, while the agent's is $U^A(y, \theta, b) \equiv -(y - (\theta + b))^2$ where $b \in (0, 1)$. As usual, b is interpreted as the "bias" of the agent. When the true state is θ, the most preferred policy to the principal is θ, while that of the agent is $\theta + b$. Without loss of generality, we assume that the set of messages, M, and the set of feasible policies, Y, are given by $[0, 1]$.

The principal's strategy is her policy choice rule $y : M \to Y$ where $y(m)$ is the action she takes after receiving message m. The agent's strategy is her reporting rule $r : \Theta \to \Delta(M)$ where $\Delta(M)$ is the set of probability measures over the set of messages, M, and $r(\theta)$ is her reporting policy conditional on observation of state θ.

Definition 1 *The strategy profile (r^*, y^*) constitutes an equilibrium if*

(1) given y^*, if m' is sent by the agent (m' is in the support of $r_b(\theta)$), then

$$m' \in \arg\max_{m \in M} U^A(y^*(m), \theta, b) = -(y^*(m) - (\theta + b))^2, \text{ and}$$

(2) given r^*, for all $m \in M$,

$$y^*(m) \in \arg\max_{y} E[U^P(y, \theta)|m].$$

Since the principal has a quadratic utility function, the second requirement reduces to $y^*(m) = E[\theta|m]$. That is, the principal's optimal strategy is always to choose the conditional expectation of θ.

2.2 Equilibrium Characterization

An equilibrium in the cheap talk game is characterized by a partition $\{\theta_0 = 0, \theta_1, ..., \theta_n = 1\}$ and a sequence $\{y_1, ..., y_n\}$. The agent sends an essentially identical message in each partition element, $[\theta_{k-1}, \theta_k]$. The principal infers only in which partition element the true state, θ, lies, and thus takes only a finite number of actions, $\{y_1, ..., y_n\}$. The following two conditions, each of which corresponds to an equilibrium requirement, are necessary and sufficient:

$$\begin{array}{ll} y_k = \frac{\theta_{k-1}+\theta_k}{2}, k = 1, ..., n., & \text{(BR: Best Response)} \\ (\theta_k + b) - y_k = y_{k+1} - (\theta_k + b), k = 1, ..., n-1., & \text{(NA: No Arbitrage)} \end{array}$$

NA states that the agent must be indifferent between y_k and y_{k+1} at state θ_k. This is necessary and sufficient for the agent's optimality because the agent's utility function is quadratic and thus the single crossing property holds.[5]

Figure 1 illustrates how the two conditions interact. When there are two partition elements, $[0, \theta_1]$ and $[\theta_1, 1]$, the principal takes either y_1 or y_2. BR requires that the principal's action (y_k) be the conditional expectation of θ, while NA states that the agent's most preferred action at state θ_1, $\theta_1 + b$, is at the average of the two actions induced in equilibrium.

---

[5] If the agent prefers y to y' at θ where $y > y'$, then he prefers y to y' at any state θ' such that $\theta' > \theta$, and vice versa.

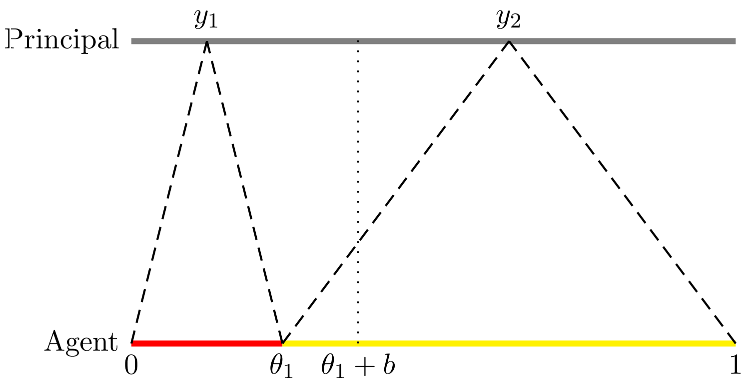

Figure 1: An equilibrium with two partition elements in the standard cheap talk game.

2.3 Incentive Compatibility Lemma

The following lemma addresses the general incentive compatibility issue of the strategic agent. Whenever the agent is rational with preferences of the form above, the results of the lemma will apply. We will extensively exploit the implications of the lemma in later sections.

Lemma 1 *Let* $z(\cdot,\cdot) : \Theta \times T \to Y$ *represent the outcome of the cheap talk game where* T *is the set of agent types and* $z(\theta, b)$ *is the policy chosen by the principal when the agent has bias* b *and observes state* θ*. Denote by* $V^S(\theta, b)$ *the indirect utility function of the agent with bias* b*, that is,* $V^S(\theta, b) = U^A(z(\theta, b), \theta, b)$*.*

$z(\cdot, b)$ *is incentive compatible for the agent with bias* b *if and only if*

(i) $z(\cdot, b)$ *is nondecreasing,*

(ii) $V^S(\cdot, b)$ *is absolutely continuous, and*

(iii) if $z_1(\theta, b) = \partial z(\theta, b)/\partial\theta$ *exists, then*

$$\frac{\partial U^A(z(\theta, b), \theta, b)}{\partial y} \cdot \frac{\partial z(\theta, b)}{\partial \theta} = 0.$$

Proof. See Appendix. ■

The lemma implies that the incentive compatibility of the strategic agent imposes a severe restriction on equilibrium outcome. In particular, part (iii) of the lemma implies that in equilibrium, for almost all states, the agent must induce either a constant action around each state ($\partial z(\theta)/\partial\theta = 0$) or induces his most preferred policy at each state ($\partial U^A(z(\theta, b), \theta, b)/\partial y = 0$). To understand this, consider an equilibrium outcome in CS and the case where the principal delegates the authority to choose a policy to the agent in the sense of Dessein (2002).[6] The agent has no incentive to deviate in both cases, but for different reasons. In the latter, the agent is achieving the first-best payoff at

---

[6] In other words, the principal is committed to following the agent's recommendation.

each state, while in the former, a small deviation does not change the outcome and a large deviation lowers the agent's payoff. In other words, in CS, $\partial z(\theta)/\partial\theta = 0$ almost everywhere, while in Dessein, $\partial U^A(z(\theta, b), \theta, b)/\partial y = 0$ everywhere.

3 Honesty

3.1 Setup

In this section, the agent is honest with probability $\mu \in (0, 1)$, and is strategic with bias b with the complementary probability. The honest agent is behavioral and his strategy is to send message θ when he observes state θ. The strategies of the principal and the biased agent are the same as in the previous section. We denote by r_b the biased agent's strategy.

Definition 2 *The strategy profile (r_b^*, y^*) constitutes an equilibrium if*

(1) given y^*, if m' is sent by the biased agent (m' is in the support of $r_b(\theta)$), then

$$m' \in \arg\max_{m\in M} U^A(y^*(m), \theta, b) = -(y^*(m) - (\theta + b))^2, \text{ and}$$

(2) given r_b^*, for all $m \in M$,

$$\begin{aligned} y^*(m) &\in \arg\max_y E_{\mu, r_b^*}[U^P(y, \theta)|m], \\ &\Leftrightarrow y^*(m) = E_{\mu, r_b^*}[\theta|m], \end{aligned}$$

where E_{μ, r_b^*} is the conditional expectation operator generated by μ, the honest agent's behavior and r_b^*.

3.2 CS Equilibrium Outcomes

We first consider an equilibrium in which every message induces the principal to take the same action. In CS, this equilibrium is called a "babbling" equilibrium because the agent essentially randomizes over the entire message space independently of his private information. Since the honest agent always reports the true state, a babbling equilibrium does not exist.

A no communication equilibrium still exists in a different form as long as $\mu \leq 1/2$. Consider the following strategy profile.

$$\begin{aligned} r_b(\theta) &= \begin{cases} 1 - \theta, & \text{with probability } \frac{\mu}{1-\mu}, \\ m \sim U[0, 1], & \text{with probability } \frac{1-2\mu}{1-\mu}, \end{cases} \\ y(m) &= \frac{1}{2}, \forall m. \end{aligned}$$

In this profile, at state θ, the biased agent sends message $1 - \theta$ with probability $\mu/(1 - \mu)$ and randomizes over the entire message space with the complementary probability. Having received any message m, the conditional probability that the message came from the honest agent is equal to the unconditional probability of the honest agent (μ). Similarly, the conditional probability that the message came from the biased agent who is reporting $1-\theta$ and the conditional probability that it came from the biased agent who is uniformly randomizing over $[0,1]$ are μ and $1-2\mu$, respectively. As such, for any m, the conditional expectation of the principal on the true state is

$$E[\theta|m] = \mu m + (1-\mu)\left(\frac{\mu}{1-\mu}(1-m) + \frac{1-2\mu}{1-\mu}\int_0^1 \theta d\theta\right) = \frac{1}{2}.$$

Therefore, the principal takes a single action independently of message. This, in turn, makes the biased agent indifferent over all messages.

We note that the principal updates her belief nontrivially. In the babbling equilibrium in CS, every posterior distribution is equal to the prior distribution. Under the strategy above, the principal's posterior puts mass on the message she received, m, and its counterpart, $1-m$. Therefore, her posterior is different for each message even though her action is independent of message. It is only the conditional expectation of the true state that is constant across messages.

This equilibrium construction generalizes to all equilibrium outcomes in CS. We can use the same trick interval by interval: the biased agent reports exactly the "opposite" state[7] in each partition element with probability $\mu/(1-\mu)$ and randomizes over the interval with the remaining probability. This, together with the subsequent discussion for the case where $\mu > 1/2$, allows us to establish the following proposition.

