Biased Recommendations from Biased and Unbiased Experts

Wonsuk Chung and Rick Harbaugh*

This Version: April 2017

Abstract

When can an expert be trusted to provide useful advice? We experimentally test a simplified recommendation game where an expert recommends one of two actions to a decision maker who may take either action or instead pursue an outside option. Consistent with predictions from the cheap talk literature, we find that decision makers partially *discount* recommendations for the action a biased expert favors, but that recommendations can still be *persuasive* in that they reduce the chance of the outside option. If the decision maker is uncertain whether the expert is biased toward an action, biased experts lie even more, while unbiased experts follow a *political correctness* strategy of recommending the opposite action so as to be more persuasive by appearing unbiased. Even if experts are known to be unbiased, experts *pander* by recommending the action that the decision maker already favors, and decision makers discount the recommendation. The results highlight that transparency of expert incentives can improve communication, but need not ensure unbiased advice.

JEL Classification: D82, C92, M3.

Key Words: cheap talk, persuasion, transparency, pandering, political correctness

---

*Chung: Korea Insurance Research Institute, woschung@indiana.edu; Harbaugh: Department of Business Economics and Public Policy, Kelley School of Business, Indiana University, riharbau@indiana.edu. For helpful comments we thank Alexei Alexandrov, Tim Cason, Archishman Chakraborty, Bill Harbaugh, Tarun Jain, Wooyoung Lim, Dmitry Lubensky, Stephen Morris, Michael Raith, Eric Rasmusen, Eric Schmidbauer, Irene Skricki, Joel Sobel, Yossi Spiegel, Lise Vesterlund, Jimmy Walker, Stephanie Wang, Alistair Wilson, and seminar and conference participants at the University of Pittsburgh, Michigan State University, Booth School of Business, Simon School of Business, Indian School of Business, the Western Economics Association meetings, the Fifth Conference on the Economics of Advertising and Marketing at Qinghua University's School of Management and Economics, and the 2013 Boulder Summer Conference on Consumer Financial Decision Making.

1 Introduction

When an expert advises a decision maker, the expert may benefit from some choices more than others, such as a salesperson who earns a higher commission on some products. Can a recommendation still be persuasive with such a conflict of interest, or will it be completely discounted? How is communication affected if the expert is suspected to benefit more from one choice, but the decision maker is not sure? And what if the decision maker is already leaning toward a choice, such as a customer who is known to favor a particular product?

Understanding these issues is important to companies, institutions, and regulators that structure the incentive and information environment in which experts provide advice. In recent years the incentives of mortgage brokers to recommend high cost loans, of credit rating agencies to overrate risky bonds, of stock analysts to push their client's stocks, of medical researchers to promote certain drugs, and of doctors to recommend expensive treatments have all come under scrutiny. Can such problems be resolved by requiring disclosure of any conflicts of interest, or is it necessary to eliminate biased incentives? And are unbiased incentives always sufficient to ensure unbiased advice?[1]

To gain insight into such questions, several papers have applied the cheap talk approach of Crawford and Sobel (1982) to discrete choice environments where an expert has private information about different actions and the decision maker has an outside option that is the expert's least favored choice (e.g., De Jaegher and Jegers, 2001; Chakraborty and Harbaugh, 2010; Che, Dessein and Kartik, 2013).[2] Based on this literature, we develop and test a simplified recommendation game where an expert knows which of two actions is better for a decision maker. Both sides benefit if the expert can credibly indicate which action is better, but the expert's incentive to lie can undermine their credibility to the detriment of both sides. For instance, a salesperson might benefit more from pushing one product than another, and both sides suffer if a suspicious customer buys nothing. Despite its simplicity, the model captures several key phenomena from the literature.

First, for sufficient payoff symmetry recommendations are "persuasive" in that they benefit the expert by reducing the chance that the decision maker walks away without taking either action. Even though a recommendation is only cheap talk it is still credible since it raises

[1] Regulations can impose more equal incentives, e.g., requirements for "firewalls" that limit the incentive of stock analysts to push their firm's clients, and incentives may also be adjusted voluntarily to increase credibility, e.g., Best Buy promotes its "Non-commissioned sales professionals" whose "first priority is to help you make the right purchasing decision". Similarly, conflict of interest disclosure may be imposed as the SEC does for investment advisors, or voluntarily adopted as many medical journals have done for authors.

