Well Informed Intermediaries in Strategic Communication*

Wei Li†

July, 2008

Abstract

This paper studies how a sender with private information can influence the decisionmaker through well informed intermediaries. Both the sender and the intermediary may be independently objective or biased: with the objective type assumed to pass on the most accurate information while the biased type wanting to push a particular agenda but also to appear objective. Although using one’s own information is a sign of objectivity, the biased intermediary selectively incorporates the sender’s information to push his agenda, and his truth-telling incentives always decrease in those of the biased sender’s. Hence measures raising the sender’s reputation cost worsen the intermediary’s distortion and may make the decisionmaker strictly worse off. In contrast, the biased sender’s and the intermediary’s truth-telling incentives are strategic complements if they report simultaneously.

JEL classification: C70, D82, M31, M38

Keywords: strategic communication through intermediaries, indirect marketing, reputational concerns.

*I am grateful for the insightful comments from Glenn Ellison, Hao Li, R. Preston McAfee, Marco Ottaviani, Edward Schlee, Jean Tirole, and the seminar participants at ASU, UCR, Fuqua School of Business, Duke University, John Hopkins University, Marshall School of Business, USC, the 2007 Duke/Northwestern/Texas IO conference, and the 83th Western Economic Association Annual Conference.

†University of California, Riverside. Email: wei.li@ucr.edu. Phone: 951-827-7809. Fax: 951-827-5685.
1 Introduction

Recommendations from well informed experts are often important to the final users of such information. In health care, patients often follow the prescriptions and other advice from their physicians. Also, the medical community learns about the efficacy of drugs and procedures through published studies and clinical trials. In electronic commerce, online customer opinions or reviews have been shown to boost sales significantly.\(^1\) Because of the importance of experts’ recommendations, it is increasingly often for businesses to attempt to influence their customers through influencing the experts’ views about their products. Some businesses even pay experts or bloggers to promote their image or products while pretending to be independent reviewers.\(^2\) In the medical and health industry, serious questions have been raised concerning pharmaceutical companies who promote their drugs through physicians and medical researchers.\(^3\) These companies exert influence through providing physicians with the company’s own information about their products (“detailing”) and other gifts; through funding research and giving perks such as paid consultancy positions; and through “ghostwriting” of journal articles where the purported academic authors have done little of the actual research.\(^4\)

These influence activities have become more prevalent in recent years. It is estimated that drug companies spend approximately $19 billion annually for marketing to doctors.\(^5\) In particular, the amount of money spent “detailing” physicians has increased from $3.0 to $4.8 billion from 1996 to 2000.\(^6\) Nearly 75 percent of physicians in a national poll said the information they received from

---

\(^1\) In a 2007 Customer Engagement Survey conducted by the global marketing research firm ACNielsen, online consumer opinions represented the third most trusted form of advertising, after word-of-mouth opinions and newspaper advertisements. Another market research firm, E-consultancy, asked online retailers about the effects of adding customer-generated reviews and ratings: 77 percent said site traffic increased, and 42 percent reported a rise in the amount of money spent.


\(^4\) One recent example concerns an Annals of Internal Medicine article on Merck’s “Advantage” trial of Vioxx, which omitted some trial participants’ deaths. The article’s first author Jeffrey Lisse, a rheumatologist at the University of Arizona, said that Merck actually wrote the report, and that “Merck designed the trial, paid for the trial, ran the trial.....Merck came to me after the study was completed.” The New York Times, “Evidence in Vioxx Suits Shows Intervention by Merck Officials”, April 24, 2005.

\(^5\) For details, please see “Health industry practices that create conflicts of interest: a policy proposal for academic medical centers,” published in the Journal of the American Medical Association (Brennan, Rothman, Blank, Blumenthal, Chimonas, Cohen, Golden, Kassirer, Kimball, Naughton, and Smelser 2006).

\(^6\) See IMS Health Inc. and Competitive Media Reporting, as reported by Kaiser Family Foundation, 2001.
pharmaceutical representatives was “very” useful (15 percent) or “somewhat” useful (59 percent).\footnote{More than half (55 percent) of a group of “high-prescribing” doctors surveyed by the industry data tracking group ImpactRx said that drug representatives serve as their primary source of information about newly approved drugs. See “Getting Doctors to Say Yes to Drugs: The Cost and Quality Impact of Drug Company Marketing to Physicians” by the BlueCross BlueShield Association. See http://www.bcbs.com/betterknowledge/cost/getting-doctors-to-say-yes.html.}

Many studies also indicate that these influence activities are effective in changing physicians’ prescription behaviors, and consequently patients’ welfare (Avorn, Chen, and Hartley 1982, Watkins, Moore, Harvey, Carthy, Robinson, and Brawn 2003).\footnote{Another example concerns General Motors Corp., which found itself spending $52 million in 2001 for prescriptions that physicians wrote for Prilosec, even though a subsequent analysis found that 91 percent of the patients receiving prescriptions had no prior diagnosis of the problem. A recent analysis of National Ambulatory Medical Care Survey data found that since 1991, “physicians are increasingly turning to expensive, broad-spectrum agents, even when there is little clinical rationale for their use” compared to simpler, generic antibiotics.}

This paper presents a simple model of strategic communication through a well informed intermediary and examines its effect on the final users of such information. This differs from many existing papers analyzing how a sender influences the decisionmaker directly by manipulating the information he sends (Crawford and Sobel 1982, Austen-Smith 1990, Dewatripont and Tirole 1999, Morris 2001, Ottaviani and Sorensen 2006, among others). Consider the medical example above, where a pharmaceutical company may promote its drugs through physicians or medical researchers. How are the physicians influenced by the pharmaceutical company, especially if their information disagree? How do the pharmaceutical company’s truth-telling incentives interact with those of the physician’s? What is the net impact on the patients (or the broader medical community)? This paper addresses these questions by investigating the strategic interactions of the sender and the intermediary. It also applies these insights to study the effectiveness of policy measures aiming at improving reporting accuracy, which has implications for disclosure laws and professional ethical rules.

In the main model, a sender receives a private signal about the state of the world and sends a message to an intermediary who also has an independent, private signal. Because this paper focuses on a well informed intermediary who has expertise such as physicians and researchers; or has experience in a market for credence goods, the intermediary’s signal is assumed to be more accurate than that of the sender’s.\footnote{In a companion piece, Li (2007b) considers the case when the intermediary has little information of his own and acts as a pure intermediary. The strategic interactions between the intermediary and the sender, as well as policy implications, are very different due to the lack of information aggregation considerations. See further discussions in Section 5.}

The intermediary in turn sends a recommendation to the decisionmaker, who takes an action based on what she hears.\footnote{Throughout this paper, the sender and the intermediary are male and the decisionmaker is female.} Afterwards, the true state becomes observable, and the decisionmaker forms
her opinion of how truthful the agents (both the sender and the intermediary) have been. The sender and the intermediary may be independently one of two types: objective or biased. An objective agent is assumed to pass on the most accurate information he has. But a biased agent wants the decisionmaker to take a particular action and he has reputational concerns: he wants to appear objective in the eyes of the decisionmaker. Thus a biased agent faces a tradeoff between reporting truthfully to appear objective and distorting to induce the decisionmaker to take an action more favorable to his bias.

The first result of this paper is that even a relatively uninformative message from the sender may have a strong impact on a biased intermediary whose own signal does not support his bias. Specifically, the biased intermediary is more likely to lie against his own, superior information if the sender’s message favors his bias than if it does not. Moreover, if the sender’s message becomes more truthful, the biased intermediary distorts more if he is sufficiently concerned about his reputation. This suggests that the decisionmaker may value experts who base their recommendations solely on their own signal.

The key to understand this result is that the biased intermediary’s truth-telling incentives always decrease in those of the sender’s. On one hand, the sender’s message affects the biased intermediary’s net benefit from distortion, which is the change in the decisionmaker’s action induced by his recommendations. If the sender’s message becomes more truthful, then the biased intermediary’s recommendation, whether it supports or opposes his bias, becomes more credible. Thus the decisionmaker’s action is more responsive to the intermediary’s recommendation, increasing his incentive to lie. On the other hand, the biased intermediary pays a reputation cost if he distorts because his recommendation is more likely to be wrong, which suggests that he is influenced by the less informative message from the sender. If the sender’s message becomes more truthful, it has two effects on the biased intermediary’s net reputation cost of distortion. First, it changes the intermediary’s own estimate of the true state through an information aggregation effect. If the message supports his bias, then the intermediary knows that if he distorts, his recommendation is less likely to be wrong, and thus he is less likely to suffer a severe loss in reputation from distortion. Second, a more informative message changes how biased intermediary’s reputation depends on his recommendations through a reputation sensitivity effect. Since the objective intermediary always reports his signal, independence from the sender’s influence is the sign of objectivity: the intermediary cannot blame any mistake on being misled by the wrong source. If the sender’s message becomes more informative, the decisionmaker assigns the intermediary less credit for
a correct recommendation because he may have followed the more accurate message; and less blame for a wrong recommendation because any mistake is more likely due to a wrong signal. Because the biased intermediary’s net reputation cost from distortion falls and his net benefit from distortion increases, he distorts more if the sender’s message is more truthful.

A biased sender, however, never reports truthfully if he can influence the intermediary. Moreover, when the intermediary has high reputational concerns, his own reputational cost from distortion decreases in the intermediary’s truth telling, because of a similar reputation sensitivity effect. In this case, the biased agents’ truth-telling incentives become strategic substitutes: if one reports more truthfully, the other distorts more. In particular, a biased sender may distort completely, regardless of how high his reputational concerns are, if the biased intermediary is sufficiently concerned about his reputation.

The decisionmaker may consider policy measures to increase the reputational cost of the biased agents to improve truth telling. The second result of this paper is that such measures may not benefit the decisionmaker. Paradoxically, making it more costly for the sender to lie, for instance by strengthening regulations over the pharmaceutical industry, may make the decisionmaker strictly worse off. This occurs when both biased agents have low levels of reputational concerns, in which case the sender’s message is not credible and unlikely to be followed by the intermediary. If the biased sender reports more truthfully under the new policy, the biased intermediary lies more against his own signal. This means that the decisionmaker may receive a recommendation against both agents’ true signals with a higher probability, and thus she makes worse decisions. She should target the intermediary instead, for instance by strengthening medical board review process and monitoring disclosure of industry ties of the physicians and researchers. The resulting gain in truth telling from the intermediary, whose signal is more accurate, outweighs any indirect effect on the sender (even if it is negative). If the intermediary has high levels of reputational concerns, however, policy measures increasing either biased agent’s reputation cost are relatively ineffective, both because the intermediary reports truthfully with a high probability, and because the biased agent’s truth-telling incentives are strategic substitutes.

In comparison, simultaneous reporting is considered where the decisionmaker receives two messages, one from each agent, before taking an action. Here each biased agent pays his own reputation cost because he is evaluated based on his message and the later observed true state. Therefore the presence of another message only affects a biased agent’s (expected) net benefit of distortion, which is shown
to decrease in the other agent’s truth telling. Intuitively, if one biased agent reports more truthfully, it increases the probability that the other agent’s distorted message is contradicted, in which case decisionmaker is less likely to change her action than if there is a concurring message. Overall, distortion becomes less effective while one’s reputation cost is unaffected, thus each biased agent lies less. The third result of this paper is that biased agents’ truth-telling probabilities are strategic complements under simultaneous reporting. Thus the decisionmaker may want to encourage simultaneous reporting, especially if both agents have high levels of reputational concerns. In this case, any policy measure targeting either the industry or the intermediary has a strictly positive effect.

The present paper focuses on how a sender can influence an intermediary by providing information and altering the intermediary’s confidence in his own signal, and thus the decisionmaker’s action. Durbin and Iyer (2008) consider the case where intermediaries (advisors) may be bribed by an uninformed and biased third party to support its bias. They show that if the advisors have reputational concerns, a bribe may be necessary for the advisor to report his true signal if it happens to favor the biased third party. In a similar vertical structure but with a different focus, Inderst and Ottaviani (2008) study the incentive problems faced by an intermediary (such as a sales agent) if he has to perform two tasks: one for the seller and the other for the buyer. They study how the firms design the intermediary’s compensation schemes given this inherent conflict of interest.

From a social network perspective, this paper is also related to DeMarzo, Vayanos, and Zwiebel (2003), who show that the influence of one’s action on others depend not only on his information accuracy, but also on his position in a given social network. Their work takes an orthogonal approach: they focus on richer sets of social networks and assume that the agents report truthfully, but fail to account for possible repetitions in the information that reaches them. Instead, the current paper focuses on a simple vertical structure and considers strategic agents who make rational inference of any information they receive given the possible source(s) and the bias involved.