Proposition 1 *Any equilibrium outcome in CS can be supported as an equilibrium outcome in the honesty model if and only if* $\mu \leq 1/2$.

3.3 Mass Balance Condition

If $\mu > 1/2$, the previous strategy profile is not well defined and thus cannot be used as part of a "no communication" equilibrium or other CS equilibria. The following lemma shows that for $\mu > 1/2$ no other strategy profile induces any CS equilibrium outcome.

Lemma 2 ***(Mass Balance Condition)*** *Suppose* $0 \leq m' < m'' \leq 1$, $0 \leq \theta' < \theta'' \leq 1$, *and*

$$\begin{aligned} \overline{y} &= B(\mu, m', m'', \theta', \theta'') \\ &\equiv \frac{\mu(m''-m')}{\mu(m''-m') + (1-\mu)(\theta''-\theta')}\frac{m'+m''}{2} + \frac{(1-\mu)(\theta''-\theta')}{\mu(m''-m') + (1-\mu)(\theta''-\theta')}\frac{\theta'+\theta''}{2}. \end{aligned}$$

There exists a collection of probability measures $\{r(\theta), \theta \in [\theta', \theta'']\} \subset \Delta([m', m''])$ *such that*

$$E_{\mu,r}[\theta|\mathcal{M}] = \overline{y} \text{ for any Borel set } \mathcal{M} \text{ in } [m', m''],$$

---

[7] Formally, if the partition element is $[\theta', \theta'']$, then the opposite state to $\theta \in [\theta', \theta'']$ is $\theta' + \theta'' - \theta$.

if and only if

$$\mu(m'' - m')^2 \leq (1-\mu)(\theta'' - \theta')^2,$$

where $E_{\mu,r}$ *is the conditional expectation operator generated by* μ *and* r.

Proof. See Appendix. ■

This lemma relates the inherent informational content of messages to the biased agent's strategy by providing a necessary and sufficient condition for the biased agent to be able to induce a single inference (and, consequently, a single optimal action for the principal) over an interval of messages. The condition shows how the honest agent's commitment imposes a lower bound on the amount of information transmission and thus highlights the effects of possible honesty on communication. Intuitively, when μ is small, the principal can be sufficiently pessimistic that she may completely ignore the agent's message. When μ is large, however, it is likely that the principal gets the correct information, and consequently, she must take into account the agent's report in her decision making.

We will make use of the following reporting strategy of the biased agent throughout the paper:

$$r_b(\theta) = \begin{cases} \frac{m''-m'}{\theta''-\theta'}(\theta'' - \theta) + m', & \text{with probability } \frac{\mu}{1-\mu}\frac{(m''-m')^2}{(\theta''-\theta')^2}, \\ m \sim U[m', m''], & \text{with probability } 1 - \frac{\mu}{1-\mu}\frac{(m''-m')^2}{(\theta''-\theta')^2}. \end{cases}$$

We denote this strategy by "$r_b(\theta) = \tilde{r}([m', m''])$ if $\theta \in [\theta', \theta'']$."

3.4 Type I Equilibrium

For $\mu > 1/2$, the following strategy profile is a natural extension of no communication equilibrium: for some $m_0 > 0$,

$$\begin{aligned} r_b(\theta) &= \tilde{r}([m_0, 1]), \forall\theta \\ y(m) &= \begin{cases} m, & \text{if } m < m_0, \\ B(\mu, m_0, 1, 0, 1), & \text{if } m \geq m_0. \end{cases} \end{aligned}$$

In this strategy profile, the biased agent sends messages above m_0. The principal believes that only the honest agent sends messages below m_0, and so perfectly trusts their contents.

The strategy profile is specifically designed to overcome the binding MB. Notice that MB is satisfied if m_0 is high enough: $\mu(1-m_0)^2 \leq 1-\mu$. Of course, we must ensure that the biased agent does not want to deviate to some message below m_0. Therefore,

$$m_0 \leq b \text{ and } |B(\mu, m_0, 1, 0, 1) - b| \leq b - m_0.$$

The first inequality guarantees that the biased agent cannot implement her most preferred policy by deviating to below m_0 at state 0. The second inequality ensures that the biased agent prefers

$B(\mu, m_0, 1, 0, 1)$ to m_0 at state 0. By the single crossing property, the biased agent does not deviate at any other state.

This strategy profile also constitutes an equilibrium when $\mu < 1/2$. Since the mass balance condition is satisfied vacuously, the strategy profile is an equilibrium as long as the incentive compatibility condition is satisfied.

Definition 3 ***(Type I Strategy Profile)*** *A type I strategy profile is represented by two partitions in a unit interval,* $\{0, m_0, m_1, ..., m_n = 1\}$ *and* $\{\theta_0 = 0, \theta_1, ..., \theta_n = 1\}$, *and a sequence,* $\{y_1, ..., y_n\}$, *such that*[8]

$$\begin{aligned} r_b(\theta) &= \tilde{r}([m_{k-1}, m_k]), \text{ if } \theta \in [\theta_{k-1}, \theta_k], \forall k = 1, .., n, \\ y(m) &= \begin{cases} m, & \text{if } m < m_0, \\ y_k, & \text{if } m \in [m_{k-1}, m_k], \forall k = 1, .., n. \end{cases} \end{aligned}$$

In this strategy profile, the biased agent sends $[m_{k-1}, m_k]$ on $[\theta_{k-1}, \theta_k]$ and does not send any message below m_0. The following conditions are necessary and sufficient for a type I strategy profile to be an equilibrium.

$$|y_1 - b| \leq b - m_0 \text{ and } m_0 \leq b \text{ if } m_0 > 0, \tag{IC}$$

$$y_k + y_{k+1} = 2(\theta_k + b), \forall k = 1, ..., n-1, \tag{NA}$$

$$y_k = B(\mu, m_{k-1}, m_k, \theta_{k-1}, \theta_k), \forall k = 1, ..., n, \tag{BR}$$

$$\mu(m_k - m_{k-1})^2 \leq (1-\mu)(\theta_k - \theta_{k-1})^2, \forall k = 1, ..., n. \tag{MB}$$

IC is the incentive compatibility condition for the biased agent to not deviate to below m_0. NA and BR are the same conditions as in Section 2. BR is modified to reflect uncertainty over the agent's type. MB is the mass balance condition in Lemma 2. As μ tends to 0, NA and BR converge to equilibrium conditions in CS, and IC and MB become negligible.

Figure 2 shows an example of type I equilibrium. The biased agent sends messages $[m_0, m_1]$ on $[0, \theta_1]$ and messages $[m_1, 1]$ on $[\theta_1, 1]$. The principal believes that if message m is below m_0, it was sent by the honest agent and thus implements m. If the message is above m_0, then she chooses y_1 or y_2 depending on whether the message lies in $[m_0, m_1]$ or $[m_1, 1]$. The two policies induced by the biased agent, y_1 and y_2, are conditional expectations of the true state (BR), and the biased agent is indifferent between y_1 and y_2 at state θ (NA). Though $m_1 - m_0 > \theta_1$, MB is satisfied because μ is small.

3.5 Type II Equilibrium

For μ sufficiently large, there cannot exist any type I equilibrium. As μ increases, m_0 must increase so that MB holds. But then IC binds because b is an upper bound of m_0 in type I equilibrium.

---

[8] We do not explicitly specify what the agent's reporting policies are at the boundary points of each partition element. They do not affect both players' ex ante utilities because the set has zero measure.

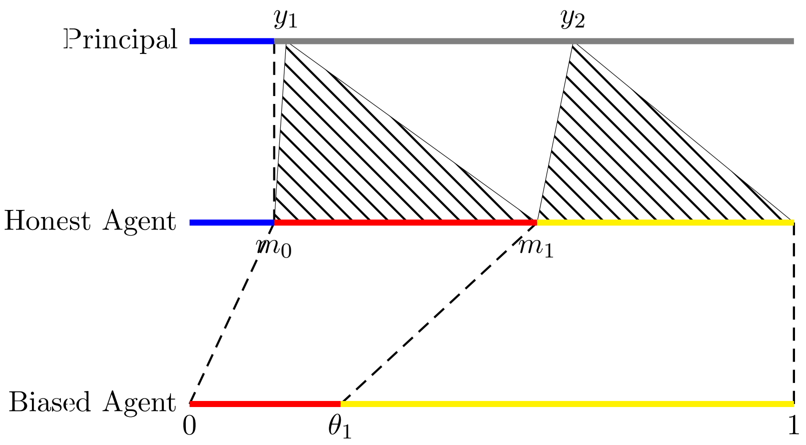

Figure 2: Type I equilibrium when $b = 0.15$ and $\mu = 0.1$.

Therefore, we need an alternative way to consume an excess of mass with the honest agent in order to satisfy both IC and MB. The key to this issue is in Lemma 1. According to Lemma 1, $\Theta = [0, 1]$ can be decomposed into the following three subsets:

$$\begin{aligned}
\Theta_1 &= \left\{\theta \in \Theta : z(\theta, b) = y^A(\theta, b) = \theta + b\right\} \\
&= \{\theta \in \Theta : \text{the agent implements his most preferred policy}\}, \\
\Theta_2 &= \left\{\theta \in \Theta : z_1(\theta, b) = \frac{\partial z(\theta, b)}{\partial \theta} = 0\right\} \\
&= \{\theta \in \Theta : \text{the same action is implemented around } \theta\}, \\
\Theta_3 &= \{\theta \in \Theta : z_1(\theta, b) \text{ does not exist}\}.
\end{aligned}$$

In many variations of the cheap talk game, Θ_2 has full measure; equilibrium features a partitioning of Θ with which a constant action is induced on each partition element. However, for μ sufficiently large, Θ_2 cannot have full measure because of the conflict between MB and IC. Then, Θ_1 is the only alternative of use, as the non-decreasing property of $z(\cdot, b)$ (see part (i) of Lemma 1) implies that Θ_3 has zero measure. On Θ_1, the biased agent induces his own optimal policy. This possibility does not arise in CS and many other contexts. However, Θ_1 is *necessary* in the honesty model for μ sufficiently large.