[2] Also see Chakraborty and Harbaugh (2007), Bolton, Freixas, and Shapiro (2007), and Inderst and Ottaviani (2012).

the expected value of one action at the opportunity cost of lowering the expected value of the other action, and it is persuasive since the higher expected value of the recommended action is now more likely to exceed the decision maker's outside option. For instance, a customer is more likely to make a purchase if a recommendation persuades him that at least one of two comparably priced products is of high quality. In our experimental results we find that recommendations are usually accepted, and are almost always accepted when the decision maker's outside option is unattractive.

Second, when the expert is biased in the sense of having a stronger incentive to push one action, a recommendation for that action is "discounted" in that the decision maker is more likely to ignore the recommendation and stick with the outside option. Therefore in equilibrium the expert faces a tradeoff where one recommendation generates a higher payoff if it is accepted but it is less likely to be accepted, while the other recommendation generates a lower payoff if it is accepted but it is more likely to be accepted.[3] Consistent with theoretical predictions, in our experiment we find that experts are significantly more likely to lie and recommend the more incentivized action even when it is worse for the decision maker, and that decision makers are significantly less likely to accept a recommendation for that action.

Third, when the decision maker is known to already favor one action, the expert benefits by "pandering" to the decision maker and recommending that action even when the other action really is better. Hence biased recommendations can result even when the expert's incentives for each action are the same (Che, Dessein, and Kartik, 2013). The decision maker anticipates such pandering and, just as in the asymmetric incentives case, discounts a recommendation for that action. In our experiment we find that experts are significantly more likely to lie and recommend the favored action even when they know it is the worse action, and decision makers are significant more likely to discount such a recommendation than when the prior distribution of decision maker values is symmetric.

Finally, when the decision maker is unsure of whether the expert is biased toward an action, a recommendation for that action is suspicious and hence discounted by the decision maker, so an unbiased expert has a "political correctness" incentive to recommend the opposite action even if it is not the best action (c.f., Morris, 2001). For instance, if a salesperson is suspected to benefit more from selling one product than another product but in fact has equal incentives, then pushing the other product is more likely to generate a sale. In our experiment we find that, as predicted, lack of transparency induces unbiased experts to make

[3] The Crawford-Sobel model also captures discounting in that equilibrium expected values of the state conditional on messages are below those implied by a symmetric, non-strategic interpretation of the messages. However the equilibrium tradeoff in the Crawford-Sobel model is different as it is driven by reluctance of the expert to recommend a higher action when the true state is sufficiently low.

the opposite recommendation as biased experts. Decision makers do not appear to sufficiently discount such recommendations, suggesting that they do not always anticipate how lack of transparency warps the incentives of even unbiased experts.

Based on the experimental literature we expect subjects to be reluctant to lie,[4] with the strength of this aversion varying across subjects (Gibson, Tanner, and Warner, 2013). Therefore we depart from a "pure" cheap talk approach and assume that experts have a lying cost that is drawn from a distribution with support that ranges from zero up to being so high as to preclude lying.[5] Such lying aversion could reflect a true preference against lying or be a reduced form for other concerns such as a fear of getting caught. Inclusion of heterogeneous lying aversion not only reduces the equilibrium amount of lying, it also makes the model more testable by eliminating additional equilibria such as strategic mixing between messages,[6] uninformative babbling, and using messages to imply the opposite of their literal meanings.

In our experimental tests we cannot control for varying subject preferences against lying, so the exact lying rates and acceptance rates cannot be predicted beforehand. However, we find that the comparative static predictions of the model for the phenomena we focus on are the same for any distribution of lying costs, so we can test these predictions even without knowing the exact distribution of subject preferences against lying. Moreover, since communication incentives in the model with lying costs are still driven primarily by the endogenous opportunity cost of messages, these comparative static predictions are the same as in the most intuitive equilibrium of the limiting case of pure cheap talk, so the predictions are consistent with the insights generated by the theoretical cheap talk literature.