In the rest of the paper, Section 2 sets up the basic model and Section 3 analyzes the equilibria behavior of the biased agents. It also considers how the decisionmaker can encourage truthful reporting. Section 4 considers the case of simultaneous reporting by the agents; and the ex ante preference of a biased sender. Section 5 discusses several main assumptions and concludes. All the proofs are contained in the Appendix.
2 The Model

There are two players in this game, agent $A$ and $B$, and there is a decisionmaker $C$. The state of the world is binary: $\eta \in \{0, 1\}$. Each state occurs with equal probability. The game proceeds in three stages: information transmission, decision making, and evaluation. In the information transmission stage, agent $A$ first observes a private signal $s_A \in \{0, 1\}$, which is equal to the true state with probability $p_A > 0.5$; otherwise it is wrong. He then sends a message $m_A \in \{0, 1\}$ to an intermediary, agent $B$. Agent $B$ observes a private signal $s_B \in \{0, 1\}$ of his own, which is equal to the true state with probability $p_B > p_A$. Agent $B$ then sends a message $m_B \in \{0, 1\}$ to the decisionmaker. Agent $A$ and $B$’s signals are independent conditional on the state. In the decision making stage, $C$ chooses an action $a \in \mathbb{R}$ given message $m_B$. In the evaluation stage, $C$ first observes the true state $\eta$ and then forms posterior beliefs about each agent’s type, to be described next. In all three stages, agent $B$ and decisionmaker $C$ only observe the message sent directly to him (her).

The decisionmaker’s payoff is represented by the quadratic loss function $-(a - \eta)^2$. Her optimal action is thus to choose $a$ equal to the probability she attaches to $\eta = 1$. An agent may be either objective (type $o$) or biased (type $b$). Each agent’s type is independently drawn from $\{o, b\}$: $\Pr(A = o) = \theta_A$, and $\Pr(A = b) = 1 - \theta_A$, with $\theta_A$ referred to as $A$’s prior objectivity. Similarly, agent $B$ is objective with probability $\theta_B$. An objective agent is assumed to report the most accurate information he has. A biased agent always wants action $a = 1$ taken, regardless of the true state. However, he also wants to be perceived as objective. Denote $C$’s posterior belief of agent $A$ and $B$ being objective, formed at the evaluation stage, as $\pi_A$ and $\pi_B$ respectively, which is referred to as $A$ and $B$’s posterior objectivity. Biased $A$ and $B$’s payoffs are assumed to be, respectively:

$$U_A = a + \alpha \pi_A \quad \text{and} \quad U_B = a + \beta \pi_B.$$  

The first half of a biased agent’s payoff function is $C$’s action. The higher is $C$’s action, the better off a biased agent is. The second half is a reduced form formulation representing a biased agent’s reputational payoffs.$^{11}$ Parameters $\alpha, \beta \in [0, \infty)$ are the weights biased $A$ and $B$ attach to their reputations. In summary, the game of communicating through a well informed intermediary is illustrated in Figure 1.

$^{11}$ A biased agent’s reputational payoffs may be convex in his posterior objectivity, reflecting the fact that information about the agent’s type itself may have value for the decisionmaker (Morris 2001, Li 2007a).
A receives $s_A$ and sends $m_A$ to $B$  

$B$ receives $m_A, s_B$ and sends $m_B$ to $C$  

$C$ takes action $a$  

State $\eta$ observed  

$C$ evaluates $A, B$  

Figure 1: Timeline

In this game, a strategy of biased $A$ consists of two probabilities of reporting truthfully, one for each signal $s_A$. Analogously, a strategy of biased $B$ consists of four probabilities of reporting truthfully, one for each combination of his own signal and $A$’s message. This paper looks for perfect Bayesian equilibrium (PBE), in which given strategies of biased $A$ and $B$, decisionmaker $C$’s action $a$ at the decision-making stage maximizes her (expected) payoff, and her posterior beliefs at the evaluation stage $\pi_A, \pi_B$ satisfy Bayes’ rule.

In the medical industry example, the true state refers to the effectiveness of a drug: it may be “useless” (state 0) or “useful” (state 1). Agent $A$ is the manufacturer of this drug and agent $B$ may be either a physician who makes a recommendation to the patients, or a medical researcher who studies the effectiveness of the drug. The decisionmaker may be a patient who needs to decide how much to rely on this drug, or the medical community who learns about the drug from the researcher. Both agent $A$ and $B$ have information about the true state, but one key assumption of this model is that agent $B$’s signal is more informative than agent $A$’s, reflecting the fact that $B$ either has experience or expertise in evaluating a product. For example, a physician is better at figuring out whether a drug is useful for his patients. In these settings, the decisionmaker prefers hearing from an independent expert or reviewer even at some cost of information aggregation, which is more reasonable if $A$’s private information is of lower quality. Moreover, agent $B$’s recommendation $m_B \in \{0, 1\}$ is best thought of as a simple, “yes-or-no” type of recommendation, which is the simplest way to illustrate the direction of biased $B$’s distortions given his information.12

An objective agent in this model always passes on the information he believes to be the most accurate. This can be justified either on the grounds of professional ethics or institutional goals.13 Clearly, objective $A$ reports his only (thus best) piece of information truthfully: $m_A = s_A$. Although objective $B$ has two pieces of information $m_A, s_B$, his message choice is also very simple:

12 Allowing $B$ to convey both $A$’s message and his own signal is discussed further in Section 5.
13 For instance, Lahey Clinic, one of the major U.S. adult care hospitals writes: “Because good ethics begins with good medicine, the patient or DPAHC must receive accurate medical information and must understand it.”
**Observation 1** Objective B always reports his own signal regardless of A’s message: \( m_B = s_B \).

Because B’s signal is a better source of information than that of A’s, whether A’s message confirms or contradicts it, his recommendation should simply reflect his signal.\(^{14}\) In other words, objective B is not influenced by A’s message.

Although messages in this model are private and unverifiable, this is not a cheap talk game and no uninformative, babbling equilibria exist. Because objective agents always send the most accurate information, biased agents must pay an endogenous cost of passing on a distorted message. By distorting, a biased agent deviates from his best estimate of the state and thus is less likely to be considered objective in the eyes of C.

Before turning to the analysis, it is useful to keep in mind a few other applications of the model. In an application to real estate business, agent A is a mortgage broker (or a bank) and the state is the value of a property. Agent B is a real estate appraiser who specializes in estimating the market value of a property. Biased A wants a higher appraisal so as to boost the commissions from financing the purchase while biased B wants to inflate the appraisal, perhaps for future referrals. Decisionmaker C is the perspective homeowner who needs to decide what to bid for the property.\(^{15}\)

In an application to political arena, agent A is a political action committee (PAC) who may be genuinely concerned about a policy or biased toward certain special interests. The state is the potential impact of this policy. Agent B is a legislator overseeing this policy related area and makes a recommendation about its impact. Biased B may want to push for the policy but is also concerned about his political future. Decisionmaker C is the legislature who decides how much to support this policy.

\(^{14}\) More precisely, \( \Pr(\eta = s_B) > \Pr(\eta = m_A) \) for any message agent A, objective or biased, may send. It also implies that within the confines of this model, information aggregation is less important to the decisionmaker than hearing from objective B, a well informed expert.

\(^{15}\) Biased appraisals have become a key issue in several recent lawsuits. For instance, New York’s attorney general announced a case against eAppraiseIT, a leading appraisal management firm, for caving in to pressure from Washington Mutual to use a list of “proven appraisers” who he claims inflated home appraisals. Associated Press, November 1, 2007.
3 Analysis

3.1 Biased Intermediary’s Best Response

Objective $B$ bases his recommendation on his own signal, but biased $B$ may be influenced by $A$ for two reasons. First, if $A$’s message contains information, it affects biased $B$’s own estimates of the true state, which in turn affect his (expected) reputation. Second, if biased $B$ may be influenced by $A$, the decisionmaker would take such influence into account, which affects the credibility of $B$’s recommendation. This subsection focuses on how biased $B$ is influenced by a potentially biased message from agent $A$.

Suppose that biased $A$ and biased $B$ adopt the following type of strategy, which is referred to as a distortion strategy from now on. Biased $A$ reports truthfully if his signal favors his bias: $m_A = 1$ if $s_A = 1$, but $m_A = 0$ with probability $x$ if $s_A = 0$. Similarly, biased $B$ reports truthfully if his signal supports his bias: $m_B = 1$ if $s_B = 1$. If $s_B = 0$, he reports $m_B = 0$ with probability $y_2$ if $A$’s message agrees with his signal ($m_A = 0$); but reports $m_B = 0$ with probability $y_1$ if $A$’s message contradicts his signal ($m_A = 1$). Note that $x, y_1$, and $y_2$, chosen later to construct an equilibrium, are all probabilities that a biased agent truthfully reports his signal. Given the distortion strategy, biased $B$ chooses $m_B$ to maximize:

$$EU_B(m_B|m_A, s_B) = \Pr(\eta = 1|m_B) + \beta \mathbb{E}_{\eta}[\Pr(B = 0|m_B, \eta)|m_A, s_B].$$

The first part is $C$’s action given $B$’s recommendation, and the second is his (expected) posterior objectivity, where the expectation is taken with respect to state $\eta$.

To begin with, the difference in $C$’s actions if biased $B$ reports $m_B = 1$ instead of $m_B = 0$, $\Pr(\eta = 1|m_B = 1) - \Pr(\eta = 1|m_B = 0)$, is always positive. This difference is biased $B$’s net benefit from distortion, because $C$ always takes a higher action, which is more favorable toward his bias, after a positive recommendation. Intuitively, even if biased $B$ lies completely, $m_B = 1$ may result from signal $s_B = 1$, but $m_B = 0$ must result from $s_B = 0$. Thus it is less likely that $\eta = 1$ if $m_B = 0$.

Biased $B$ is also concerned about how $m_B$ affects his (expected) posterior objectivity. Biased $B$’s net cost of distortion is the difference in his posterior objectivity if he reports $m_B = 0$ instead of $m_B = 1$:

$$\sum_{\eta} \Pr(\eta|m_A, s_B) [\Pr(B = 0|m_B = 0, \eta) - \Pr(B = 0|m_B = 1, \eta)],$$
which is always positive regardless of state $\eta$, because $m_B = 0$ is more likely to come from objective $B$ and is thus a better sign of $B$’s objectivity. Clearly, $m_A$ affects biased $B$’s estimate of the true state $\Pr(\eta|m_A,s_B)$: it increases biased $B$ belief in his own signal if they agree, but decreases it otherwise.

It is worth noting that because objective $B$ never follows $m_A$, not being influenced by $A$’s message is a sign of objectivity. Thus biased $B$ may have an incentive to show his recommendation is independent. In particular, given the distortion strategy, biased $B$’s posterior objectivity $\Pr(B = o|m_B = 0, \eta = 1) > \Pr(B = o|m_B = 0, \eta = 0)$: he is considered more objective despite a wrong recommendation. Intuitively, if $m_B = 0$ but $\eta = 1$, $C$ thinks that it is likely that $m_A = 1$, but $B$ did not follow $m_A$. This effect is particularly strong if $A$’s signal is very informative, because biased $B$ is less confident about his own signal, and may want to appear independent by reporting $m_B = 0$ when $m_A = 0$, $s_B = 1$. To ensure that biased $B$ distorts only when his signal does not support his bias, we limit our attention from now on to the case when $p_A$ is sufficiently close to 0.5.\footnote{All the results in this paper hold if in equilibrium, biased $B$ is willing to report $s_B = 1$ truthfully. One sufficient, but not necessary condition for this to occur is if $p_A$ is sufficiently close to 0.5; another is for $\theta_B$ to be sufficiently close to 1. In both cases, $m_B = 0$ is still a sign of objective $B$ reporting $s_B = 0$ rather than a sign of whether biased $B$ is influenced by $A$ or not. If $p_A$ is sufficiently high, then biased $B$ may lie against $s_B = 1$ and report $m_B = 0$ in a perverse equilibrium.}

If his signal does not favor his bias ($s_B = 0$), biased $B$ distorts and reports $m_B = 1$ if, given his information, his net benefit from distortion strictly exceeds his net reputation cost; otherwise he reports $m_B = 0$ with some probability. For any given $x$, biased $B$’s best response is to choose probabilities $y_1$ and $y_2$ to maximize his expected payoff, incorporating how $C$ chooses her actions given $m_B$; and how she updates her belief about $B$’s objectivity in the evaluation stage given their distortion strategies.

**Proposition 1** For any given $x$, biased $B$ has a unique pair of best response $(y_1, y_2)$ which satisfies the following properties: (1) $y_1$ and $y_2$ cannot both be strictly between 0 and 1; (2) $y_2 \geq y_1$, and the inequality is strict if $\max\{y_1, y_2\} > 0$; (3) If $y_1 \in (0,1)$, it decreases in $x$; similarly, if $y_2 \in (0,1)$, it decreases in $x$.