Definition 4 ***(Type II Strategy Profile)*** *A type II strategy profile is represented by two partitions*

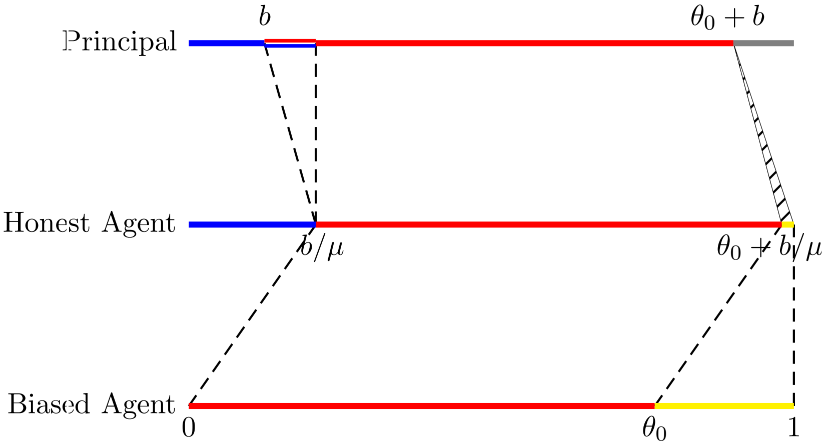

Figure 3: Type II equilibrium when $b = 1/8$ and $\mu = 0.6$.

in a unit interval, $\{0, m_0, ..., m_n\}$ *and* $\{0, \theta_0, ..., \theta_n\}$*, and a sequence,* $\{y_1, ..., y_n\}$*, such that*

$$r_b(\theta) = \begin{cases} \theta + b/\mu, & \text{if } \theta \in [0, \theta_0], \\ \tilde{r}([m_{k-1}, m_k]), & \text{if } \theta \in [\theta_{k-1}, \theta_k], \forall k \geq 1, \end{cases}$$

$$y(m) = \begin{cases} m, & \text{if } m \leq b/\mu, \\ \mu m + (1-\mu)(m - b/\mu), & \text{if } b/\mu < m \leq m_0, \\ y_k, & \text{if } m \in [m_{k-1}, m_k], \end{cases}$$

In this strategy profile, the biased agent induces his own optimal policy on $[0, \theta_0]$. For example, suppose $\theta = 0$. Then the biased agent sends message b/μ. When the principal receives this message, her inference on θ is $\mu \cdot (b/\mu) + (1-\mu) \cdot 0 = b$, which is optimal to the biased agent. The conditions required for this strategy profile to be an equilibrium are as follows:

$$\begin{aligned} & b/\mu \leq \theta_0 + b, && \text{(IC)}, \\ & m_0 = \theta_0 + b/\mu, && \text{(EL)}, \\ & y_1 = \theta_0 + b, && \text{(NA0)}, \\ & y_k + y_{k+1} = 2(\theta_k + b), \forall k \geq 1, && \text{(NA)}, \\ & y_k = B(\mu, m_{k-1}, m_k, \theta_{k-1}, \theta_k), \forall k \geq 1, && \text{(BR)}, \\ & \mu(m_k - m_{k-1})^2 \leq (1-\mu)(\theta_k - \theta_{k-1})^2, \forall k \geq 1, && \text{(MB)}. \end{aligned}$$

NA, BR, and MB are the same as before. NA0 is the condition required to prevent the biased agent from deviating to $[b/\mu, m_0]$ for $\theta > \theta_0$. IC guarantees that the deviation to $[0, b/\mu)$, where the principal perfectly trusts messages, is not profitable. EL (equal length) is an obvious requirement from the structure of equilibrium.

Example 1 *For* $\mu \geq 1/2$ *and* $b \leq \mu / \left(1 + \sqrt{\mu(1-\mu)}\right)$*, the following strategy profile is a Type II* *equilibrium.*

$$r_b(\theta) = \begin{cases} \theta + b/\mu, & \text{if } 0 \leq \theta < \theta_0, \\ [\theta_0 + b/\mu, 1], & \text{if } \theta \in [\theta_0, 1], \end{cases}$$

$$y(m) = \begin{cases} m, & \text{if } m \leq b/\mu, \\ m - \frac{1-\mu}{\mu} b, & \text{if } b/\mu < m < \theta_0 + b/\mu, \\ \theta_0 + b, & \text{if } m \in [\theta_0 + b/\mu, 1], \end{cases}$$

where $\theta_0 = 1 - b - b\sqrt{(1-\mu)/\mu}$.

Figure 1 shows the structure of this equilibrium. The biased agent sends message $\theta + b/\mu$ *if* $\theta < \theta_0$, *and sends messages* $[\theta_0 + b/\mu, 1]$ *on* $[\theta_0, 1]$. *The principal perfectly trust the agent's report if the message is below* b/μ, *while she believes that any message above* μ *might be sent by the biased agent. The principal discounts the agent's recommendation by the same amount,* $(1-\mu)/\mu \cdot b$, *on* $[b/\mu, \theta_0 + b/\mu]$, *while she takes a single action,* $\theta + b$, *on* $[\theta_0 + b/\mu, 1]$. *Notice that the biased agent induces her most preferred policy on* $[0, \theta_0]$ *and actions* $[b, b/\mu]$ *are chosen both when* $m \in [b, b/\mu]$ *and when* $m \in [b/\mu, 2b/\mu - b]$.

3.6 Other Possibilities

The two equilibrium structures we have introduced do not exhaust all equilibria in the honesty model. But they are the only classes of equilibria that satisfy two natural properties of strategy profile, one for each player. We introduce each property and present an example of equilibrium that violates the property.

Definition 5 *(Convexity) The biased agent's reporting strategy is convex if there exists* $m_0 \in [0,1]$ *such that the biased agent never sends messages below* m_0 *and sends all messages above* m_0.

That is, if the biased agent's strategy is convex, then there is a cutoff point in the message space such that all messages above the point are contaminated by the biased agent. This is natural because the biased agent, due to his positive bias, is less willing to deviate to lower messages, but there is an equilibrium that violates this property.

Example 2 *Suppose* $\mu = 1/8$ *and* $b = 2/5$. *The following strategy profile is an equilibrium but does not satisfy convexity. See the left panel of Figure 4 for the equilibrium structure.*

$$r_b(\theta) = m \sim U[0, 1/8] \cup [1/4, 1], \forall \theta,$$

$$y(m) = \begin{cases} m, & \text{if } m \in [1/8, 1/4], \\ 509/1008, & \text{otherwise.} \end{cases}$$

Definition 6 *(Monotonicity) The principal's strategy* y *is monotone if for all* $m' > m$ *such that* $m \in \text{supp } r_b(\theta)$ *and* $m' \in \text{supp } r_b(\theta')$ *for some* θ *and* θ', $y(m') \geq y(m)$.

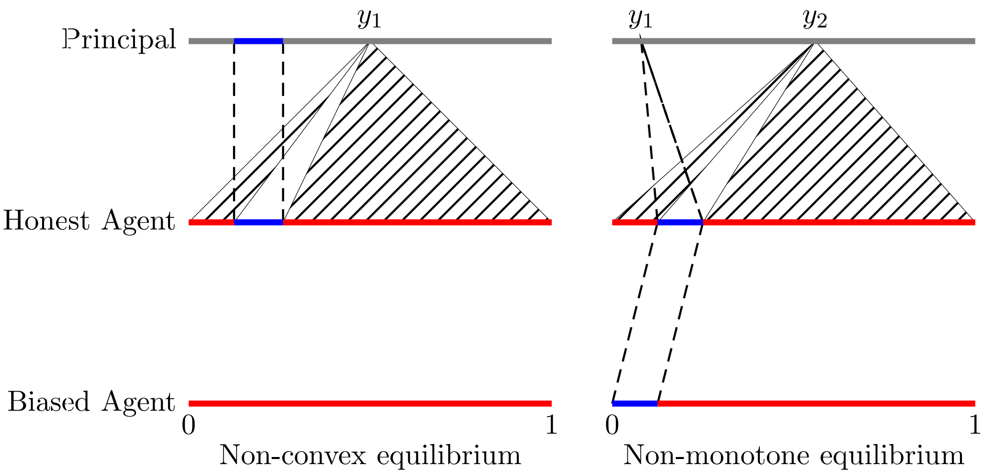

Figure 4: Equilibria that do not satisfy convexity or monotonicity

In words, for those messages that the biased agent may send, the higher message the principal gets, the weakly higher action she implements. This restriction makes the biased agent's strategy weakly monotone in the sense that a strictly higher action can be induced only by sending a message higher than any message that would induce a lower action.