The "deception game" literature following Gneezy (2005) tests a game where the expert and decision maker payoffs are opposed, so lying aversion is necessary for communication.[7] In our game lying aversion reduces lying but is not necessary for communication, and different strengths of lying aversion have no effect on the comparative static predictions.[8] Method-

[4] Such aversion is consistent with the empirical pattern of overcommunication in cheap talk games (Dickhaut et al., 1995; Cai and Wang, 2006; Sánchez-Pagés and Vorsatz, 2007) and also with more direct evidence on pupil dilation in such games (Wang, Spezio, and Camerer, 2010).

[5] Since messages have a direct effect on payoffs as in a (costly) signaling game, the model is a form of "costly talk" (Kartik, Ottaviani, and Squintani, 2007) or, as lying costs become arbitrarily small, a form of "almost cheap talk" (Kartik, 2009). If costs preclude lying the game is a disclosure/persuasion game (Milgrom, 1981).

[6] With no lying aversion or homogeneous lying aversion the expert is indifferent in equilibrium and uses a mixed strategy of lying or not. Heterogeneous lying aversion purifies these strategies since experts with lower lying costs strictly prefer to lie while those with higher lying costs strictly prefer to tell the truth.

[7] Unlike in Sánchez-Pagés and Vorsatz (2007), the decision maker is not told that payoffs are opposed, so heterogeneous beliefs can make inferences about player preferences and strategies difficult (Sutter, 2009).

[8] We do not consider the range of factors that can affect lying aversion. Gneezy (2005) and others find that subject reluctance to lie increases with the (expected) impact of the lie on the decision maker, but for simplicity

ologically we differ from the deception game literature in that we give the subjects common knowledge information about the distribution of payoffs so we can analyze how communication changes as the expert and decision maker strategically respond to changes in the information and payoff structure.

Most of the experimental literature on strategic information transmission via cheap talk has focused on testing different implications of the original Crawford and Sobel (1982) model in which the decision maker can take a range of possible actions from low to high and the expert has a limited bias toward a higher action (e.g., Dickhaut, McCabe and Mukherji, 1995; Cai and Wang, 2006; Wang, Spezio and Camerer, 2010).[9] This literature has been extended to testing results by Battaglini (2002) for cheap talk by competing experts in a multi-dimensional version of the Crawford-Sobel model (Lai, Lim, and Wang, 2015; Vespa and Wilson, 2016). In a costly communication model, Lafky and Wilson (2015) test how incentives for information provision can both encourage communication and make it less credible. Their misaligned incentives treatment has a similar structure and similar findings as our treatment with a biased expert.

Our theoretical and experimental results on the benefits of transparency contrast with the mostly negative results in the literature based on different models. Cain, Loewenstein, and Moore (2005) and others find experimentally that subjects do not fully discount recommendations by biased experts when the biases are disclosed, and that disclosure can even degrade communication by allowing experts to feel more license to exaggerate. Li and Madarász (2008) find theoretically that disclosure of the expert's bias in the Crawford-Sobel model often reduces communication. In a model that is closer to ours, Inderst and Ottaviani (2012) find that a salesperson's commissions are endogenously lower when they are disclosed, but also find that disclosure can inefficiently induce commissions to favor the higher cost producer.

The model's predictions are related to the literature on credence goods which examines recommendations to buy a cheap or expensive version of a product (Darby and Karni, 1973).[10] Building on the Pitchik and Schotter (1987) model, De Jaegher and Jegers (2001) consider a doctor who recommends either a cheap or expensive treatment to a patient whose condition is severe or not, where an expensive treatment works for both conditions but a cheap treatment works only if the condition is not severe. For some parameter ranges their model has a mixed

we assume that all lies are equally costly. See Sobel (2016) for the distinction between lying and deception. Abeler, Nosenzo, and Raymond (2016) find that reputational costs are an important factor in lying aversion. To reduce this factor, in our experiment we use random matching and emphaze the anonymity of the procedure to subjects.