Biased $B$’s net benefit from distortion does not vary with his signal and $A$’s message, but his net reputation cost of distortion does. In particular, if $s_B = 0$, his net reputation cost if $m_A = 0$ is higher than that if $m_A = 1$, because lying against two signals is more likely to lead to the worst reputation ($\Pr(B = o|m_B = 1, \eta = 0)$). Therefore biased $B$ cannot be indifferent between reporting truthfully and distorting both when $m_A = 0$ and when $m_A = 1$: if so, his net benefit from distortion must be the
same as his net reputation cost in both cases, which is impossible. Hence \( y_1 \) and \( y_2 \) cannot be both strictly between 0 and 1. Moreover, Proposition 1 shows that if \( s_B = 0 \), biased \( B \) is more apt to lie if \( A \)'s message agrees with, rather than contradicts, his bias: \( y_2 > y_1 \) if \( \max\{y_1, y_2\} > 0 \). This is because \( m_A = 1 \) lends some support for his bias while \( m_A = 0 \) is another strike against it. If \( s_B = 0 \) and biased \( B \) is indifferent between reporting \( m_B = 1 \) and \( m_B = 0 \) when \( m_A = 0 \); he must strictly prefers distorting and reporting \( m_B = 1 \) when \( m_A = 1 \): \( y_1 = 0, y_2 \in (0,1] \). If biased \( B \) never distorts when \( m_A = 0 \), he reports \( m_B = 0 \) with some probability strictly smaller than 1 when \( m_A = 1 \): \( y_1 \in (0,1), y_2 = 1 \). However, biased \( B \) is never completely truthful, because his recommendation is credible if \( y_1 = 1 \); yet he pays no reputation cost for a wrong recommendation, because \( C \) rationally attributes the mistake to a faulty signal. Also, the bigger is the gap between \( y_1 \) and \( y_2 \), the more biased \( B \)'s recommendation hinges on biased \( A \)'s relatively uninformative message.

Note that biased \( B \)'s net benefit from distortion increases in his truth-telling probability while his net reputation cost decreases in it. Intuitively, the more truthful he is, the more credible his recommendation is; also, his reputation depends less on his recommendations, because any mistake is less likely to result from his distortion. Hence biased \( B \) has a unique best response to any given \( x \). He may strictly prefers reporting \( m_B = 1 \) and thus \( y_1 = 0, y_2 = 0 \). Otherwise, because truthful reporting is never a best response, and because of the above monotonicity, there exists a unique probability strictly between 0 and 1 such that he is indifferent between reporting \( m_B = 0 \) or \( m_B = 1 \).

Somewhat surprisingly, it is cheaper for biased \( B \) to distort if \( A \)'s message is more truthful. More formally, biased \( B \)'s truth-telling probability, if strictly between 0 and 1, decreases in \( \theta_A + (1 - \theta_A)x \), the perceived probability that agent \( A \) reports \( m_A = 0 \). \( A \)'s message may become more truthful either because biased \( A \) reports more truthfully (\( x \) increases); or because \( A \) is more likely to be objective (\( \theta_A \) increases). One may think that biased \( B \) should lie less instead: since \( C \) cannot observe \( A \)'s message, \( A \) and \( B \) should share the blame if \( B \)'s recommendation turns out to be wrong. And if \( C \) attributes less blame to \( A \), whose message is more truthful, it is more costly for biased \( B \) to lie. In this model, however, any sign of being influenced is a sign of bias, thus biased \( B \) cannot shift any blame to \( A \). Instead, if \( s_B = 0 \), a more truthful message from biased \( A \) has two effects on biased \( B \).

The first is an information aggregation effect: how likely biased \( B \) believes that \( \eta = 0 \) given \( m_A \), holding \( C \)'s evaluations of \( B \)'s posterior objectivity fixed. If \( m_A \) is more truthful, \( m_A = 1 \) is more likely...
to result from $s_A = 1$ and thus is more credible. Therefore biased $B$ is more likely to follow $A$’s message and report $m_B = 1$: everything being equal, he is less likely to receive the worst posterior objectivity $\Pr(B = o|m_B = 1, \eta = 0)$. Second, there is a reputation sensitivity effect: how responsive biased $B$’s posterior objectivity is to a more truthful $m_A$, holding biased $B$’s belief about $A$’s message’s credibility fixed. Observe that since $y_2 > y_1$ if $y_2 > 0$, it is more likely for biased $B$ to report $m_B = 0$ if biased $A$ reports $m_A = 0$ with a higher probability. In this way, $m_B = 0$ becomes a less positive signal of independence (and hence $B$’s objectivity). Similarly, $m_B = 1$ becomes a less negative signal of $B$’s objectivity because the chance of being influenced by $m_A = 1$ is smaller. Hence biased $B$’s reputation is less sensitive to his recommendation. Both these effects reduce biased $B$’s net reputation cost of distortion if $m_A$ is more truthful. Since biased $B$’s net benefit from distortion increases if $m_A$ is more truthful, he lies more as a consequence.

Another case of interest is how biased $B$’s truth telling changes with $A$’s signal quality. For instance, do physicians give more truthful recommendations when they have access to better information?

**Corollary 1** For a given $x$, if $y_1 \in (0, 1)$, it strictly decreases in $p_A$.

Biased $B$ may distort more, not less, if $A$’s signal quality increases. First, if $A$’s message becomes more informative, biased $B$ has more to gain from distortion. As $p_A$ increases, both $m_B = 1$ and $m_B = 0$ become more credible, thus biased $B$’s net benefit from distortion increases in $p_A$. If biased $B$ reports $m_B = 0$ with some probability when $m_A = 1$, his net reputation cost falls in $p_A$ due to the information aggregation effect and the reputation sensitivity effect. Since his net benefit from distortion increases in $p_A$, biased $B$ distorts more as $p_A$ increases. This result provides a rationale for why the decisionmaker may encourage well informed experts to give independent recommendations instead of incorporating weak sources of information. The decisionmaker benefits little from the additional information in $m_A$, but she loses strictly because biased $B$ distorts more against his more informative signal.

### 3.2 Equilibrium

This section focuses on the strategic interactions between biased $A$ and $B$. Given biased $A$ and $B$’s distortion strategy, biased $A$ chooses a message $m_A$ to maximize his expected payoff:

$$E_{m_B} \left[ \Pr(\eta = 1|m_B) + \alpha \Pr(A = o|m_B, \eta) \right] | s_A, m_A.$$
Note that biased A not only takes expectation with respect to the true state, but also with respect to B’s recommendation, because he can only influence the decisionmaker through B. Biased A knows that if \( s_B = 1 \), B always reports \( m_B = 1 \). Thus the pivotal event for biased A, which determines his message choice, is if biased B truthfully reports \( m_B = 0 \) when \( s_B = 0 \). Biased A’s message only matters if biased B changes his recommendation from \( m_B = 0 \) to \( m_B = 1 \) because of \( m_A \). Let \( \kappa \equiv (1 - \theta_B)(y_2 - y_1) \). If \( s_A = 0 \), then the net benefit from distortion if biased A reports \( m_A = 1 \) instead of \( m_A = 0 \) is:

\[
\kappa \Pr(s_B = 0|s_A = 0)[\Pr(\eta = 1|m_B = 1) - \Pr(\eta = 1|m_B = 0)].
\]

This benefit is simply the net benefit of biased B multiplied by the probability that biased B changes his recommendation because of \( m_A \). Intuitively, if B is known to be objective (\( \theta_B = 1 \)); or if biased B always distorts (\( y_1 = y_2 = 0 \)), A’s message has no impact on C’s action (\( \kappa = 0 \)). Moreover, the larger is the gap between \( y_2 \) and \( y_1 \), the stronger is A’s (indirect) influence on C’s action. If \( s_A = 0 \), biased A’s net reputation cost if he distorts is:

\[
\kappa \Pr(s_B = 0|s_A = 0) \sum_\eta \Pr(\eta|s_A = 0, s_B = 0)[\Pr(A = o|m_B = 0, \eta) - \Pr(A = o|m_B = 1, \eta)],
\]

which only depends on the event that both their signals are against their bias. Otherwise, biased A’s reputation does not depend on his own message. Biased A chooses a message by comparing his net benefit from distortion with his net reputation cost from doing so.

A distortion equilibrium is one in which both biased agents use distortion strategy. Since a distortion equilibrium consists of three probabilities: \( x, y_1, y_2 \), biased A’s best response to biased B’s truth telling depends on both \( y_1 \) and \( y_2 \); and biased B’s best response \((y_1, y_2)\) depends on biased A’s truth-telling probability \( x \). The next result summarizes the biased agents’ equilibrium behavior.

**Proposition 2** There exist cutoff values \( \beta^c, \beta^s, \beta^w \) satisfying \( \beta^c < \beta^s < \beta^w \) such that: (1) a complete distortion equilibrium exists if \( \beta \leq \beta^c \), with \( y_1 = 0, y_2 = 0 \) and \( x \in [0,1] \); (2) a strong distortion equilibrium exists if \( \beta \in [\beta^c, \beta^s] \), with \( y_1 = 0, y_2 \in (0,1) \) and \( x \in [0,1] \) such that \( x = 0 \) if \( \beta \) is sufficiently close to \( \beta^c \); (3) a weak distortion equilibrium exists if \( \beta > \beta^w \), with \( y_2 = 1, y_1 \in (0,1) \) and \( x \in [0,1] \) such that \( x = 0 \) if \( \beta \) is sufficiently high; (4) either a strong or a weak distortion equilibrium exists, or both exist, if \( \beta \in (\beta^s, \beta^w) \).
If biased $B$ has very low reputational concerns ($\beta \leq \beta^c$), biased $B$ always reports $m_B = 1$ because his net benefit from distortion is always higher than his net reputation cost. Because objective $B$ always reports his true signal, and biased $B$ always reports $m_B = 1$, biased $A$ has no influence on the decisionmaker: his message does not affect $C$’s action, nor does it affect his own reputation. Therefore biased $A$ is free to choose any truth-telling probability $x \in [0,1]$ since his message does not matter. Biased $A$ only has influence if biased $B$ may alter his recommendation because of $m_A (y_1 \neq y_2)$. In this case, biased $A$ never reports completely truthfully: if so, then $m_A$ can indirectly induce a higher action from $C$ at no reputation cost. For biased $A$, if there is influence, there is distortion.

If biased $B$ has high reputational concerns, he does not want to distort at all when both signals are against his bias. The cutoff value is defined such that at $\beta = \beta^w$, $y_1 = 0, y_2 = 1$ is biased $B$’s best response to $x = 1$ if $m_A = 1, s_B = 0$. That is, $\beta^w$ is the lowest weight on reputation such that, if biased $A$ reports truthfully, biased $B$ can be indifferent between reporting $m_B = 1$ and $m_B = 0$ if $m_A = 1, s_B = 0$. Because biased $B$’s net benefit from distortion increases in $x, y_1$ and $y_2$ while his net reputation cost decreases in them, if $\beta > \beta^w$, $y_1 = 0, y_2 \leq 1$ is never a best response to any $x$. Intuitively, biased $B$’s maximum net benefit from distortion is smaller than his minimum net reputation cost in any strong distortion equilibrium, which is impossible.

In any weak distortion equilibrium, biased $A$’s net reputation cost of distortion falls if biased $B$ reports more truthfully, due to the aforementioned reputation sensitivity effect. If $y_1$ increases, $C$ knows that biased $B$ is less influenced by $A$’s message, thus her posterior belief of $A$’s objectivity varies less with $B$’s recommendation. In particular, $m_B = 0$ is a less positive sign of $A$’s objectivity because $C$ believes that it is more likely that biased $B$ follows his own signal; while $m_B = 1$ is a less negative sign of $A$’s objectivity because $C$ thinks that it is more likely that biased $B$ receives the wrong signal. This also explains why, if $y_1$ is sufficiently close to 1, biased $A$ always reports $m_A = 1$. Because biased $B$ (almost) reports his signals truthfully, his recommendation is credible. Thus biased $A$’s net benefit from distortion is strictly positive while his net reputation cost approaches zero, and he lies with probability one. Therefore, even if biased $A$ is very concerned about his reputation, he may nonetheless distort with a high probability when indirect influence is involved. In addition, because the biased agents’ net benefit from distortion strictly increases in $x$ and $y_1$, and their net reputation cost decreases in the $\beta^c$ is independent of $x$. Instead, it depends on $B$’s characteristics and $A$’s signal quality, which affects $B$’s private estimate of the true state if $m_A = 0, s_B = 0$. 

17
other’s truth-telling probability, $x$ and $y_1$, if positive, are strategic substitutes: if biased $A$ reports more truthfully, biased $B$ distorts more.