Example 3 *Suppose $\mu = 1/8$ and $b = 87/448$. The following strategy profile is an equilibrium but does not satisfy monotonicity. See the right panel of Figure 4 for the equilibrium structure.*

$$r_b(\theta) = \begin{cases} m \sim U[1/8, 1/4], & \textit{if } \theta \in [0, 1/8] \\ m \sim U[0, 1/8] \cup [1/4, 1], & \textit{otherwise,} \end{cases}$$

$$y(m) = \begin{cases} 5/64, & \textit{if } m \in [1/8, 1/4], \\ 251/448,, & \textit{otherwise.} \end{cases}$$

Proposition 2 *Any equilibrium in which the biased agent's strategy is convex and the principal's strategy is monotone is either Type I or II.*

Proof. See Appendix. ■

4 White Lie

4.1 Setup

In this section, the principal is uncertain about the agent's bias. The agent has no bias with probability μ, and has bias b with the complementary probability. We denote by r_0 and r_b the white liar's and the biased agent's strategies, respectively.

Definition 7 *The strategy profile* (r_0^*, r_b^*, y^*) *constitutes an equilibrium if*

(1) given y^**, if* m' *is sent by the biased agent at state* θ *(in the support of* $r_b^*(\theta)$*), then*

$$m' \in \arg\max_{m \in M} U^A(y^*(m), \theta, b) = -(y^*(m) - (\theta + b))^2,$$

(2) given y^**, if* m' *is sent by the white liar at state* θ *(in the support of* $r_0^*(\theta)$*), then*

$$m' \in \arg\max_{m \in M} U^A(y^*(m), \theta, 0) = -(y^*(m) - \theta)^2, \text{ and}$$

(3) given (r_0^*, r_b^*)*, for all* $m \in M$*,*

$$\begin{aligned} y^*(m) &\in \arg\max_y E_{\mu, r_0^*, r_b^*}[U^P(y, \theta)|m] \\ &\Leftrightarrow y^*(m) = E_{\mu, r_0^*, r_b^*}[\theta|m]. \end{aligned}$$

where E_{μ, r_0^*, r_b^*} *is the conditional expectation operator generated by* μ*,* r_0^* *and* r_b^**.*

4.2 CS Equilibrium Outcomes

Proposition 3 *For* $\mu > 0$*, only the no communication outcome in CS is supported as an equilibrium outcome in the white lie model.*

A babbling equilibrium always exists in the white lie model. If both types of agent randomize over the entire message space independently of state, then the principal cannot make any meaningful inference and thus takes a single action independently of message. This in turn makes both types of agent indifferent over all messages independently of state.

Any other equilibrium outcome in CS is not supported as an equilibrium outcome in the white lie model. To see this, take an equilibrium that is characterized by a partition $\{\theta_0, \theta_1, ..., \theta_n\}$ and a sequence $\{y_1, ..., y_n\}$ where $n > 1$. Suppose this outcome is supported as an equilibrium outcome in the white lie model. By Lemma 1, the biased agent must be indifferent between y_k and y_{k+1} at state θ_k. But the white liar wants to induce a lower action than the biased agent at any state and thus strictly prefers y_k to y_{k+1} at state θ_k. Consequently, the white liar deviates when she observes θ slightly above θ_k.

4.3 No Arbitrage Condition for the White Liar

The previous discussion suggests a new equilibrium condition in the white lie model, No Arbitrage condition for the White liar (NAW): at the boundary state of two partition elements, the white liar must be indifferent between the two induced actions. This condition is a consequence of the white liar's flexibility: whenever this condition does not hold, the white liar can increase the principal's expected payoff by adjusting his report. Together with No Arbitrage condition for the Biased agent (NAB), this implies that if two actions, and no action in between, are induced by both types in equilibrium, then both the white liar and the biased agent must be indifferent between the two actions at their own boundary states.

4.4 Equilibrium Characterization

The following proposition characterizes the set of *all* equilibria in the white lie model.

Proposition 4 *Any equilibrium in the white lie model is characterized by two finite partitions in a unit interval,* $\{0, \theta_0^0, \theta_1^0, ..., \theta_n^0 = 1\}$ *and* $\{\theta_0^b = 0, \theta_1^b, ..., \theta_n^b = 1\}$*, a finite sequence,* $\{y_1, ..., y_n\}$*, and* $n_0 \in N \cup \{\infty\}$ *if* $\theta_0^0 > 0$ *such that*[9]

$$\begin{array}{ll} y_1 + \frac{2n_0-1}{2n_0}\theta_0^0 \leq 2b \text{ if } \theta_0^0 > 0, & (ICB) \\ y_1 = \frac{2n_0+1}{2n_0}\theta_0^0 \text{ if } \theta_0^0 > 0, & (ICW) \\ y_k + y_{k+1} = 2(\theta_k^b + b), & (NAB) \\ y_k + y_{k+1} = 2\theta_k^0, & (NAW) \\ y_k = B(\mu, \theta_{k-1}^0, \theta_k^0, \theta_{k-1}^b, \theta_k^b). & (BR) \end{array}$$

Proof. See Appendix. ■

Figure 5 illustrates the equilibrium structure. Type i agent induces the principal to take an action y_k when she observes θ in $[\theta_{k-1}^i, \theta_k^i]$. The white liar separates from the biased agent when he observes θ in $[0, \theta_0^0]$. If $n_0 < \infty$ then the white liar induces the principal to take an action $\frac{2k-1}{2n_0}\theta_0^0$ when he observes $\theta \in [\frac{k-1}{n_0}\theta_0^0, \frac{k}{n_0}\theta_0^0]$ (see the left panel of Figure 5). If $n_0 = \infty$ then the white liar induces the principal to take their most preferred action when the true state is in $[0, \theta_0^0]$ (see the right panel of Figure 5). The latter case is similar to type I equilibrium in the honesty model. The principal chooses an action, taking into account her uncertainty over the agent's bias.

BR and NAB are straightforward generalizations of the equilibrium conditions in CS. Also, as we explained, NAW must hold due to the rationality of the white liar. ICB guarantees that the biased agent does not deviate to any policy below y_1. Similarly, ICW ensures that the white liar does not deviate to any policy below y_1 at state above θ_0^0.[10]

Example 4 *Consider the following strategy profile.*

$$\begin{aligned} r_0(\theta) &= \begin{cases} m \sim U[\frac{k-1}{n_0}\theta^*, \frac{k}{n_0}\theta^*], & \text{if } \theta \in [\frac{k-1}{n_0}\theta^*, \frac{k}{n_0}\theta^*], \\ m \sim U[\theta^*, 1], & \text{if } \theta \in [\theta^*, 1], \end{cases} \\ r_b(\theta) &= \quad m \sim U[0, 1], \quad \text{if } \theta \in [0, 1], \\ y(m) &= \begin{cases} \frac{2k-1}{2n_0}, & \text{if } m \in [\frac{k-1}{n_0}\theta^*, \frac{k}{n_0}\theta^*], \\ \frac{2n_0+1}{2n_0}\theta^*, & \text{if } m \in [\theta^*, 1], \end{cases} \end{aligned}$$

---

[9] Morgan and Stocken (2003) characterized the cases where $n_0 = 0$ ($\theta_0^0 = 0$) or $n_0 = \infty$.

[10] ICW can be understood as another NAW.

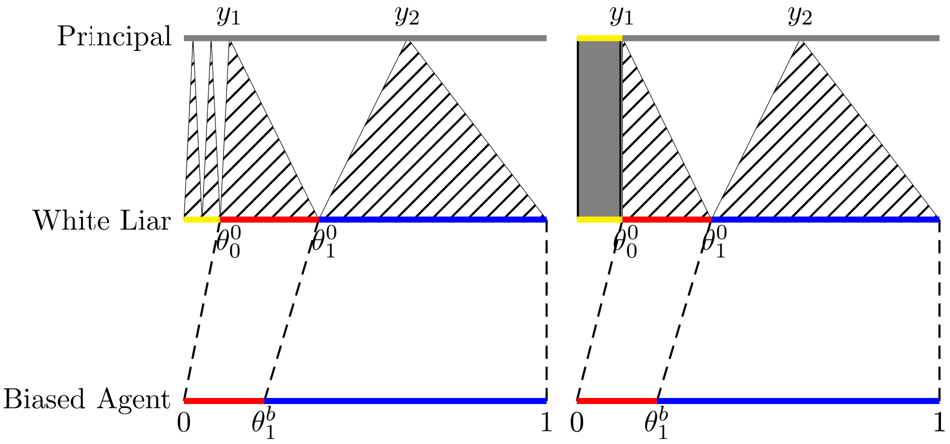

Figure 5: Equilibrium structure in the white lie model. There are two partition elements below θ_0^0 in the left panel, while the white liar induces her most preferred policy on $[0, \theta_0^0]$ in the right panel.

where

$$\theta^* = \frac{2n_0}{2n_0 + 1 + \sqrt{4(1-\mu)n_0^2 + 4(1-\mu)n_0 + 1}}.$$

This strategy profile is an equilibrium if and only if

$$\frac{4n_0+1}{2n_0}\theta^* \leq 2b.$$

An interesting case is when n_0 is equal to infinity (see the right panel of Figure 5). In this case, the white liar induces the principal to take their most preferred policy on $[0, \theta^*]$. The equilibrium structure is very similar to that of type I equilibrium in the honesty model (in particular, that of Example 1).

Here we present a way to find the set of equilibria. First, given μ and b, suppose there exists an equilibrium with n partition elements and $\theta_0^0 = 0$. From all equality conditions, we find

$$\begin{aligned}
\theta_k^b &= k\theta_1^b + b(k-1)\left[2(1-\mu)k + \mu - \frac{\mu(1-\mu)b}{\theta_1^b + \mu b}\right], k = 2, .., n-1, \qquad (1) \\
\theta_k^0 &= \theta_k^b + b, \forall k = 1, ..., n-1 \\
3\theta_{n-1}^b &- \theta_{n-2}^b + 2b(2-\mu) = 2B(\mu, \theta_{n-1}^0, \theta_n^0, \theta_{n-1}^b, \theta_n^b)
\end{aligned}$$

Combined with $\theta_n^b = \theta_n^0 = 1$, this system of equations has a unique solution for each n. The necessary condition for this equilibrium to exist is $\theta_1^b > 0$. As in CS, this condition provides an upper bound on the possible number of partition elements.