[9] Blume, DeJong, Kim, and Sprinkle (1998, 2001) test the ability of players to successfully coordinate on the equilibrium meaning of messages when preferences are aligned or partly aligned. There is also a large literature on pre-play communication about strategic intentions in games with complete information (Crawford, 2003).

[10] Dulleck, Kerschbamer, and Sutter (2011) test the problem in a pricing rather than cheap talk context.

strategy equilibrium with aspects of pandering and discounting since a cheap treatment is more attractive for the patient and an expensive treatment is more lucrative for the doctor.

The predictions are also related to the literature on cheap talk in repeated games with a binary decision and no outside option. Sobel (1985) shows that recommendations are discounted based on the strength of the expert's incentive to push the favored action relative to reputational costs. Gentzkow and Shapiro (2006) show that an expert panders to the decision maker's prior beliefs to maintain credibility for future advice. Morris (2001) shows how an unbiased expert avoids making the same recommendation as a biased expert so as to maintain a reputation for not being biased. Our one-period model captures similar tradeoffs in an easily testable framework. Instead of a future reputational loss to being perceived as biased, there is an immediate loss that the decision maker is more likely to take the outside option.

In the following section we present the simplified recommendation game and provide testable hypotheses based on its equilibrium properties. We then outline how we test the hypotheses experimentally, and finally we report on the experiment results.

2 Recommendation Game

A decision maker must choose one of two actions, A or B, or an outside option C with respective values to the decision maker of v_A, v_B, and v_C. An expert knows v_A and v_B but the decision maker only knows that one is good and one is bad with equal chance, $\Pr[(v_A, v_B) = (a, 0)] = \Pr[(v_A, v_B) = (0, b)] = 1/2$ for given $0 < a, b \leq 1$.[11] The decision maker knows v_C but the expert only knows that it is independently and uniformly distributed on $[0, 1]$. The expert receives $\pi_A > 0$ or $\pi_B > 0$ if the decision maker chooses action A or B respectively, but receives $\pi_C = 0$ if the outside option C is chosen.

The expert first sends a message $m \in \{m_A, m_B\}$. Sending m_A when $v_B > v_A$ or sending m_B when $v_A > v_B$ incurs a lying cost d. The expert knows his own lying cost but the decision maker only knows that d comes from a distribution G. In the special case where $d = 0$ with certainty the game is a "pure cheap talk" game. The decision maker observes the message m, learns the value v_C, and then chooses the action A or B or the outside option C to maximize her expected payoffs given her beliefs.

Our equilibrium concept is Perfect Bayesian Equilibrium so along the equilibrium path

[11] This simplifying two-state assumption captures the feature from a two-dimensional quality space that exact equality of both values is a zero probability event. With multiple dimensions comparative statements across dimensions are credible even when absolute statements within each dimension are not (Chakraborty and Harbaugh, 2007, 2010), but with only two states any statement is inherently comparative so we cannot test the persuasiveness of comparative versus absolute statements.

the decision maker's beliefs follow Bayes Rule based on the expert's communication strategy which maps the expert's information (the realization of (v_A, v_B) and also the realization of d when it is uncertain) to the message space. Given the expert's communication strategy let the probabilities of a "lie" or "false claim" be $\alpha = \Pr[m_A|v_B > v_A]$ and $\beta = \Pr[m_B|v_A > v_B]$.

Suppose, as we will confirm below for sufficient symmetry, that in equilibrium the expected value of the recommended action exceeds that of the unrecommended action so the decision maker takes the recommended action if and only if its expected value exceeds v_C. The respective probabilities that recommendations for actions A and B are accepted are then

$$\begin{aligned} P_A &= \Pr[v_C \leq E[v_A|m_A]] = E[v_A|m_A] = \Pr[v_A > v_B|m_A]a = \frac{1-\beta}{1-\beta+\alpha}a \\ P_B &= \Pr[v_C \leq E[v_B|m_B]] = E[v_B|m_B] = \Pr[v_B > v_A|m_B]b = \frac{1-\alpha}{1-\alpha+\beta}b. \end{aligned} \tag{1}$$