If biased $B$ has low reputational concerns ($\beta \in [\beta^c, \beta^w]$), he wants to distort with some probability when both signals are against his bias. The cutoff value is defined such that at $\beta = \beta^s$, $y_1 = 0, y_2 = 1$ is biased $B$’s best response to $x = 0$ if $m_A = 0, s_B = 0$. That is, $\beta^s$ is the highest weight on reputation such that, if biased $A$ lies completely, biased $B$ can be indifferent between reporting $m_B = 1$ and $m_B = 0$ if $m_A = 0, s_B = 0$. If $\beta < \beta^s$, $y_1 > 0, y_2 = 1$ is never a best response to any $x$. In this case, only strong distortion equilibrium exists, because biased $B$’s minimum net benefit from distortion still exceeds his maximum net reputation cost in any weak distortion equilibrium, which is impossible.

In a strong distortion equilibrium, biased $A$’s net reputation cost increases, rather than decreases, in biased $B$’s truth-telling probability $y_2$. This is because biased $B$ (primarily) follows $A$’s message if $s_B = 0$: he reports $m_B = 1$ if $m_A = 1$ and $m_B = 0$ with probability $y_2$ if $m_A = 0$. Thus if $m_B = 1$, $C$ assigns some blame of the possibly distorted recommendation to $A$ for misleading biased $B$. If biased $B$ is more truthful, $m_B = 0$ becomes a better sign of $A$’s objectivity because it is more likely that biased $B$ has followed $m_A = 0$; while $m_B = 1$ is a worse sign of $A$’s objectivity because biased $B$ is less likely to report $m_B = 1$ if $m_A = 0$. This asymmetry in biased $A$’s net reputation cost with respect to biased $B$’s truth-telling probability $y_1$ and $y_2$ is driven by the fact that in a weak distortion equilibrium, if biased $B$ reports more truthfully, he does so despite $A$’s message; and thus biased $A$’s reputation depends less on $B$’s recommendation if $y_1$ increases.

If biased $B$ has intermediate levels of reputational concerns ($\beta \in (\beta^s, \beta^w]$), his equilibrium behavior is very sensitive to how truthful biased $A$ is. Specifically, he prefers strong distortion if $A$’s truth-telling probability $x$ is above a cutoff value; but weak distortion if $x$ is smaller than this cutoff value. If biased $A$ reports truthfully with a higher probability than this cutoff when his message has the maximum influence on biased $B$ ($y_1 = 0, y_2 = 1$), a strong distortion equilibrium exists. Otherwise, a weak distortion equilibrium exists. However, in this range, multiple equilibria involving both strong distortion and weak distortion may exist.$^{18}$

$^{18}$ Because biased $A$’s best response to $y_2$ may not be monotonic, multiple equilibria cannot be ruled out. As a numerical example, if $\alpha = 10, \theta_A = \theta_B = 0.5, p_A = 0.7, p_B = 0.75$, multiple equilibria exist for $\beta \in [1.55, 1.62]$. 

15
3.3 Information Loss for the Decisionmaker

If well informed intermediaries are influenced by sources with inferior and potentially distorted information, their biased recommendations may lead the decisionmaker to take wrong actions. The decisionmaker may want to reduce such losses by encouraging truthful reporting from the agents. Since $C$’s optimal action given $m_B$ is $a = \Pr(\eta = 1|m_B)$, her ex ante expected payoff is simply:

$$-E_\eta E_{m_B}[(\Pr(\eta = 1|m_B) - \eta)^2|m_B] = -0.5 \sum_{m_B} \Pr(\eta = 1|m_B)(1 - \Pr(\eta = 1|m_B)).$$

Note that the decisionmaker is better off if $B$’s recommendations become more credible, that is, her expected payoff increases in the probability that $m_B$ is correct: $\Pr(\eta = 1|m_B = 1)$ and $\Pr(\eta = 0|m_B = 0)$. The biased agents may lower $C$’s expected payoffs in two ways: they make $m_B = 1$ less credible by distorting $s_A = 0$ or $s_B = 0$; and to a lesser extent, even when biased $B$ reports $m_B = 0$ truthfully, some useful information from $A$ may be lost if $s_A = 1$.

One way for the decisionmaker to increase her expected payoff is to take measures to change the net reputation cost of the intermediary by changing $\beta$; or she may change that of the source by changing $\alpha$. In practice, such measures may take the form of more stringent and thorough medical board reviews; or different regulations on disclosure by the pharmaceutical companies. For such measures to be effective, we must first consider how biased $A$ and $B$’s truth-telling incentives vary with changes in their reputational concerns. The following result shows that in a strong distortion equilibrium, changes in $\alpha$ and $\beta$ are not equivalent, and the decisionmaker may be worse off if she chooses the wrong measure.

**Proposition 3** If $x > 0, y_2 > 0$ in a strong distortion equilibrium, then $x$ increases in $\alpha$ and $y_2$ increases in $\beta$; and if further $\beta$ is sufficiently close to $\beta^c$ or if $\theta_B$ is sufficiently high, then $x$ increases in $\beta$ and $y_2$ decreases in $\alpha$.

To begin with, decisionmaker $C$’s expected payoff increases in both $x$ and $y_2$ in a strong distortion equilibrium. It increases in $x$ because, holding biased $B$’s truth telling fixed, the more truthful $A$ is, the

---

19 The settlement of the $185 million class action lawsuit against Bristol-Myers Squibb in January 2006 shows that they paid physicians to exaggerate, in major medical meetings, the benefits of their drug for patients with high blood pressure and heart failure; these physicians also failed to report publicly on substantial numbers of life-threatening drug complications which they knew, from their close relationship to the company, to exist.

20 Even though here we only consider measures affecting the reputational cost of the agents and ignore monetary fines or any cost required to implement and enforce such measures, such costs can be easily incorporated. For instance, if $C$’s cost of imposing stronger regulations are sufficiently convex, then only small changes in $\alpha, \beta$ may be feasible, which may be counterproductive as implied by Proposition 3.
more likely biased A reports \( m_A = 0 \) which is passed on by biased B. It also increases in \( y_2 \) because biased B lies against both signals in a strong distortion equilibrium, which is the worst distortion. Even though C prefers biased A and B to report more truthfully, changes in \( \alpha \) or \( \beta \) may have different net effect due to their strategic interactions. Surprisingly, Proposition 3 suggests that, increasing the reputation cost of pharmaceutical companies in the opening example may make the patients worse off.

To see this, observe that when biased B’s reputational concerns are sufficiently low or he is considered very objective, \( y_2 \) is close to 0, and thus \( x \) is close to 0 because biased A has little influence on B. If \( \alpha \) increases, biased A reports more truthfully because it is more costly for him to lie, which leads biased B to distort more because of the reputation sensitivity effect. The increase in \( x \), however, has a negligible impact on C because biased A still primarily reports \( m_A = 1 \). But biased B lies against both signals with a higher probability. Thus the negative effect from the decrease in \( y_2 \) dominates, and C is strictly worse off.\(^{21}\) Intuitively, a small increase in the source’s reputational concerns may worsen the intermediary’s incentives when it counts the most.

Instead, the decisionmaker should increase \( \beta \), which makes biased B more truthful. For biased A, the increase in his net reputation cost due to the blame sharing effect exceeds the gain in his net benefit from distortion, which is negligible because \( y_2 \) is still close to 0. Thus both \( y_2 \) and \( x \) increase in \( \beta \), and C is strictly better off. In the medical industry example, strengthening medical board review process or disclosure rules may improve the credibility of the medical profession both because the physicians are less influenced, and because the drug companies may become more truthful in revealing side effects.

In a weak distortion equilibrium, measures that increase biased A or B’s reputation cost tend to have a positive, but weak effect on their truth telling:

**Proposition 4** Suppose that \( x > 0, y_1 > 0 \) in a weak distortion equilibrium, then \( x \) increases in \( \alpha \) and \( y_1 \) increases in \( \beta \). Also, \( y_1 \) decreases in \( \alpha \) and \( x \) decreases in \( \beta \).

Because A’s signal is not very informative, the loss of A’s possibly distorted message is small. In a weak distortion equilibrium, decisionmaker C’s expected payoff increases in \( y_1 \). Because their truth-telling probabilities are substitutes, an increase in \( \beta \) leads biased A to lie more (or completely if \( \beta \) is sufficiently high). In this region, however, because B reports truthfully with a high probability, C’s expected payoff

---

\(^{21}\) As a numerical example, if \( \theta_A = \theta_B = 0.5, p_A = 0.7, p_B = 0.95 \), then for \( \beta \in [0.7, 1] \), an increase from \( \alpha = 1 \) to \( \alpha = 2 \) reduces C’s expected payoff.
increases gradually in $\beta$, and the marginal benefit of raising $\beta$ approaches zero if $\beta$ is sufficiently high.\footnote{Hence measures raising $\beta$ for intermediaries sufficiently concerned about their reputations makes $C$ worse off if a small fixed cost is involved.}

Similarly, an increase in $\alpha$ has a negligible positive effect on $C$.

4 Comparing Communication Methods

For legal or institutional reasons, in environments such as marketing and lobbying, exerting influence through intermediaries is common; in other environments such as advertising, information is often conveyed directly. Simultaneous reporting serves as a natural benchmark against the main model, because both agents have informative signals and their messages are potentially useful if they influence the decisionmaker directly. This section studies the biased agents’ truth-telling incentives under simultaneous reporting and how these incentives affect the decisionmaker’s expected payoff.

4.1 Simultaneous Reporting

Suppose instead of communicating through an intermediary, $A$ sends a message to $C$, who now receives both $m_A$ and $m_B$ before taking an action. All the other assumptions remain. Biased $A$ then chooses $m_A$ to maximize:

$$E_{m_B}[\Pr(\eta = 1|m_A, m_B)|s_A] + E_\eta[\Pr(A = o|m_A, \eta)|s_A];$$

and biased $B$ chooses $m_B$ in a similar way. Observe that the first part of biased $A$’s expected payoff depends on both $A$ and $B$’s messages. The presence of multiple messages affects the tradeoff a biased agent faces primarily by changing the marginal impact of his message on the decisionmaker’s action, and thus his (expected) net benefit from distortion. Note that the second part of biased $A$’s expected payoff, which is his (expected) posterior objectivity, is independent of $m_B$. Because $A$ and $B$’s signals are independent conditional on the state, which is observed by $C$ at the evaluation stage, $B$’s message does not impose additional discipline on biased $A$ through his reputational concerns.\footnote{Additional discipline is present if $A$ and $B$’s signals are correlated. This effect surfaces in Chan, Li, and Suen (2007), who study grade inflation and show that the market judges how likely one school has distorted in terms of grade inflation based on the other school’s grades if their students’ quality is correlated. Even if the signals are independent conditional on the state, Gerardi, McLean, and Postlewaite (2008) show that a decisionmaker who receives reports from multiple partially informed sources can extract more truthful reports by exploiting the correlation of the sources’ signals with the true state.}
Under simultaneous reporting, then, all the strategic interactions between the biased agents enter through their messages’ influence on C’s action; their posterior objectivity is evaluated independent of the other message. Each biased agent chooses a message after comparing his (expected) net benefit from distortion with his net reputation cost.

Proposition 5 There exists a distortion equilibrium in which biased agent $i = A, B$ reports $m_i = 1$ if $s_i = 1$. If $s_i = 0$, biased $i$ always reports $m_i = 1$ if he attaches little weight to his reputation ($\alpha$ or $\beta$ sufficiently low). Biased $i$ reports $m_i = 0$ truthfully with probability $x_i > 0$ if his weight on reputation is sufficiently high. Moreover, if positive, $x_A$ and $x_B$ are strategic complements.

The key of Proposition 5 is that, for each biased agent, the (expected) net benefit of distortion decreases in the other agent’s truth telling. Suppose that $s_A = 0$ and biased $B$ reports more truthfully ($x_B$ increases), then biased $A$ has a smaller marginal impact on $C$’s action if he reports $m_A = 1$ instead of $m_A = 0$. There are two cases. First, given $m_B = 1$, then as $x_B$ increases, $m_B = 1$ is more credible and $C$ becomes more convinced that the state is 1, and thus $m_A = 1$ induces a smaller additional change in her action. Second, if $m_B = 0$, then the marginal impact of $A$’s message does not vary with $x_B$ because $C$ knows that $s_B = 0$ in this case. However, because $C$ believes that $\eta = 0$ is more likely, $m_A = 1$ appears to be distorted and less credible, and thus $C$ changes her action less than if $m_B = 1$. Since $s_A = 0$, biased $A$ knows that $B$ is more likely to receive $s_B = 0$ and to report $m_B = 0$ as biased $B$ becomes more truthful. Because biased $A$ is less effective at changing $C$’s action in either case, his marginal impact on $C$ falls in $x_B$, but his net reputation cost is unaffected. Consequently, biased $A$ distorts less than he would have as the sole source of information. Moreover, this also implies that with simultaneous reporting, biased $A$ and $B$’s truth-telling probabilities $x_A, x_B$ are complements.24

The complementarity between biased agents’ truth telling implies that $C$ can increase both $x_B$ and $x_A$ by increasing either agent’s reputation cost. As both $A$ and $B$’s messages become more accurate, $C$’s expected payoff increases. Moreover, the decisionmaker should encourage simultaneous reports if biased $B$ is sufficiently concerned about his reputation, in which case biased $A$ and $B$’s truth-telling incentives are substitutes in the main model. Proposition 4 shows that in this case the marginal benefit of raising

---

24 This complementarity is driven by each agent’s decreasing marginal impact on $C$ instead of changes in reputation cost as in Li (2007b), who shows that with sequential communication, the uncertainty of which agent initiates a wrong message creates a blame sharing effect, which reduces all biased agents’ reputation costs and encourages them to lie more.
\( \alpha \) or \( \beta \) is small and it approaches zero if \( \beta \) is sufficiently high. Instead, with simultaneous reports, an increase in \( \alpha \) makes both biased \( A \) and \( B \) report more truthfully. In particular, biased \( A \) distorts with a smaller probability, both because of the strategic complementarity and because his reputation is more sensitive to his own message. In practice, this suggests that policies aiming to improve reporting accuracy from potentially biased sources should not be limited to stricter regulations against any single source of distortion; it may be more effective to encourage simultaneous reports. For example, it may be more effective to reduce the direct linkage between medical researchers, doctors and pharmaceutical companies. The University of Pennsylvania Health System took a step in this direction by requiring that industry can make gifts to departments to support educational programs (but not to individual faculty), while the money is disbursed at the discretion of department chairs.