Now suppose $\theta_0^0 > 0$ and there are n_0 partition elements below θ_0^0. By a similar algebra, we get

$$\begin{aligned} \theta_k^b &= 2k\theta_1^b + (k-1)\left[2(1-\mu)bk - \frac{\mu(b+\theta_0^0)(\theta_1^b + b - \theta_0^0)}{\theta_1^b + \mu(b-\theta_0^0)}\right], k = 2, .., n-1 \qquad (2) \\ \theta_0^0 &= \frac{2n_0}{2n_0+1}B(\mu, \theta_0^0, \theta_1^0, 0, \theta_1^b), \\ \theta_k^0 &= \theta_k^b + b, \forall k = 1, ..., n-1. \\ 3\theta_{n-1}^b &- \theta_{n-2}^b + 2b(2-\mu) = 2B(\mu, \theta_{n-1}^0, \theta_n^0, \theta_{n-1}^b, \theta_n^b) \end{aligned}$$

Again using $\theta_n^b = \theta_n^0 = 1$, we can identify all equilibrium values. In this equilibrium, there are two restrictions on the number of partition elements above θ_0^0. The first one is the same as above: $\theta_1^b > 0$. This condition also converges to the one in CS and imposes an upper bound on the possible number of partition elements. The second condition is the incentive compatibility for the biased agent: $\theta_0^0 = \frac{2n_0}{2n_0+1}y_1 \leq b$. This condition imposes a lower bound on n. Different from the case with $\theta_0^0 = 0$, an equilibrium with $\theta_0^0 > 0$ exists only when there are enough number of partition elements above θ_0^0.

5 Welfare Comparison

Now we compare the maximal expected utilities the principal can achieve in the two models.

Almost Honest Agent vs. Almost White Liar

Proposition 5 *For any $b \in (0,1)$, there exists a $\bar{\mu}(b) < 1$ such that if $\mu > \bar{\mu}(b)$, then the principal is strictly better off with a possibly honest agent than with a potential white liar.*

Suppose μ is close to 1 and consider the honesty model. Fix any strategy for the biased agent and suppose the principal simply follows the recommendation of the agent. That is, suppose the principal's strategy is $y(m) = m$. Then, the principal takes her most preferred policy with at least probability μ. Since the principal cannot do no worse than this strategy, his ex-ante expected utility is close to that of the perfect communication. Intuitively, when $\mu = 1$, there is a unique equilibrium in which the first-best outcome is achieved. Therefore, as μ tends to 1, any equilibrium outcome must converge to the perfect communication outcome.

Next, consider the white lie model. It is straightforward from Equations (1) and (2) that the number of possible partition elements above θ_0^0 is bounded independently of μ. For example, $1/b$ is an upper bound on the number of partition elements for the case where $\theta_0^0 = 0$. In other words, the perfect communication equilibrium in the white lie model when $\mu = 1$ is not lower hemi-continuous. Therefore, the loss from imperfect communication does not vanish even when the agent is the white liar almost for sure.

Small Probabilities of Honesty and White Lie

Now we consider the case where μ is small, that is, the agent is biased with a high probability. The following result shows that if μ is sufficiently close to 0, for at least 79.1% of bias values, the principal is better off with a possibly honest agent than with a potential white liar.

Proposition 6 *Whenever* $b \in \left(\frac{1}{2n(n+1)}, \frac{1}{2n(n+0.5)}\right) \cup \left(\frac{1}{2n^2}, \frac{1}{2n(n-1)}\right)$ *for some natural number* n*, there exists* $\underline{\mu}(b) > 0$ *such that if* $\mu \leq \underline{\mu}(b)$ *then the maximal utility the principal can achieve is weakly greater in the honesty model than in the white lie model.*

The following three lemmas establish the proof of the proposition.

Lemma 3 *Unless* $b = \frac{1}{2n(n-1)}$*, for* μ *sufficiently close to* 0*, any equilibrium outcome in the white lie model in which* $\theta_0^0 = 0$ *(that is,* $n_0 = 0$*) or the white liar induces his more preferred policy on* $[0, \theta_0^0]$ *(that is,* $n_0 = \infty$*) can be replicated in the honesty model.*

Proof. See Appendix. ■

The intuition behind this result is as follows. For μ small, equilibrium is Type I in the honesty model. Then, the honesty model and the white lie model (for the cases where $n_0 = 0$ or $n_0 = \infty$) share equilibrium conditions other than MB for the former and NAW for the latter. As μ tends to 0, MB becomes negligible and thus imposes no restriction on equilibrium outcome. To the contrary, NAW is independent of μ and thus still binds. Consequently, any equilibrium outcome in the white lie model with the same structure as that of type I equilibrium is supported as an equilibrium outcome in the honesty model.

Lemma 4 *For* $b \in \left(\frac{1}{2n(n+1)}, \frac{1}{2n(n+0.5)}\right)$*, if* μ *is sufficiently small, there is no equilibrium in the white lie model in which there are* n *partition elements in the biased agent's strategy and the white liar separates from the biased agent on a positive measure of states, that is,* $\theta_0^0 = 0$ *in any equilibrium with* n *partition elements above* θ_0^0 *in the white lie model.*

To see this, suppose $\theta_0^0 > 0$ and let y_1 be the lowest policy induced by the biased agent. By the incentive compatibility of the biased agent, $\theta_0^0 = \frac{2n_0}{2n_0+1} y_1 \leq b$. This condition does not hold for any $n_0 \geq 1$ if and only if $y_1 > 1.5b$. When μ is sufficiently close to 0, the strategies of the biased sender and the principal are close to those of CS, and thus $y_1 \simeq 1/(2n) + b(1-n)$ (see Crawford and Sobel (1982), Section 4). Applying this result to $y_1 > 1.5b$, we get the condition in the lemma.

Lemma 5 *For* $b \in \left(\frac{1}{2n^2}, \frac{1}{2n(n-1)}\right)$*, if* μ *is sufficiently small, for any equilibrium in the white lie model, there exists a corresponding equilibrium in the honesty model which yields a weakly better utility to the principal.*

If $b \in \left(\frac{1}{2n^2}, \frac{1}{2n(n-1)}\right)$, for μ sufficiently small, then the lowest policy induced by the biased agent, y_1, is smaller than b. Then, the biased agent has no incentive to deviate to below θ_0^0, even though the principal perfectly trusts messages below θ_0^0. Then consider a type I strategy profile that inherits all the properties of the original white lie equilibrium except that the principal follows the agent's recommendation whenever the message is below θ_0^0 (that is, set $m_0 = \theta_0^0$). This is obviously an equilibrium in the honesty model and yields at least as much utility to the principal as the original white lie equilibrium.[11]

The following example shows that the principal can be *strictly* better off in the honesty model than in the white lie model.

Example 5 *Consider the case where* $b \in \left(\frac{1}{4}, \frac{1}{3}\right)$. *When* μ *is sufficiently close to* 0, *by Lemma 4, the babbling equilibrium is a unique equilibrium in the white lie model. To the contrary, in the honesty model, there exists an equilibrium in which the biased agent does not send messages below* $m_0 \in (0, 2b - 0.5)$. *The latter yields a strictly greater utility to the principal than the former.*

The following proposition is a partial converse to Proposition 6.

Proposition 7 *For* μ *sufficiently close to 0, if* $b \in \left(\frac{1}{2n(n+0.5)}, \frac{1}{2n^2}\right)$ *for some natural number* n, *there is an equilibrium in the white lie model whose outcome cannot be replicated in the honesty model. Also, for a fixed (small)* μ *there exists a* $\bar{b}_n(\mu)$ *such that for* $b \in [\frac{1}{2n(n+1)}, \bar{b}_n(\mu)]$ *there exists an equilibrium in the white model whose outcome cannot be replicated in the honesty model.*

Proof. See Appendix. ■

The two cases have different reasons. For μ positive, if b is slightly greater than one of CS critical values, the number of partition elements increases, but such increase is faster in the white lie model than in the honesty model. Therefore, there may be $(n+1)$-partition-element (above θ_0^0) equilibrium in the white lie model, while the maximal number of partition elements is still n in the honesty model.

For $b \in \left(\frac{1}{2n(n+0.5)}, \frac{1}{2n^2}\right)$, it is because there is an equilibrium with a finite number of partition elements below $\theta_0^0 > 0$ whose outcome cannot be replicated in the honesty model. To see this, notice that $\theta_0^0 = \frac{2n_0}{2n_0+1} y_1$ in an equilibrium with n_0 positive and finite. For this to be an equilibrium, the biased agent must prefer to induce y_1 to $\frac{2n_0-1}{2n_0+1} y_1$ at state 0. If we try to replicate the equilibrium outcome in the honesty model by setting $m_0 = \theta_0^0$, the biased agent must now prefer y_1 to $\theta_0^0 = \frac{2n_0}{2n_0+1} y_1 > \frac{2n_0-1}{2n_0+1} y_1$, which cannot be the case when $b \in \left(\frac{1}{2n(n+0.5)}, \frac{1}{2n^2}\right)$ and n_0 is the maximal number.