One caveat concerning policies encouraging simultaneous reporting is that it may be counterproductive if both agents have sufficiently low reputational concerns that they always distort \( (x_i = 0) \). Because in the main model, \( C \) only receives a biased recommendation from \( B \). Under simultaneous reporting, however, she receives two distorted messages which reinforce and strengthen each other. Thus in situations where agents are very likely biased and have little reputational concerns, it may be better to discourage simultaneous reporting to reduce the number of distorted messages the decisionmaker receives, such as the amount of direct TV advertisements for prescription medicines.

4.2 Biased \( A \)'s Preference

Having examined some implications of simultaneous reporting on the decisionmaker’s expected payoff, this subsection proceeds to consider a related question: the ex ante preference of biased \( A \). That is, whether sending his own message or communicating through \( B \) gives biased \( A \) a higher expected payoff if he chooses an environment before receiving his signal.\(^{25}\) Knowing this helps us understand when influencing well informed experts is more useful to a biased source.

Biased \( A \)'s ex ante expected payoff if he communicates through \( B \) is simply:

\[
0.5 \left[ E_{m_B} \left[ \Pr(\eta = 1|m_B) + \alpha E_{\eta}[\Pr(A = o|m_B, \eta)] \right] \right].
\]

Note that if \( A \) has no influence on \( B \), then the first half of the above expression, \( C \)'s (expected) action,

\(^{25}\) This is the case where biased \( A \) commits to his choice before observing his signal, thus his preference does not reveal information about his signal.
purely depends on the characteristics of $B$; and the second half is just $\alpha\theta_A$, because his reputation is not affected at all. But if he has influence on $B$, biased $A$ needs to compare his impact on $C$’s action with his reputation cost.

**Proposition 6** If $\alpha$ and $\beta$ are sufficiently low, biased $A$ prefers simultaneous reporting if $\theta_B$ is sufficiently close to 0. If $\alpha$ and $\beta$ are sufficiently high, biased $A$ prefers using an intermediary.

The first part of Proposition 6 shows that if biased $A$ and $B$ have sufficiently low reputational concerns, biased $A$ may prefer sending his own message. Because here $A$’s message has little influence on biased $B$ due to the latter’s low reputational concerns, he has little influence on the decisionmaker, as shown in Proposition 2. Therefore biased $A$ prefers an environment where $C$’s action favors his bias more. If $B$’s recommendation is not credible due to his low prior objectivity, $m_A$ has a stronger impact on $C$’s action than $B$’s. Reporting $m_A = 1$ directly hurts $A$’s reputation, but it is negligible in this case. This suggests that if the intermediary, despite his better signal, is perceived to be very biased and not credible, biased $A$ should try to persuade the decisionmaker directly instead of wasting his signal.

Next, if in equilibrium, biased $A$ reports truthfully with positive probabilities in both these environments and biased $B$ does not lie completely, then the law of iterated expectation applies. Biased $A$’s ex ante expected payoff is simply the sum of his priors: $0.5 + \alpha\theta_A$, hence he is indifferent between simultaneous reporting and influencing $B$. To see the second part of Proposition 6, recall from Proposition 2 that biased $A$ always reports $m_A = 1$ if $\beta$ is sufficiently high for any $\alpha$. In this case, biased $B$ reports very truthfully ($y_1$ sufficiently close to 1) that biased $A$ pays a negligible reputation cost, which is proportional to $1 - y_1$; but he still has some influence on the decisionmaker because $B$’s recommendation is highly credible. Thus he can “afford” to lie while he could not do so, at the same level of reputational concerns, if he sends a message of his own. Hence if biased $B$ has sufficiently high reputational concerns, biased $A$ is better off influencing him than risking losing his own reputation. Therefore a pharmaceutical company concerned about its reputation prefers influencing physicians or researchers with high levels of reputational concerns.
5 Discussion and Concluding Remarks

This section discusses several main assumptions on the agents’ information quality, preferences of the objective agent, as well as the distribution of state $\eta$. It also suggests how agents’ behavior may change if these assumptions are varied.

A. Relative informativeness of signals. Intermediary B’s signal is assumed to be more informative than that of A’s, which implies that B is considered more objective if he appears free from A’s influence. This model thus fits more closely the environments where the intermediary possesses good information such as experts or critics. In some environments, agent A has superior and/or exclusive information, for example, military intelligence. In that case, A’s signal may be far more informative and an objective intermediary behaves differently. He follows A if it is unlikely that A has lied, but may dismiss it if biased A lies a lot: $m_B(m_A = 1, s_B = 0) = 0$ if $x \leq x_1$, $m_B(m_A = 1, s_B = 0) = 1$ otherwise. The cutoff value $x_1$ increases in $p_B$, and in particular, if B’s signal is sufficiently uninformative, an objective B always follows $m_A$ despite A’s potential distortion. Li (2007b) shows that when B serves as a pure intermediary, biased A and B’s truth-telling incentives are complements because they share the blame of any distortion. Moreover, biased A always lies more using an intermediary — and he may prefer a more biased intermediary to a more objective one — because the decrease in his reputation cost strictly outweighs any loss in his message’s credibility due to biased B’s possible distortion.

B. Message space of agent B. The well informed intermediary here sends a simple message $m_B \in \{0, 1\}$ to the decisionmaker C. One question is why the intermediary does not indicate both A’s message as well as his own signal. One reason for restricting intermediary’s message space is that often experts only convey their recommendations as opposed to its intensity to the decisionmaker, especially if the issue at stake is complex, for example the effectiveness of a new medicine. Another reason is that this restriction helps illustrate the direction of biased B’s distortions. The information aggregation effect still exists and may even be intensified under a richer message space. The objective intermediary reports a vector of messages $(m_A, s_B)$ truthfully, but the biased type still has incentives to push one or both messages toward 1.\footnote{Sixteen, instead of four, incentive constraints must be satisfied for biased B to report truthfully.} Giving B a richer message space may not reduce the information loss of the decisionmaker. For instance, if biased B’s reputational concerns are sufficiently low, he may report $m_B = (1, 1)$ if $s_B = 0$, which is more credible than $m_B = 1$ because it may be true signals from two
objective agents.

C. Role of the objective types. In this model, the objective agent has a preference for accuracy: he conveys the best information available to him. This assumption, consistent with the stated goals of experts such as stock analysts, critics and physicians, also greatly simplifies the inference problem of decisionmaker $C$. An objective agent, however, may be concerned about his reputation as well as passing on accurate information. Morris (2001) shows in a model of direct communication that an objective expert may lie even though he wants the decisionmaker to take the correct action, because she doesn’t want to be confused with a biased type and lose future influence. In the present model, similar incentives may arise if the objective intermediary is also concerned about his reputation. Recall that $m_B = 0$ but $\eta = 1$ is a good sign of $B$’s objectivity because it suggests that he is not influenced by $m_A = 1$. Thus an objective $B$ may have stronger incentives to report 0 if $m_A = 0, s_B = 1$, driving biased $B$ to do so as well.

D. Asymmetric state distribution. In this model, the states are assumed to be ex ante symmetric, but the agents may believe that one state is more likely than the other, for instance patients may be suspicious of the effectiveness of a new drug: $\Pr(\eta = 0) > 0.5$. In this case, objective $B$ is less likely to give a positive recommendation than if he has symmetric priors. That is, he may report $m_B = 0$ with a positive probability even if $s_B = 1$, which makes it less likely for biased $B$ to report $m_B = 1$. Because it increases biased $B$’s net reputational cost and decreases the net benefit of distortion. This effect may be very pronounced if the biased agents have high reputational concerns: sometimes the only equilibrium is an uninformative one in which only $m_B = 0$ reaches the decisionmaker.

E. Conclusion. In the opening examples, recommendations from physicians or other well informed experts are important for the patients or the consumers. However, even though independent judgment is a sign of objectivity, biased experts may selectively incorporate inferior information from potentially biased sources before making their recommendations. Thus the decisionmaker may want to encourage the experts to give recommendations based on their own information. However, policy measures designed to make the source more truthful may reduce the truthfulness of the expert’s recommendation to such an extent that the decisionmaker is strictly worse off. It may be more effective to either increase the expert’s cost of distortion; or to encourage simultaneous reports from the sources and the experts by severing the ties between them.
Appendix

Proof of Proposition 1: We first consider biased $B$’s truth-telling incentives given his signal $s_B$ and message $m_A$ and then prove Proposition 1. Given biased $A$, $B$’s distortion strategy, biased $B$ chooses a message $m_B$ to maximize his expected payoff. If he reports $m_B = 1$ instead of $m_B = 0$, the difference in biased $B$’s expected payoffs, $EU_B(m_B = 1|m_A, s_B) - EU_B(m_B = 0|m_A, s_B)$, is:

$$
Pr(\eta = 1|m_B = 1) - Pr(\eta = 1|m_B = 0) - \beta \sum_\eta Pr(\eta|m_A, s_B)[Pr(B = o|m_B = 0, \eta) - Pr(B = o|m_B = 1, \eta)].
$$

Let the probabilities that $C$ believes agent $A$ reports $m_A = 0$ be $N_x = \theta_A + (1 - \theta_A)x$; and the probabilities that agent $B$ reports $m_A = 0$ be $N_{y_1} \equiv \theta_B + (1 - \theta_B)y_1$ and $N_{y_2} \equiv \theta_B + (1 - \theta_B)y_2$, depending on $m_A$. Also, denote $C$’s actions given $B$’s recommendation as $\hat{\eta}_1^B \equiv Pr(\eta = 1|m_B = 1)$ and $\hat{\eta}_0^B \equiv Pr(\eta = 1|m_B = 0)$, then we have:

$$
\hat{\eta}_1^B = \frac{1 - (1 - p_B)[N_{y_1}(1 - (1 - p_A)N_x) + N_{y_2}(1 - p_A)N_x]}{2 - (1 - p_B)(N_{y_1}(1 - (1 - p_A)N_x) + N_{y_2}(1 - p_A)N_x) - p_B[N_{y_1}(1 - p_A N_x) + N_{y_2}p_A N_x];}
$$

$$
\hat{\eta}_0^B = \frac{(1 - p_B)[N_{y_1}(1 - (1 - p_A)N_x) + N_{y_2}(1 - p_A)N_x] - p_B[N_{y_1}(1 - p_A N_x) + N_{y_2}p_A N_x]}{(1 - p_B)[N_{y_1}(1 - (1 - p_A)N_x) + N_{y_2}(1 - p_A)N_x] + N_{y_2}(1 - p_A)N_x] + p_B[N_{y_1}(1 - p_A N_x) + N_{y_2}p_A N_x].}
$$

The difference in biased $B$’s information is captured in his estimate of the true state $Pr(\eta = 0|m_A, s_B)$, which are respectively:

$$
\Gamma_1 \equiv Pr(\eta = 0|m_A = 1, s_B = 0) = \frac{(1 - p_A N_x) p_B}{(1 - p_A N_x)p_B + (1 - (1 - p_A) N_x)(1 - p_B)};
$$

$$
\Gamma_2 \equiv Pr(\eta = 0|m_A = 0, s_B = 0) = \frac{p_A(1 - p_B)}{p_A(1 - p_B) + p_B(1 - p_A)};
$$

$$
\Gamma_3 \equiv Pr(\eta = 0|m_A = 0, s_B = 1) = \frac{(1 - p_A N_x)(1 - p_B)}{(1 - p_A N_x)(1 - p_B) + (1 - (1 - p_A) N_x)p_B};
$$

$$
\Gamma_4 \equiv Pr(\eta = 0|m_A = 1, s_B = 1) = \frac{(1 - p_A N_x)(1 - p_B)}{(1 - p_A N_x)(1 - p_B) + (1 - (1 - p_A) N_x)p_B}.
$$

Because $p_B > p_A$, simple calculations can show that $\Gamma_2 > \Gamma_1 > 0.5 > \Gamma_3 > \Gamma_4$. To simplify notations, denote $B$’s posterior objectivity as $\tau_{m_B, \eta}^B$, and we have:

$$
\tau_{1,0}^B \equiv Pr(B = o|m_B = 1, \eta = 0) = \frac{(1 - p_B)\theta_B}{1 - p_B + (1 - \theta_B)p_B[(1 - y_1)(1 - p_A N_x) + (1 - y_2)p_A N_x]};
$$

$$
\tau_{0,0}^B \equiv Pr(B = o|m_B = 0, \eta = 0) = \frac{\theta_B}{\theta_B + (1 - \theta_B)[y_1(1 - p_A N_x) + y_2p_A N_x]};
$$

$$
\tau_{0,1}^B \equiv Pr(B = o|m_B = 0, \eta = 1) = \frac{\theta_B}{\theta_B + (1 - \theta_B)[y_1(1 - (1 - p_A) N_x) + y_2(1 - p_A) N_x]};
$$
\[
\tau_{1,1}^B \equiv \Pr(B = 0|m_B = 1, \eta = 1) = \frac{p_B \theta_B}{p_B + (1 - \theta_B)(1 - p_B)(1 - (1 - p_A)N_x) + (1 - y_2)(1 - p_A)N_x}.
\]

For biased \( B \) to report \( s_B \) truthfully, the following four incentive constraints must be satisfied:

\[
\begin{align*}
\hat{\eta}_1^B - \tilde{\eta}_1^B &\leq \Delta_1 \equiv \beta \left[ \Gamma_1(\tau_{0,0}^B - \tau_{1,0}^B) + (1 - \Gamma_1)(\tau_{0,1}^B - \tau_{1,1}^B) \right]; \\
\hat{\eta}_1^B - \tilde{\eta}_0^B &\leq \Delta_2 \equiv \beta \left[ \Gamma_2(\tau_{0,0}^B - \tau_{1,0}^B) + (1 - \Gamma_2)(\tau_{0,1}^B - \tau_{1,1}^B) \right]; \\
\hat{\eta}_1^B - \tilde{\eta}_0^B &\geq \Delta_3 \equiv \beta \left[ \Gamma_3(\tau_{0,0}^B - \tau_{1,0}^B) + (1 - \Gamma_3)(\tau_{0,1}^B - \tau_{1,1}^B) \right]; \\
\hat{\eta}_1^B - \tilde{\eta}_0^B &\geq \Delta_4 \equiv \beta \left[ \Gamma_4(\tau_{0,0}^B - \tau_{1,0}^B) + (1 - \Gamma_4)(\tau_{0,1}^B - \tau_{1,1}^B) \right].
\end{align*}
\]

The first two incentive constraints, IC (1) and IC (2), concern the case when \( s_B = 0 \); and the last two, IC (3) and IC (4), concern the case when \( s_B = 1 \). The left hand side (LHS) of all the above ICs is the same, which measures biased \( B \)'s net benefit from reporting \( m_B = 1 \) instead of reporting \( m_B = 0 \). The right hand side (RHS) of the above ICs measures \( B \)'s net reputation cost if he reports \( m_B = 1 \) versus \( m_B = 0 \) given \( m_A \) and his signal \( s_B \). Biased \( B \)'s truthful reporting depends on how large the net benefit of distortion is relative to his net reputation cost. Because \( \Gamma_2 > \Gamma_1 > \Gamma_3 > \Gamma_4 \), we can rank the RHS of these ICs. For instance, \( \Delta_2 - \Delta_1 = (\Gamma_2 - \Gamma_1)(\tau_{0,0}^B - \tau_{1,0}^B + \tau_{1,1}^B - \tau_{0,1}^B) \), and other comparisons are similar.

The term \( (\tau_{0,0}^B - \tau_{1,0}^B + \tau_{1,1}^B - \tau_{0,1}^B) \) is positive if \( B \)'s expected posterior objectivity of giving correct recommendations \( (\tau_{0,0}^B + \tau_{1,1}^B) \) is larger than that of giving wrong recommendations \( (\tau_{0,1}^B + \tau_{1,0}^B) \); it is negative otherwise. One sufficient condition for \( (\tau_{0,0}^B - \tau_{1,0}^B + \tau_{1,1}^B - \tau_{0,1}^B) > 0 \) is if \( p_A \) is sufficiently close to 0.5; or if \( \theta_B \) is sufficiently close to 1. More specifically, for every \( p_A \), there exists a cutoff \( \bar{\theta}_B \) such that this condition holds if \( \theta_B > \bar{\theta}_B \). Also, \( \bar{\theta}_B = 0 \) if \( p_A \) is sufficiently close to 0.5; and \( \bar{\theta}_B \) is arbitrarily close to 1 if \( p_A \) is arbitrarily close to \( p_B \).

By assumption, \( p_A \) is sufficiently close to 0.5, and thus \( (\tau_{0,0}^B - \tau_{0,1}^B + \tau_{1,1}^B - \tau_{1,0}^B) > 0 \). Let the probabilities that biased \( B \) reports his signal \( s_B = 1 \) truthfully be:

\[
z_1 \equiv \Pr(m_B = 1|m_A = 0, s_B = 1)
\]

and

\[
z_2 \equiv \Pr(m_B = 1|m_A = 1, s_B = 1)
\]

If biased \( B \) reports both signals truthfully, then \( y_1 = y_2 = z_1 = z_2 = 1 \). The following lemma establishes claim (1) and (2) of Proposition 1.

**Lemma 1** In any (continuation) equilibrium, \( z_1 = 1, z_2 = 1, \) and \( y_1 \leq y_2 \). The inequality is strict if \( y_2 > 0 \). Moreover, \( y_1 \) and \( y_2 \) cannot both be strictly between 0 and 1.

**Proof:** Because \( (\tau_{0,0}^B - \tau_{0,1}^B + \tau_{1,1}^B - \tau_{1,0}^B) > 0 \), biased \( B \)'s net reputation cost can be ranked such that \( \Delta_2 > \Delta_1 > \Delta_3 > \Delta_4 \). Because the LHS of all the ICs is the same, it is not possible for more than one
IC to bind, thus at most one of the four probabilities \( y_1, y_2, z_1, z_2 \) can be strictly between zero and one. Also, truth telling cannot be part of the equilibrium. If \( y_1 = 1, y_2 = 1 \), the LHS of IC (1) is strictly positive, while the RHS of IC (1) and IC (2) are zero, which is impossible.

If any of the ICs binds, there are four possibilities. First, suppose that IC (3) binds, then \( 0 < z_1 < 1 \). This implies that \( z_2 = 1 \), and \( y_1 = y_2 = 1 \). This possibility corresponds to the case when biased \( B \) reports \( s_B = 0 \) truthfully but lies with some probability if \( s_B = 1, m_A = 0 \). In this case, \( C \) knows that \( s_B = 1 \) if \( m_B = 1 \), which implies that \( \hat{\eta}_1^B = p_B > \hat{\eta}_0^B \). Thus the net benefit for biased \( B \) to send \( m_B = 1 \) versus \( m_B = 0 \) is positive \( (\hat{\eta}_1^B - \hat{\eta}_0^B > 0) \). On the reputation side, however, we can show that \( \tau_{0,0}^B < \tau_{1,0}^B \) and \( \tau_{0,1}^B < \tau_{1,1}^B \). That is, biased \( B \)'s net reputation cost is negative. Together, biased \( B \) is strictly better off reporting \( m_B = 1 \), thus he will deviate, a contradiction. Intuitively, in this putative equilibrium, \( m_B = 1 \) results from signal \( s_B = 1 \), and it is more likely to come from objective \( B \), therefore it is both credible and a good sign of objectivity. Similarly, IC (4) binds and thus \( z_1 = 0, z_2 > 0 \) cannot be part of an equilibrium. The only remaining possibility is for either IC (1) or IC (2) to bind. In either case, \( z_1 = 1, z_2 = 1 \). If IC (1) binds, then \( y_1 > 0, y_2 = 1 \); and if IC (2) holds or binds, then \( y_1 = 0, y_2 \leq 1 \), thus \( 0 \leq y_1 < y_2 \). Finally, if none of the ICs binds, then \( z_1 = z_2 = 1, y_1 = y_2 = 0 \), in which case biased \( B \) always reports \( m_B = 1 \) regardless of his information. □

For claim (3) of Proposition 1, note that for biased \( B \), the net benefit from distortion is \( \hat{\eta}_1^B - \hat{\eta}_0^B \), which has the same sign, and is proportional to \( (2p_B - 1)N_{y_1} + (p_Bp_A - (1 - p_B)(1 - p_A))N_y(N_{y_2} - N_{y_1}) \). This benefit is positive because \( m_B = 1 \) is more likely to have come from \( s_B = 1 \). Thus \( C \) always takes a higher action after \( m_B = 1 \) than \( m_B = 0 \). Moreover, note that \( \hat{\eta}_1^B \) increases in \( y_2 \) and \( \hat{\eta}_0^B \) decreases in \( y_2 \), thus the net benefit strictly increases in \( y_2 \). Intuitively, if \( m_A = 0, s_B = 0 \), biased \( B \)'s distortion is the most inefficient, because neither signal supports his bias. Thus the more truthful he is, the more \( C \) changes her action.

Next, in terms of biased \( B \)'s benefit from distortion if \( m_A = 1, s_B = 0 \), we can show that both \( \hat{\eta}_1^B \) and \( \hat{\eta}_0^B \) increase in \( y_1 \). Here, \( \hat{\eta}_1^B \) increases in \( y_1 \) because message \( m_B = 1 \) is more likely to have come from \( s_B = 1 \); and \( \hat{\eta}_0^B \) increases in \( y_1 \) because the higher is \( y_1 \), the more likely that biased \( B \) does not use \( A \)'s information, which may result from \( s_A = 1 \). But when \( p_A \) is sufficiently close to 0.5 (or if \( \theta_B \) is sufficiently large), \( \hat{\eta}_1^B - \hat{\eta}_0^B \) increases in \( y_1 \). Also, the net reputation cost of distortion for biased \( B \) strictly decreases in \( y_1 \) and \( y_2 \). To see this, note that both \( \tau_{0,0}^B - \tau_{1,0}^B \) and \( \tau_{0,1}^B - \tau_{1,1}^B \) decrease in \( y_1, y_2 \):
the more truthful biased $B$ is, the less $C$ changes her estimates about $B$’s objectivity. Finally, simple calculations can show that $\tilde{\eta}_1^B - \hat{\eta}_0^B$ increases in $N_x$, and biased $B$’s net reputation cost decreases in $N_x$. By the implicit function theorem, if $y_1 \in (0, 1)$, it strictly decreases in $N_x$. Similarly, if $y_2 \in (0, 1)$, it strictly decreases in $N_x$. □

**Proof of Corollary 1:** Simple calculations can show that $\tilde{\eta}_1^B - \hat{\eta}_0^B$ increases in $p_A$, and the increase is proportional to $(1 - \theta_B)(y_2 - y_1)$. Differentiate biased $B$’s net reputation cost with respect to $p_A$:

$$\frac{\partial \Delta_1}{\partial p_A} = \frac{\partial \Gamma_1}{\partial p_A} [r_{0,0}^B - \tau_{1,0}^B - (\tau_{0,1}^B - \tilde{\tau}_{1,1}^B)] + \left[ \Gamma_1 \frac{\partial}{\partial p_A} (r_{0,0}^B - \tau_{1,0}^B) + (1 - \Gamma_1) \frac{\partial}{\partial p_A} (\tilde{\tau}_{0,1}^B - \tilde{\tau}_{1,1}^B) \right].$$

And that of $\Delta_2$ is very similar. Note that the first half is the information aggregation effect: it measures the change in biased $B$’s net reputation cost due to $m_A$. The second part is the reputation sensitivity effect: it measures how $B$’s posterior objectivity responds to the fact that biased $B$ may use $m_A$. Also, $\tau_{0,0}^B - \tau_{1,0}^B$ decreases in $p_A$ and $\tau_{0,1}^B - \tau_{1,1}^B$ increases in $p_A$. For $p_A$ sufficiently close to 0.5, the reputation sensitivity effect is strictly negative. Moreover, $\frac{\partial \Delta_1}{\partial p_A} < 0$, $\frac{\partial \Delta_2}{\partial p_A} > 0$. Thus if $y_1 \in (0, 1)$, by the implicit function theorem, $y_1$ decreases in $p_A$. If $y_2 \in (0, 1)$, then at $y_2$ sufficiently close to zero, we can use the indifference condition of biased $B$ and show that $y_2$ increases in $p_A$. □

**Proof of Proposition 2:** Given the distortion strategy of biased $A$ and $B$, biased $A$’s (expected) net benefit from reporting $m_A = 1$ instead of $m_A = 0$ given his signal is:

$$E_{m_B} [\Pr(\eta = 1|m_B)|m_A = 1, s_A = 0] - E_{m_B} [\Pr(\eta = 1|m_B)|m_A = 0, s_A = 0] = (1 - \theta_B)(y_2 - y_1)p_A p_B + (1 - p_A)(1 - p_B)(\tilde{\eta}_1^B - \hat{\eta}_0^B);$$

$$E_{m_B} [\Pr(\eta = 1|m_B)|m_A = 1, s_A = 1] - E_{m_B} [\Pr(\eta = 1|m_B)|m_A = 0, s_A = 1] = (1 - \theta_B)(y_2 - y_1)[p_A(1 - p_B) + p_B(1 - p_A)](\tilde{\eta}_1^B - \hat{\eta}_0^B).$$

Note that $m_A$ only influences $C$’s action if biased $B$ may change his recommendation based on what he hears. Recall that $\kappa \equiv (1 - \theta_B)(y_2 - y_1)$, if $\kappa = 0$, then biased $A$’s net benefit from distortion is 0.