Proposition 7 does not establish that when b satisfies conditions in the proposition white lie necessarily yields greater utility to the principal than honesty. We show for the special case of $n = 1$ that white lie may or may not be preferable to honesty in such conditions and explain its underlying reason. Numerical analysis suggests that the considerations for the $n = 1$ yield the same conclusions for the case where $n > 1$, though we are unable to prove the result analytically.

---

[11] If $n_0 < \infty$ ($n_0 = \infty$), then the new equilibrium yields a strictly greater (the same) utility to the principal.

Consider the case where $b \in \left(\frac{1}{3}, \frac{1}{2}\right)$. Given μ and b, let $U_W(\mu, b)$ and $U_H(\mu, b)$ be the maximal principal utilities in the white lie model and in the honesty model, respectively. When μ is sufficiently close to 0, there can exist only one partition element above θ_0^0 and m_0 in each model. Therefore,

$$U_W(\mu, b) \quad = \quad -\mu\left[\sum_{k=1}^{n_0}\int_{\frac{(k-1)}{n_0}\theta_0^0}^{\frac{k}{n_0}\theta_0^0}\left(\theta - \frac{(2k-1)}{2n}\theta_0^0\right)^2 d\theta + \int_{\theta_0^0}^{1}(\theta - y_1)^2 d\theta\right] - (1-\mu)\int_0^1(\theta - y_1)^2 d\theta$$

where n_0 is the maximal number of partition elements below θ_0^0. Also,

$$U_H(\mu, b) \quad = \quad -\mu\int_{m_0}^{1}(\theta - y_1)^2 d\theta - (1-\mu)\int_0^1(\theta - y_1)^2 d\theta$$

where m_0 is the largest value subject to the incentive compatibility of the biased agent. Then,[12]

$$\lim_{\mu\to 0}\frac{\partial U_W(\mu, b)}{\partial\mu} \quad = \quad -\left[n_0\int_0^{\frac{\theta_0^0}{n_0}}\left(\theta - \frac{\theta_0^0}{2n}\right)^2 d\theta + \int_{\theta_0^0}^{1}(\theta - y_1)^2\, d\theta\right] - \int_0^1(\theta - y_1)^2\, d\theta$$

$$\lim_{\mu\to 0}\frac{\partial U_H(\mu, b)}{\partial\mu} \quad = \quad -\int_{2b-y_1}^{1}(\theta - y_1)^2\, d\theta - \int_0^1(\theta - y_1)^2\, d\theta.$$

As μ tends to 0, y_1 converges to $\frac{1}{2}$, n_0 to the largest integer that is weakly smaller than $\frac{b}{1-2b}$, θ_0^0 to $\frac{n_0}{2n_0+1}$, and m_0 to $2b - y_1$. Hence,

$$\lim_{\mu\to 0}\frac{\partial U_W(\mu, b)}{\partial\mu} - \lim_{\mu\to 0}\frac{\partial U_H(\mu, b)}{\partial\mu} = -n_0\int_0^{\frac{\theta_0^0}{n_0}}\left(\theta - \frac{\theta_0^0}{2n_0}\right)^2 d\theta + \int_{2b-y_1}^{\theta_0^0}(\theta - y_1)^2\, d\theta. \tag{3}$$

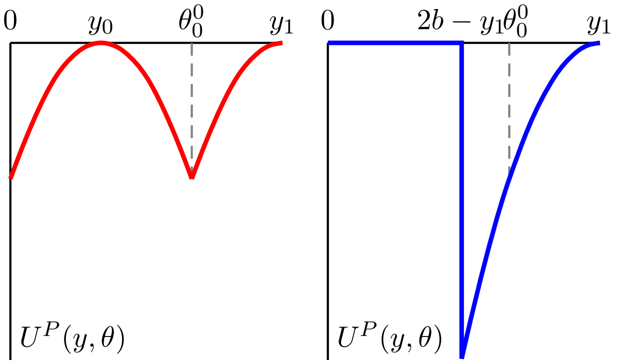

Figure 6: The principal's utilities in the white lie model (left) and in the honesty model (right) when $b \in \left(\frac{1}{2n(n+0.5)}, \frac{1}{2n(n+0.2)}\right)$, the agent is either honest or a white liar, and μ is sufficiently small.

---

[12] The terms involving $\frac{\partial y_1}{\partial\mu}$ vanish as $\mu \to 0$.

The two terms in Equation (3) show that when μ is sufficiently close to 0 the difference between equilibrium welfare in the two models arises from the outcomes when the agent is not biased and the state is in $[0, \theta_0^0]$. In the honesty model, when the agent is honest, perfect communication occurs on $[0, 2b - y_1]$ and y_1 is taken on $[2b - y_1, \theta_0^0]$ (see the right panel of Figure 6). In the white lie model, a finite number of actions that are uniformly distributed are taken on $[0, \theta_0^0]$ (see the left panel of Figure 6). The first term represents the losses in the white lie model due to imperfect communication between the white liar and the principal, while the second term represents the excessive losses in the honesty model due to the honest agent's inflexibility.[13] As shown in Figure 6, the latter can be relatively large because of the concavity of quadratic utility function.[14]

Figure 7 shows when white lie is preferable to honesty for μ sufficiently close to 0. Each kink occurs when n_0 jumps. As we increase b, the maximum n_0 is locally constant and consequently, the maximum welfare in the white lie model is fixed. However, the maximum welfare in the honesty model increases, as higher b allows for a larger region of perfect communication with the honest agent. Thus, the negative welfare effect in the honesty model arising from the quadratic utility loss shrinks as we increase b. For some ranges of b, this allows for a higher maximum welfare under the honesty model.

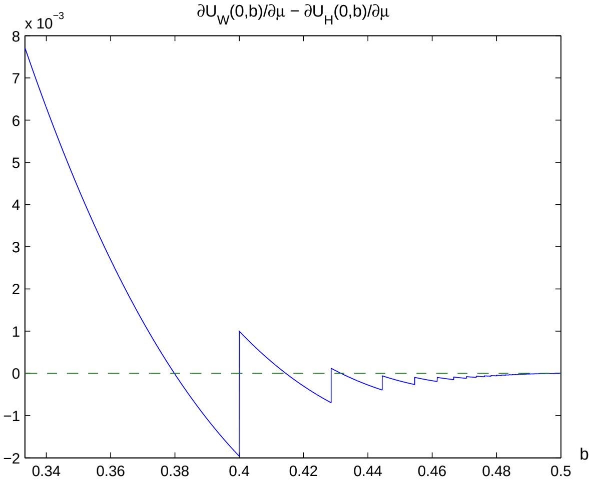

Figure 7: Welfare comparison for $b \in (1/3, 1/2)$ when μ is sufficiently small.

Intermediate Cases

For intermediate values of μ, we explain the underlying reasons why the principal can be better off in the honesty model than in the white lie model. We support our arguments by providing some

---

[13] Perfect communication must occur whenever the principal knows that the agent is honest. Therefore, it is harder in the honesty model to deter the biased agent from deviating to lower messages than in the white lie model.

[14] For the case of $n = 1$, the welfare maximizing equilibrium strategy profiles satisfy convexity and monotonicity. Therefore, the restriction to such equilibria in the welfare comparison is without loss of generality.

numerical examples.[15]

In the intermediate case, honesty mainly improves the welfare of the principal by reducing the excessive losses in communication due to strategic considerations of players. To see this, recall that in the original CS game, partition element size increases for higher θ. This is true in the white model as well, though the increase in partition element size is decreasing in μ. In the honesty model, however, this is not necessarily true. The inequality property of MB does not require that partition element sizes increase in the honesty model as they must in the white lie model. Partition elements can be adjusted (as long as all MB constraints are satisfied) so that higher partition elements need not be larger than lower partition elements. This has the consequence of allowing better communication under honesty where the welfare losses were largest under white lie.

The first consequence of this is that such a freedom in the honesty model allows for more uniform partition element sizes by manipulating the partitioning of the message space. While this may reduce the expected welfare conditional on the agent being honest, it also makes the biased agent's partitioning more uniform, which is directly beneficial to the principal due to the concavity of quadratic utility function. This effect is highlighted in Example 6.

Example 6 *Suppose* $b = 0.15$ *and* $\mu = 0.1$*. In the white lie model, there exist 2-partition-element equilibria for each* $n_0 \in \mathbf{N} \cup \infty$ *and the equilibrium yielding the maximal principal utility, -0.0392, is characterized by*

$$\theta^0 = \{0.126, 0.371, 1\}, \theta^b = \{0, 0.221, 1\}, y = \{0.126, 0.617\}, \textit{and } n_0 = \infty.$$

The expected loss from communication with the white liar and biased agent are -0.0029 *and* -0.0363*, respectively. In the honesty model, there exists a 2-partition-element equilibrium that is characterized by*

$$m = \{0.138, 0.575, 1\}, \theta^b = \{0, 0.2478, 1\}, y = \{0.162, 0.634\}.$$

In this equilibrium, the principal achieves utility -0.0374 *where the loss from communication with the honest and biased agent are* -0.004 *and* -0.0335*, respectively. Note that the principal is worse off in the honest equilibrium if the agent is not biased, though is better off when the agent is biased.*

The second consequence of allowing for more uniform partition element sizes is that as μ increases, the maximal number of partition elements increases faster in the honesty model than in the white lie model. In the white lie model, the increasing partition element sizes for both the white liar and biased agent inhibit the ability for larger number of partition elements. In the honesty model, smaller mass from the honest agent may be placed on higher messages, raising the higher actions, shrinking the largest partition elements, and allowing for greater number of intermediate

---

[15] The analytical difficulty for intermediate cases lies in characterizing the set of equilibria in the honesty model. The set of equilibria is very large (due to the inequalities of MB), not convex (IC is necessary only when $m_0 > 0$), and possibly not closed (due to $m_0 < m_1 < .. < m_n$). Furthermore, we cannot fix the dimension of control variables, because the maximal number of partition elements is not analytically available.

equilibrium actions. This is illustrated in Example 7.