Biased $A$ is also concerned about how objective $C$ thinks he is at the evaluation stage, given $B$’s recommendation and the observed true state. To simplify notations, let $A$’s posterior objectivity be $\tau_{m_B,\eta}^A$. Using Bayes’ rule, we have:

$$\tau_{0,0}^A \equiv \Pr(A = o|m_B = 0, \eta = 0) = \frac{[p_A N_{y_2} + (1 - p_A) N_{y_1}] \theta_A}{p_A N_x N_{y_2} + (1 - p_A N_x) N_{y_1}};$$
Clearly, if \( \beta \) is any way. Hence if \( \gamma = A_s \) in IC (6) holds strictly, he reports a cutoff value \( A_s \). Because \( \kappa = 0 \) and \( \eta(m) \equiv 1 \), because agent \( A_s \) is more likely to report \( \gamma = A_s \) truthfully, the following two incentive constraints must hold:

\[
\tau_{0,0}^A \equiv \Pr(A = o|m_B = 0, \eta = 0) = \frac{[p_A(1 - p_B N_{y_2}) + (1 - p_A)(1 - p_B N_{y_1})] \theta_A}{p_A N_x(1 - p_B N_{y_2}) + (1 - p_A N_x)(1 - p_B N_{y_1})};
\]

\[
\tau_{0,1}^A \equiv \Pr(A = o|m_B = 0, \eta = 1) = \frac{[(1 - p_A) N_{y_2} + p_A N_{y_1}] \theta_A}{(1 - p_A) N_x N_{y_2} + (1 - (1 - p_A) N_x) N_{y_1}};
\]

\[
\tau_{1,1}^A \equiv \Pr(A = o|m_B = 1, \eta = 1) = \frac{[(1 - p_A)(1 - (1 - p_B) N_{y_2}) + p_A (1 - (1 - p_B) N_{y_1})] \theta_A}{(1 - p_A) N_x (1 - (1 - p_B) N_{y_2}) + (1 - (1 - p_A) N_x)(1 - (1 - p_B) N_{y_1})}.
\]

Biased \( A_s \)'s net reputation cost if he reports \( m_A = 0 \) instead of \( m_A = 1 \) given his signal is:

\[
E_{\eta, m_B} [\Pr(A = o|m_B, \eta)|m_A = 0, s_A = 0] - E_{\eta, m_B} [\Pr(A = o|m_B, \eta)|m_A = 1, s_A = 0]
\]

\[
= (N_{y_2} - N_{y_1})[p_A p_B \tau_{0,0}^A + (1 - p_A)(1 - p_B) \tau_{0,1}^A] - (N_{y_2} - N_{y_1})[p_A p_B \tau_{1,0}^A + (1 - p_A)(1 - p_B) \tau_{1,1}^A];
\]

Because \( \kappa = N_{y_2} - N_{y_1}, \) if \( \kappa = 0 \), biased \( A_s \)'s net reputation cost is zero. If \( \kappa > 0 \), for biased \( A_s \) to report \( s_A = 0 \) and \( s_A = 1 \) truthfully, the following two incentive constraints must hold:

\[
\bar{\eta}_0^B - \bar{\eta}_0^B \leq \alpha \left[ \Pr(\eta = 0|s_A = 0, s_B = 0) \left( \tau_{0,0}^A - \tau_{0,1}^A \right) + \Pr(\eta = 1|s_A = 0, s_B = 0) \left( \tau_{0,1}^A - \tau_{1,1}^A \right) \right] ; \tag{5}
\]

\[
\bar{\eta}_1^B - \bar{\eta}_0^B \geq \alpha \left[ \Pr(\eta = 0|s_A = 1, s_B = 0) \left( \tau_{0,0}^A - \tau_{0,1}^A \right) + \Pr(\eta = 1|s_A = 1, s_B = 0) \left( \tau_{0,1}^A - \tau_{1,1}^A \right) \right]. \tag{6}
\]

The LHS of IC (5) and (6) is the same, and simple calculations can show that the RHS of IC (5) is always larger than the RHS of IC (6). Thus whenever biased \( A_s \) prefers reporting \( m_A = 1 \) if \( s_A = 0 \), he strictly prefers reporting \( m_A = 1 \) if \( s_A = 1 \). Intuitively, biased \( A_s \) has less to lose in terms of reputation if \( s_A = 1 \), because agent \( B \) is more likely to report \( s_B = 1 \) correctly, which is a less negative sign of \( A_s \)'s objectivity. Thus for biased \( A_s \), if neither IC holds, he always reports \( m_A = 1 \); and if IC (5) binds and IC (6) holds strictly, he reports \( m_A = 0 \) if \( s_A = 0 \) with some positive probability.

Next, at \( y_1 = y_2 = 0 \), biased \( B_s \)'s highest reputation cost occurs if \( s_B = 0, m_A = 0 \). Thus there exists a cutoff value \( \beta^c \) such that, at \( y_1 = y_2 = 0 \), IC (2) binds and biased \( B_s \)'s net benefit from distortion is equal to his net reputation cost from distortion:

\[
\frac{2p_B - 1}{2 - \theta_B} = \beta^c (1 - \theta_B) \left[ \frac{\Gamma_2}{1 - p_B \theta_B} + \frac{1 - \Gamma_2}{1 - (1 - p_B) \theta_B} \right].
\]

Clearly, if \( \beta \leq \beta^c \), biased \( B_s \) prefers reporting \( m_B = 1 \) with probability one. And since biased \( A_s \) has no influence on \( C \) and incurs no reputation cost, he is indifferent between his messages and he can report in any way. Hence if \( \beta \leq \beta^c, x \in [0,1], y_1 = 0, y_2 = 0 \) in equilibrium.
If $\beta > \beta^c$, then $y_1 = 0, y_2 = 0$ cannot be part of an equilibrium because biased $B$’s net benefit from distortion is strictly smaller than his net reputation cost. Thus he must report $m_B = 0$ truthfully with some probability. Also, for biased $B$, IC (1) implicitly defines $\tilde{y}_1(x)$, a function of $y_1$ with respect to $x$ when $y_2 = 1$. Similarly, IC (2) implicitly defines $\tilde{y}_2(x)$, a function of $y_2$ with respect to $x$ when $y_1 = 0$. Biased $A$’s IC (5) also defines two functions of $x$ with respect to $y_1$ and $y_2$: $\tilde{x}(y_1)$ if $y_2 = 1$; and $\tilde{x}(y_2)$ if $y_1 = 0$. Moreover, because the biased agents’ net benefit and net reputation cost are continuous in $x, y_1, y_2$, these best responses are continuous. From Proposition 1, both $\tilde{y}_1(x)$ and $\tilde{y}_2(x)$ decreases in $x$. Also, biased $A$ and $B$’s net benefit from distortion $\tilde{n}_1^B - \tilde{n}_0^B$ strictly increases in $y_1, y_2, x$. For biased $A$, his net reputation cost decreases in $x$. And because both $(\tau_{0,0}^A - \tau_{1,0}^A)$ and $(\tau_{0,1}^A - \tau_{1,1}^A)$ decrease in $y_1$ and increase in $y_2$, his net reputation cost decreases in $y_1$ and increases in $y_2$. By the implicit function theorem, $\tilde{x}(y_1)$ decreases in $y_1$.

To begin with, observe that if $\tilde{y}_2(0) \leq 1$, there does not exist a weak distortion equilibrium. In this case, biased $B$’s IC (2) does not hold at $x = 0$ and $y_1 = 0, y_2 = 1$, i.e., biased $B$ prefers reporting $m_B = 1$ if $s_B = 0, m_A = 0$. Because biased $B$’s net reputation cost $\Delta_2 > \Delta_1$, and $\tilde{y}_1(x)$ decreases in $x$, IC (1) cannot bind for any $x$, thus $y_1 > 0$ cannot be part of the equilibrium. One sufficient condition for this to occur is if $\beta \leq \beta^s$, where IC (2) binds at $\beta = \beta^s$, and $y_1 = 0, y_2 = 1, x = 0$. In this case, biased $B$’s best response to $x$, $y^{BR}(x)$, is simply $\tilde{y}_2(x)$ and biased $A$’s best response to $y_2$ is $x^{BR}(y_2)$. Also, $x^{BR}(y_2) = 0$ if $y_2$ is sufficiently close to 0, and $x^{BR}(1) < 1$. To see this, note that $y_2$ increases in $\beta$, everything else being equal. If $y_2$ is sufficiently close to 0, which occurs if $\beta$ is sufficiently close to $\beta^c$, then simple algebra can show that biased $A$’s net reputation cost (the RHS of IC (5)) is sufficiently close to 0, which is strictly smaller than his benefit from distortion (the LHS of IC (5)). Given that $y_1 = 0$, consider $x^{BR}(y_2)$ and $\tilde{y}_2(x)$. Because $x^{BR}(0) \in [0,1]$, $x^{BR}(1) < 1$; and $\tilde{y}_2(0) > 0, \tilde{y}_2(1) < \tilde{y}_2(0) \leq 1$, by the intermediate value theorem, the two best responses intersect. Thus if $\beta \in (\beta^c, \beta^s]$, there exists a strong distortion equilibrium in which $x \in [0,1], y_2 \in (0,1)$.

If $\tilde{y}_1(1) > 0$, however, there does not exist a strong distortion equilibrium. In this case, biased $B$’s IC (1) holds at $x = 1$ and $y_1 = 0, y_2 = 1$, i.e., biased $B$ prefers reporting $m_B = 0$ with some probability if $s_B = 0, m_A = 1$. Thus IC (2) holds strictly for any $y_2, x$. One sufficient condition for this to occur is if $\beta > \beta^w$, where the cutoff is defined such that at $\beta = \beta^w$, IC (1) binds at $y_1 = 0, y_2 = 1, x = 1$. Because IC (1) binds at $\beta > \beta^w$ and $x = 1, y_1 = 0, y_2 = 1$, IC (2) strictly holds at these values. Because the LHS
of IC (2) increases in \(x\) and the RHS decreases in it, \(\beta^s < \beta^w\). In this case, biased \(B\)'s best response \(y^{BR}(x)\) is simply \(\hat{y}_1(x)\) and biased \(A\)'s best response \(x^{BR}(y_1)\) is simply \(\hat{x}(y_1)\). Also, \(x^{BR}(y_1) = 0\) if \(y_1\) is sufficiently close to 1, which occurs if \(\beta\) is sufficiently large. To see this, note that \(y_1\) increases in \(\beta\).

If \(y_1\) is sufficiently close to 1, then simple algebra can show that biased \(A\)'s net reputation cost (the RHS of IC (5)) is sufficiently close to 0, which is strictly smaller than his benefit from distortion (the LHS of IC (5)). Given that \(y_2 = 1\), consider \(x^{BR}(y_1)\) and \(\hat{y}_1(x)\). Because \(x^{BR}(0) \geq 0\), \(x^{BR}(1) \in [0, 1]\); and \(\hat{y}_1(0) > 0\), \(\hat{y}_1(1) < 1\), by the intermediate value theorem, the two best responses intersect. Thus if \(\beta \geq \beta^w\), there exists a weak distortion equilibrium in which \(x \in [0, 1], y_1 \in (0, 1)\).

Finally, if \(\beta \in (\beta^s, \beta^w]\), then \(\hat{y}_2(0) > 1\) and \(\hat{y}_1(1) = 0\), then there exists a cutoff value \(\hat{x}\) such that

\[
y^{BR}(x) = \begin{cases} 
\hat{y}_1(x), & \text{if } x \leq \hat{x} \\
\hat{y}_2(x), & \text{if } x > \hat{x}.
\end{cases}
\]

As for biased \(A\), at \(y_1 = 0, y_2 = 1\), either his IC (5) fails to hold, or it binds at some \(x\). Let \(x' = 0\) in the former case and \(x' = x\) in the latter case. First, if \(x' \leq \hat{x}\), then \(0 < \hat{y}_1(x') < 1\), \(0 < \hat{y}_1(\hat{x}) < \hat{y}_1(x')\), \(x^{BR}(y_1) = 0\) for \(y_1\) sufficiently close to 1, \(x^{BR}(0) = x'\). Because of continuity, and by the intermediate value theorem, the two best responses intersect and there exists a weak distortion equilibrium. However, because \(x^{BR}(y_2)\) depends on parameter values, a strong distortion equilibrium cannot be ruled out.

Second, if \(x' > \hat{x}\), then because \(\hat{y}_2(\hat{x}) = 1\), \(0 < \hat{y}_1(x') < \hat{y}_1(\hat{x}) < 1\), and \(x^{BR}(y_2) = 0\) for \(y_2\) sufficiently close to 0, \(x^{BR}(1) = x'\), the two best responses intersect and there exists a strong distortion equilibrium. However, because the curvatures of \(x^{BR}(y_1)\) and \(\hat{y}_1(x)\) depend on parameter values, a weak distortion equilibrium cannot be ruled out. Thus if \(\beta \in (\beta^s, \beta^w]\), there may exist multiple equilibria. \(\square\)

**Proof of Proposition 3:** Suppose that \(x > 0, y_2 > 0\) in a strong distortion equilibrium. Recall that biased \(A\) and \(B\)'s incentive constraints IC (5) and IC (2) bind respectively in this case. These two ICs can be rewritten as: \(\xi^s(x, y_2, \alpha) = 0\) and \(\psi^s(x, y_2; \beta) = 0\). Let \(\xi_1^s, \xi_2^s, \xi_3^s\) respectively be the partial derivative of \(\xi^s\) with respect to \(x, y_2\) and \(\alpha\); and \(\psi_1^s, \psi_2^s\) are similarly defined. Differentiate with respect to \(\alpha\), then we have:

\[
\frac{dx}{d\alpha} = \frac{\xi_3^s \psi_2^s}{\xi_2^s \psi_1^s - \xi_1^s \psi_2^s}; \text{ and } \frac{dy}{d\alpha} = -\frac{\xi_3^s \psi_1^s}{\xi_2^s \psi_1^s - \xi_1^s \psi_2^s}.
\]

Signs of some of the above partial derivatives are straightforward, namely, \(\xi_1^s > 0, \psi_2^s > 0, \psi_1^s > 0, \xi_3^s < 0\). The key is the sign of \(\xi_3^s\). Differentiate the LHS and the RHS of IC (5), then use biased \(A\)'s indifference condition \(\xi^s(x, y_2, \alpha) = 0\) itself, we can show that if \(y_2\) is sufficiently close to 0, which occurs if \(\beta\) is
sufficiently close to $\beta^*$, or if $\theta_B$ is sufficiently large, $\xi_2 < 0$. Therefore in this case $y_2$ decreases in $\alpha$ while $x$ increases in $y_2$. □

**Proof of Proposition 4:** Suppose that $x > 0, y_1 > 0$ in a weak distortion equilibrium. Recall that biased $A$ and $B$’s incentive constraints IC (5) and IC (1) bind respectively in this case. Also, from the proof of Proposition 2, we know that biased $A$’s best response decreases in $y_1$ and biased $B$’s best response decreases in $x$. Therefore $x$ and $y_1$ are substitutes, $x$ increases in $\alpha$ while $y_1$ decreases in it; also $x$ decreases in $\beta$ while $y_1$ increases in it. □

**Proof of Proposition 5:** Let the weights on biased $i, j$’s reputation be $\alpha_i, \alpha_j$ respectively. Similarly, the signal quality is $p_i, p_j$ respectively and their truthful reporting probabilities be $x_i, x_j$ respectively. Recall from the text that here biased $i = A, B$ maximizes:

$$E_{m_j}[\Pr(\eta = 1|m_i, m_j)|s_i] + E_{\eta}[\Pr(i = o|m_i, \eta)|s_i].$$

To simplify notations, denote $\beta$’s action given the messages $\Pr(\eta = 1|m_i, m_i)$ be $\Pr(1|m_i, m_j)$. Then for biased $i$ to report $s_i = 0$ and $s_i = 1$ truthfully, the following two incentive constraints must hold:

$$\Pr(m_j = 1|s_i = 0)[\Pr(1|1, 1) - \Pr(1|0, 1)] + \Pr(m_j = 0|s_i = 0)[\Pr(1|1, 0) - \Pr(1|0, 0)]$$

$$\geq \alpha_i \sum_{\eta} \Pr(\eta|s_i = 0)[\Pr(i = o|m_i = 0, \eta) - \Pr(i = o|m_i = 1, \eta)];$$

(7)

$$\Pr(m_j = 1|s_i = 1)[\Pr(1|1, 1) - \Pr(1|0, 1)] + \Pr(m_j = 0|s_i = 1)[\Pr(1|1, 0) - \Pr(1|0, 0)]$$

$$\geq \alpha_i \sum_{\eta} \Pr(\eta|s_i = 1)[\Pr(i = o|m_i = 0, \eta) - \Pr(i = o|m_i = 1, \eta)].$$

(8)

Let $N_i \equiv \theta_i + (1 - \theta_i)x_i, N_j \equiv \theta_j + (1 - \theta_j)x_j$. Then given the distortion strategies, the decisionmaker’s action after hearing both messages becomes:

$$\Pr(1|0, 0) = \frac{(1 - p_i)(1 - p_j)}{p_ip_j + (1 - p_i)(1 - p_j)};$$

$$\Pr(1|0, 1) = \frac{(1 - p_i)[1 - (1 - p_j)N_j]}{(1 - p_i)[1 - (1 - p_j)N_j] + p_i[1 - p_jN_j]};$$

$$\Pr(1|1, 0) = \frac{(1 - p_j)[1 - (1 - p_i)N_i]}{(1 - p_j)[1 - (1 - p_i)N_i] + p_j[1 - p_iN_i]};$$

$$\Pr(1|1, 1) = \frac{[1 - (1 - p_j)N_j][1 - (1 - p_i)N_i]}{[1 - (1 - p_j)N_j][1 - (1 - p_i)N_i] + [1 - p_jN_j][1 - p_iN_i]}.$$

Moreover, because of the presence of objective agent, it is simple to show that $\Pr(1|1, 1) > \Pr(1|0, 1)$ and $\Pr(1|1, 0) > \Pr(1|0, 0)$. Next, it can also be shown that the difference $\Pr(1|1, 1) - \Pr(1|0, 1) -$
Intuitively, $m_i = 1$ has a higher marginal impact on $C$ if $m_j$ supports rather than contradicts $m_i$. Because biased $i$ is more likely to distort when $\eta = 0$, in which case his signal is more likely to be $s_A = 0$ than when $\eta = 1$.

This also implies that if IC (7) binds or fails to hold, IC (8) holds strictly. The reason is that if $s_i = 1$, it is more likely that $s_j = 1$ as well, thus the benefit from distortion is higher for biased $i$ to report $m_j = 1$ as opposed to $m_i = 0$. Moreover, the net reputation cost is smaller for $i$ if $s_i = 1$ because $m_i = 1$ is more likely to be correct. Thus biased $i$ always reports $m_i = 1$ truthfully. Also, if biased $i$’s weight on his reputation $\alpha_i$ is sufficiently small, IC (7) cannot hold, thus biased $i$ always report $m_i = 1$. Otherwise, IC (7) binds. Differentiate biased $i$’s net benefit from distortion (the LHS of IC (7)) with respect to $N_j$, we have:

$$\frac{\partial}{\partial N_j} \Pr(1|1,1) - \frac{\partial}{\partial N_j} \Pr(1|0,1)$$

For the discussion above, the second line is negative. Moreover, $\frac{\partial}{\partial N_j} \Pr(1|1,1) < \frac{\partial}{\partial N_j} \Pr(1|0,1)$. Thus biased $i$’s net benefit from distortion is decreasing in $x_j$. Together, $x_i$ increases in $x_j$. The biased agents’ truth-telling probabilities are complements because if one becomes more truthful, the other has a smaller (expected) marginal impact on $C$, thus he lies less as well. □

**Proof of Proposition 6:** consider simultaneous reporting first. Suppose that in equilibrium, biased $A$ reports $m_A = 0$ with positive probability $x_A$, then his ex ante expected payoff becomes:

$$E_{U}^{1}_A \equiv E_{s_A}[E_{m_B}[\Pr(\eta = 1|m_A, m_B)|s_A] + \alpha E_{\eta}[\Pr(A = o|m_A, \eta)|s_A]]$$

$$= 0.5[\theta_B + (1 - \theta_B)x_B)\Pr(1|1,1) + (\theta_B + (1 - \theta_B)x_B)\Pr(1|0,1)]$$

$$+ \alpha \Pr(A = o|m_A = 1). \quad (9)$$

Because biased $A$ is indifferent between reporting $m_A = 1$ or $m_A = 0$ if $s_A = 0$, while he strictly prefers reporting $m_A = 1$ if $s_A = 1$, his expected payoff is the same as if he always reports $m_A = 1$. Moreover, compare biased $A$’s ex ante payoff with a sum of $C$’s prior beliefs $0.5 + \alpha \theta_A$:

$$E_{U}^{1}_A - (0.5 + \alpha \theta_A)$$

$$= E_{U}^{1}_A - \sum_{l,k \in \{0,1\}} \Pr(m_A = l, m_B = k)\Pr(\eta = 1|m_A = l, m_B = k) - \sum_{l \in \{0,1\}} \alpha \Pr(m_A = l)\Pr(A = o|m_A = l) \quad (10)$$
\[
= \Pr(m_A = 0) \left[ \Pr(m_B = 1|s_A = 0)\Pr(1|1) - \Pr(1|0, 1) \right] + \Pr(m_B = 0|s_A = 0)\Pr(1|1, 0) - \Pr(1|0, 0) \\
- \alpha \Pr(m_A = 0) \left[ \Pr(A = o|m_A = 0) - \Pr(A = o|m_A = 1) \right]
\]
\[
= 0.
\]

The first equality is due to the law of iterated expectations, while the last one is due to biased A’s indifference condition IC (7), as given in Proposition 5. Therefore if biased A cares sufficiently about his reputation to report \( m_A = 0 \) sometimes, his expected payoff is simply the sum of C’s prior beliefs.

Second, suppose that biased A communicates through intermediary B, and he has influence in equilibrium \( \max\{y_1, y_2\} > 0 \). If \( x > 0 \) in equilibrium, biased A’s ex ante expected payoff if he sends a message to B is:

\[
EU^2_A = 0.5[1 + (1 - \theta_B)(1 - y_1)][\hat{\eta}_1^{B} + \alpha \tau_{1,1}^A + \alpha \tau_{1,0}^A] + 0.5[\theta_B + (1 - \theta_B)y_1][\hat{\eta}_0^{B} + \alpha \tau_{0,1}^A + \alpha \tau_{0,0}^A].
\]

Similar argument can show that \( EU^2_A = 0.5 + \alpha \theta_A \) in this case. Thus if \( x_A > 0, x > 0, \max\{y_1, y_2\} > 0 \), biased A is indifferent in ex ante terms between these two types of communication.

Next, if biased A strictly prefers lying, then his ex ante expected payoff is strictly higher than the prior \( 0.5 + \alpha \theta_A \). For instance, suppose that \( x_A = 0 \). Then biased A’s reputational cost is so low that \( EU^1_A(m_A = 1|s_A = 0) > EU^1_A(m_A = 0|s_A = 0) \) at \( x_A = 0 \), thus he is worse off if he reports truthfully with any infinitesimally small probability. To see this, note that expression (10) is strictly positive because IC (7) no long binds. Recall from Proposition 5 that biased A always reports \( m_A = 1 \) if \( \alpha \) is sufficiently low. A simple comparison of his expected payoffs can show that if \( \beta \) is sufficiently low that \( y_1 = 0 \), then biased A prefers simultaneous reporting if \( \theta_B \) is sufficiently close to 0. Formally, this is because \( \Pr(\eta = 1|1, 1) > \Pr(\eta = 1|m_B = 1) \). Otherwise, biased A prefers using intermediary B.

Finally, if \( \beta \) is sufficiently high, biased A always reports \( m_A = 1 \) for any given \( \alpha \) if he communicates through B. Therefore biased A with moderately high reputational concerns—who would report \( s_A = 0 \) truthfully with some probability if he sends his own message—always prefers communicating through B if biased B is very concerned about his reputation. □

References