Example 7 *Consider the case where* $b = 1/12$ *and* $\mu = 1/4$. *The largest number of partition elements (above* θ_0^0*) in the white lie model is 3 and any* $n_0 \in \mathbf{N} \cup \infty$ *is possible, with* $n_0 = \infty$ *yielding the maximal principal utility. There exists a 5-partition equilibrium in the honesty model that is characterized by*

$$m = \{0.083, 0.228, 0.539, 0.756, 0.796, 1\},$$
$$\theta^b = \{0, 0.083, 0.263, 0.387, 0.565, 1\},$$
$$y = \{0.083, 0.250, 0.443, 0.498, 0.798\}.$$

Note that the partition element sizes for the biased agent are $0.083, 0.180, 0.124, 0.178$ *and* 0.435*, respectively. If we consider any 5-partition equilibrium in the white lie model, then it must be that the third and fourth partition elements increase in size by* $4b(1-\mu) = 1/4$ *from the preceding element. Relative to honesty, this is a constraining factor in allowing partition equilibria with greater number of elements as* μ *increases. The maximal principal utility in the white lie model is* -0.0198*, while it is* -0.0147 *in the honest model.*

6 Conclusion

In the paper, we explore the effects of honesty and white lie on communication. They are highlighted by new equilibrium conditions, mass balance condition in the honesty model and no arbitrage condition for the white liar in the white lie model. By doing so, we could compare the behavioral consequences of honesty and white lie, and conclude that the principal is often better off with a possibly honest agent than with a potential white liar.

Our results shed some light on the question on honesty versus white lie, which is relevant in many contexts but has not been answered in a formal model. In addition, our favorable conclusion to honesty can be understood as a micro-foundation for the preference for honesty in particular and behavioral character in general.

Another important contribution of this paper is identifying new channels through which commitment is valuable. We find that commitment simplifies communication protocol and thus reduces the loss due to strategic considerations. Also, we show that commitment has the effect of enriching language used in communication, which allows more possibilities in a strategic environment.

Appendix: Omitted Proofs

Proof of Lemma 1. Let $z^+(\theta', b) = \lim_{\theta \to \theta'+} z(\theta, b)$, $z^-(\theta', b) = \lim_{\theta \to \theta'-} z(\theta, b)$.

($\Rightarrow$) (i) $z(\cdot, b)$ is nondecreasing.

Suppose $z(\cdot, b)$ is strictly decreasing on (θ^1, θ^2) with $\theta^1 < \theta^2$. For $z(\cdot, b)$ to be incentive compatible, $U_S(z(\theta^1, b), \theta^1, b) \geq U_S(z(\theta, b), \theta^1, b)$ and $U_S(z(\theta^2, b), \theta^2, b) \geq U_S(z(\theta, b), \theta^2, b)$ for all $\theta \in \Omega$. Since $z(\cdot, b)$ is strictly decreasing on (θ^1, θ^2), $z(\cdot, b)$ is continuous except countably many points.

Pick some $\theta \in (\theta^1, \theta^2)$ at which $z(\cdot, b)$ is continuous. If $\partial U_S(z(\theta,b),\theta,b)/\partial y \neq 0$, then the biased agent has a profitable deviation (If $\partial U_S(z(\theta,b),\theta,b)/\partial y > (<)0$ then The biased agent deviates to $\theta' < (>)\theta$). Hence $\partial U_S(z(\theta,b),\theta,b)/\partial y = 0$ almost everywhere on (θ^1, θ^2). But this is a contradiction to the single-crossing property of the utility function.

(ii) $V^S(\cdot, b)$ is continuous.

Suppose $V^S(\cdot, b)$ is not continuous at $\theta' \in (0,1)$. Then $z(\cdot, b)$ cannot be continuous at θ'. Since $z(\cdot, b)$ is nondecreasing, this means $z(\cdot, b)$ has jump at θ'. Pick θ^+ and θ^- sufficiently close to θ' so that $z^+(\theta', b) \leq z(\theta^+, b)$ and $z(\theta^-, b) \leq z^-(\theta', b)$. We have the following three cases: (1) $z(\theta^+, b) \leq y^S(\theta', b)$, (2) $z(\theta^-, b) < y^S(\theta', b) < z(\theta^+, b)$ and (3) $y^S(\theta', b) \leq z(\theta^-, b)$. In case (1), the biased agent has an incentive to deviate at θ^-, while he does at θ^+ in case (3). In case (3), no type has an incentive to deviate only when $\lim_{\theta \to \theta'+} V^S(\theta, b) = \lim_{\theta \to \theta'-} V^S(\theta, b)$. Hence $V^S(\cdot, b)$ is continuous.

(iii) If $z_1(\theta, b)$ exists, then $U_1^S(z(\theta,b),\theta,b) \cdot z_1(\theta, b) = 0$.

The differentiability of z at θ implies the differentiability of $V^S(\cdot, b)$ at θ, because $V^S(\theta, b) = U^A(z(\theta,b),\theta,b)$. By the Envelope theorem,

$$\frac{\partial U_S(z(\theta,b),\theta,b)}{\partial y}\frac{\partial z(\theta,b)}{\partial \theta} = 0$$

($\Leftarrow$) We want to show that $U^A(z(\theta'',b),\theta'',b) \geq U^A(z(\theta',b),\theta'',b)$ for all $\theta', \theta'' \in \Omega$.

$$\begin{aligned}
U^A(z(\theta'',b),\theta'',b) - U^A(z(\theta',b),\theta'',b) &= [U^A(z(\theta'',b),\theta'',b) - U^A(z(\theta',b),\theta',b)] \\
&\quad -[U^A(z(\theta',b),\theta'',b) - U^A(z(\theta',b),\theta',b)] \\
&= \int_{\theta'}^{\theta''} \frac{\partial V^S(\theta,b)}{\partial \theta} d\theta - \int_{\theta'}^{\theta''} U_2^S(z(\theta',b),\theta,b) d\theta \\
&= \int_{\theta'}^{\theta''} U_2^S(z(\theta,b),\theta,b) d\theta - \int_{\theta'}^{\theta''} U_2^S(z(\theta',b),\theta,b) d\theta \\
&= \int_{\theta'}^{\theta''} \left[\int_{z(\theta',b)}^{z(\theta,b)} U_{12}^S(z(t,b),\theta,b) dz \right] d\theta.
\end{aligned}$$

V^S is absolutely continuous via application of an Envelope theorem for this environment (see Milgrom and Segal (2002)).

If $\theta'' > \theta'$, then $z(\theta,b) \geq z(\theta',b), \forall \theta \in [\theta', \theta'']$, and so $U^A(z(\theta'',b),\theta'',b) - U^A(z(\theta',b),\theta'',b) \geq 0$. Similarly, if $\theta'' < \theta'$, then $z(\theta,b) \leq z(\theta',b), \forall \theta \in [\theta', \theta'']$, and again $U^A(z(\theta'',b),\theta'',b) - U^A(z(\theta',b),\theta'',b) \geq 0$. **Q.E.D.**

Proof of Lemma 2. ($\Leftarrow$) For each θ, consider the following probability measure.

$$r(\theta)=\begin{cases}\frac{m''-m'}{\theta''-\theta'}(\theta''-\theta)+m', & \text{with probability } \frac{\mu}{1-\mu}\frac{(m''-m')^2}{(\theta''-\theta')^2},\\ m\sim U[m',m''], & \text{with probability } 1-\frac{\mu}{1-\mu}\frac{(m''-m')^2}{(\theta''-\theta')^2}.\end{cases}$$

This probability measure is well-defined when $\mu(m''-m')^2\leq(1-\mu)(\theta''-\theta')^2$. Then for any $m\in[m',m'']$,

$$\begin{aligned}E_{\mu,r}[\theta|m] &= \frac{\mu\left(m''-m'\right)}{\mu\left(m''-m'\right)+\left(1-\mu\right)\left(\theta''-\theta'\right)}m\\ &\quad+\frac{\left(1-\mu\right)\left(\theta''-\theta'\right)}{\mu\left(m''-m'\right)+\left(1-\mu\right)\left(\theta''-\theta'\right)}\frac{\mu}{1-\mu}\frac{\left(m''-m'\right)^2}{\left(\theta''-\theta'\right)^2}\left(\theta''-\frac{\theta''-\theta'}{m''-m'}(m-m')\right)\\ &\quad+\frac{\left(1-\mu\right)\left(\theta''-\theta'\right)}{\mu\left(m''-m'\right)+\left(1-\mu\right)\left(\theta''-\theta'\right)}\left(1-\frac{\mu}{1-\mu}\frac{\left(m''-m'\right)^2}{\left(\theta''-\theta'\right)^2}\right)\frac{\theta''+\theta'}{2}\\ &= \overline{y}.\end{aligned}$$

($\Rightarrow$) Let $m^*\in[m',m'']$ be the value such that $m^*-m'=m''-m^*$. Let θ^* be the value such that

$$B(\mu,m',m^*,\theta^*,\theta'')=\overline{y}.$$

Such θ^* is well-defined if and only if $\mu\left(m''-m'\right)^2\leq\left(1-\mu\right)\left(\theta''-\theta'\right)^2$.

Suppose there exists a collection of probability measures, $\{r\left(\theta\right),\theta\in[\theta',\theta'']\}$, that satisfies the given property. Then by construction,

$$\int_{\theta'}^{\theta''}r\left(\theta\right)\left(\left[m',m^*\right]\right)d\theta\geq\theta''-\theta^*.$$

Therefore, θ^* must be well-defined, which is true only when $\mu\left(m''-m'\right)^2\leq\left(1-\mu\right)\left(\theta''-\theta'\right)^2$. **Q.E.D.**

Proof of Proposition 2: The proof is straightforward from the following claim (together with Lemma 1), which establishes that there can be no partition element below the region on which the agent induces his most preferred policy.

Lemma 6 *Fix $\theta_0\in(0,1-b/\mu)$ and $m_0=\theta_0+b/\mu$. There cannot exist $\theta\in[0,\theta_0)$ and $m\in[0,m_0)$ such that $B(\mu,m,m_0,\theta,\theta_0)=\theta_0+b$.*

Proof. Suppose not. Then for some θ and m,

$$\theta_0+b=\frac{\mu\left(m_0-m\right)}{\mu\left(m_0-m\right)+\left(1-\mu\right)\left(\theta_0-\theta\right)}\frac{m+m_0}{2}+\frac{\left(1-\mu\right)\left(\theta_0-\theta\right)}{\mu\left(m_0-m\right)+\left(1-\mu\right)\left(\theta_0-\theta\right)}\frac{\theta+\theta_0}{2}.$$

Re arranging terms,

$$\begin{aligned} & \theta_0^2 - 2\left(\mu m + (1-\mu)\,\theta - b\right)\theta_0 + 2b\left(b - \mu m - (1-\mu)\,\theta\right) - \frac{b^2}{\mu} + \mu m^2 + (1-\mu)\,\theta^2 \\ = \; & \left[\theta_0^2 - \left(\mu m + (1-\mu)\,\theta - b\right)\right]^2 + \mu\,(1-\mu)\left[(m-\theta)^2 - \left(\frac{b}{\mu}\right)^2\right] = 0. \end{aligned}$$

For the solution to exist, the right-hand side should be not more than 0 when $\theta_0 = \mu m + (1-\mu)\,\theta - b$. But in that case,

$$(m-\theta)^2 - \left(\frac{b}{\mu}\right)^2 = \frac{1}{\mu^2}\left[(\theta_0-\theta)^2 + 2b\,(\theta_0-\theta)\right] > 0.$$

Hence there cannot exist such θ and m. ■ **Q.E.D.**

Proof of Proposition 4: Take any equilibrium in the white lie model. Due to the flexibility of the white liar, the biased agent cannot induce his most preferred policy on a positive measure of states.

Now let Θ^* be the set of states on which the white liar induces actions that are not induced by the biased agent and let $\overline{\theta}$ be the supremum of Θ^*. $\overline{\theta}$ is less than or equal to any equilibrium action induced by the biased agent, otherwise there would be some state at which the biased agent would deviate to induce it. In addition, Θ^* is convex. This is because by part (i) of Lemma 1, the white liar induces a (weakly) lower action than $\overline{\theta}$ at any state $\theta < \overline{\theta}$.

All other arguments are straightforward from equilibrium necessary conditions. **Q.E.D.**

Proof of Lemma 3: Notice that all partition element sizes for the white liar are less than or equal to the partition element sizes for the biased agent except the first partition element, because $\theta_k^0 = \theta_k^b + b, \forall k = 1, .., n-1$. Therefore, given an equilibrium in the white lie model, MB trivially holds for all but the first partition element as long as $\mu \leq 1/2$.

Suppose that as $\mu \to 0$, $\theta_1^b \to \bar{\theta} > 0$. Then the proof is immediate as any equilibrium outcome in the white lie model $\{\theta^b, \theta^0, y\}$ can be replicated as an equilibrium in the honest model for μ sufficiently small.

Now consider the case in which $\theta_1^b \to 0$. First, suppose $\theta_0^0 = 0$. It is straightforward to show that $\theta_1^b \to \frac{1-2n(n-1)b}{n}$. Therefore, $\theta_1^b \to 0$ only when b is a critical CS value. Second, suppose $\theta_0^0 > 0$. Let $\theta_0^0 \to \bar{\theta}_0$ and consider the case where $\theta_1^b \to 0$. Then $\theta_n^b \to 2(nb - \bar{\theta}_0)(n-1) = 1$ so that $\bar{\theta}_0 = \frac{2(n-1)nb-1}{2(n-1)}$. Respecting $\theta_0^0 \leq y_2$ and noting that $\lim_{\mu\to 0} y_2 = \theta_2^b/2$ and $\lim_{\mu \to 0} \theta_2^b = \frac{1-2n(n-1)b}{n}$ it must be that $\frac{2n(n-1)b-1}{2(n-1)} \leq \frac{1-2n(n-1)b}{2n}$. Combined with the fact that $\bar{\theta}_0 \geq 0$, this implies that, $b = \frac{1}{2n(n-1)}$. **Q.E.D.**

Proof of Proposition 7: We first prove the first half of the statement. Suppose μ is arbitrarily close to 0 and $\frac{2n_0+3}{2n_0+2}b < y_1 \leq \frac{2n_0+1}{2n_0}b$ for some natural number n_0. Then there exists an equilibrium in which there are n_0 partition elements below θ_0^0. In such an equilibrium, $\theta_0^0 = \frac{2n_0}{2n_0+1}y_1$. This equilibrium outcome cannot be replicated in the honesty model, because $m_0 \leq 2b - y_1$ due to the incentive compatibility of the biased agent, while $\theta_0^0 > 2b - y_1$ by construction for the case with $n_0 = 1$). Since this argument holds for any natural number n_0, there is an equilibrium with $\theta_0^0 > 0$ whose outcome cannot be replicated in the honesty model when $b < y \leq 1.5b$, whose condition coincides with $b \in \left(\frac{1}{2n(n+0.5)}, \frac{1}{2n^2}\right)$.

Next, we prove the case where b is greater but sufficiently close to $\frac{1}{2n(n-1)}$. For small but positive μ, there exists an n-partition equilibrium in the white lie model with θ_1^b small but positive. MB requires that $\mu(\theta_1^b + b - \theta_0^0)^2 \leq (1-\mu)(\theta_1^b)^2$. We show that this condition does not hold when $b = \frac{1}{2n(n-1)}$. By continuity, the condition does not hold for b is greater but sufficiently close to $\frac{1}{2n(n-1)}$ as well.

Fix μ close to 0. From Equation (1), we know that θ_1^b satisfies

$$1 = n\theta_1^b + (n-1)\left[2\left(1-\mu\right)nb + \mu b - \frac{\mu\left(1-\mu\right)b^2}{\theta_1^b + \mu b}\right].$$

Since $b = \frac{1}{2n(n-1)}$,

$$1 = n\theta_1^b + (1-\mu) + \frac{\mu}{2n} - \frac{\mu\left(1-\mu\right)}{2n\left(n-1\right)\theta_1^b + \mu}\frac{1}{2n}.$$

Arranging terms,

$$\mu\left[2n\left(n-1\right)\left(2n-1\right)\theta_1^b + \mu\left(2n-1\right) + \left(1-\mu\right) - 2n^2\theta_1^b\right] = 4n^3\left(n-1\right)\left(\theta_1^b\right)^2.$$

When μ is close to 0, θ_1^b is also close to 0 because b is near a critical value, and

$$\mu \approx 4n^3\left(n-1\right)\left(\theta_1^b\right)^2.$$

Therefore,

$$\begin{aligned}\mu(\theta_1^b + b - \theta_0^0)^2 &\approx 4n^3\left(n-1\right)\left(\theta_1^b\right)^2 \frac{1}{4n^2\left(n-1\right)^2} = \frac{n}{n-1}\left(\theta_1^b\right)^2 \\ &> \left(\theta_1^b\right)^2 > (1-\mu)(\theta_1^b)^2.\end{aligned}$$

Q.E.D.

References

Benabou, R. and Laroque, G. (1992), 'Using privileged information to manipulate markets: Insiders, gurus, and credibility', *Quarterly Journal of Economics* **107**(3), 921–958.

Blume, A., Board, O. J. and Kawamura, K. (2007), ‘Noisy talk’, *Theoretical Economics* **2**(4), 395–440.

Chen, Y., Kartik, N. and Sobel, J. (2008), ‘Selecting cheap-talk equilibria’, *Econometrica* **76**(1), 117–136.

Crawford, V. P. and Sobel, J. (1982), ‘Strategic information transmission’, *Econometrica* **50**(6), 1431–1451.

Dessein, W. (2002), ‘Authority and communication in organizations’, *Review of Economic Studies* **69**, 811–838.

Li, M. and Madarasz, K. (2007), ‘When mandatory disclosure hurts: Agent advice and conflicting interests’, *Journal of Economic Theory* **139**, 47–74.

Milgrom, P. and Segal, I. (2002), ‘Envelope theorems for arbitrary choice sets’, *Econometrica* **70**, 583–601.

Morgan, J. and Stocken, P. C. (2003), ‘An analysis of stock recommendations’, *RAND Journal of Economics* **34**, 183–203.

Morris, S. (2001), ‘Political correctness’, *Journal of Political Economy* **109**, 231–265.

Olszewski, W. (2004), ‘Informal communication’, *Journal of Economic Theory* **117**, 180–200.

Sobel, J. (1985), ‘A theory of credibility’, *Review of Economic Studies* **52**, 557–573.