We first analyze the special case of pure cheap talk where is no aversion to lying, $d = 0$. If one message offers a higher expected payoff then the expert will always send it, but then the decision maker will heavily discount such a message, thereby reducing its expected payoff. In equilibrium this discounting must equalize the payoffs from each message,

$$\pi_A P_A = \pi_B P_B. \tag{2}$$

Without loss of generality we focus on the case of $\pi_A a \geq \pi_B b$ so that the overall incentive and value asymmetry favors A. Consider a conjectured equilibrium where the expert might claim that A is better when it is not, $\alpha \geq 0$, but never claims B is better when it is not, $\beta = 0$. Then, from (1), the probability that A is really better when it is claimed to be better is $\Pr[v_A > v_B|m_A] = 1/(1+\alpha)$ and the probability that B is really better when it is claimed to be better is $\Pr[v_B > v_A|m_B] = 1$, so the acceptance probabilities are

$$P_A = \frac{a}{1+\alpha},\ P_B = b. \tag{3}$$

Substituting and solving for α from (2), in this equilibrium the expert lies in favor of A and B with respective probabilities

$$\alpha = \frac{\pi_A a}{\pi_B b} - 1,\ \beta = 0. \tag{4}$$

Notice that in our environment with an uncertain v_C the acceptance and lying probabilities P_A and α each depend on expert payoffs and decision maker preferences.[12]

[12] If v_C is fixed each side randomizes to keep the other side indifferent, so that the expert's lying rate depends only on decision maker preferences, $\alpha = a/v_C - 1$, and the decision maker's acceptance rate depends only on expert payoffs, $P_A = \pi_B/\pi_A$. Hence allowing for uncertain v_C as in a standard discrete choice model both purifies the decision maker's strategy and implies more intuitive behavior. We introduce uncertain lying costs below to purify the expert's strategies and eliminate unintuitive equilibria.

We impose two symmetry conditions. First is a recommendation condition ensuring that $\alpha \in [0,1]$, which from (4) holds if and only if $\pi_A a/\pi_B b \in [1,2]$. Second is an acceptance condition ensuring that, when the recommendation condition is satisfied, the expected value of the recommended action exceeds that of the unrecommended action, $\Pr[v_A > v_B|m_A]a \geq \Pr[v_B > v_A|m_A]b$ and $\Pr[v_B > v_A|m_B]b \geq \Pr[v_A > v_B|m_B]a$. Substituting α and β from (4), these reduce to $\pi_A a/\pi_B b \leq (a+b)/b$ and $b \geq 0$. Therefore, combining the recommendation and acceptance conditions when incentives and values weakly favor A, we assume

$$1 \leq \frac{\pi_A a}{\pi_B b} \leq \min\{\frac{a+b}{b}, 2\}. \tag{5}$$

We also use this restriction in our experimental parameterizations.[13]

Proposition 1 *For incentives and values satisfying (5), there exists a pure cheap talk equilibrium with acceptance and lying rates given by (3) and (4).*

Other pure cheap talk equilibria can also exist, including a babbling equilibrium in which the expert randomizes to reveal no information. However, the only equilibrium that survives when we allow for any lying aversion has the same main properties and same testable implications as the above equilibrium.

Turning to our main case of lying aversion, we assume that the expert knows his own lying cost d but the decision maker only knows that d has a common knowledge distribution G with no mass points and with support on $[0, \overline{d}]$ for some $\overline{d} > 0$. Since experts have heterogeneous lying costs, if there is any incentive to lie then those experts with lower lying costs lie while those with higher lying costs tell the truth, i.e., we are now interested in pure strategies rather than mixed strategies for the expert.

Continuing to focus on the case where A is favored, conjecture an equilibrium where the expert never lies in favor of B, but some α fraction of experts lies in favor of A. Letting d_A be the lying cost of the marginal expert type who is just indifferent between lying in favor of A or not, the equilibrium condition is

$$\pi_A P_A - d_A = \pi_B P_B \tag{6}$$

where P_A and P_B are still given by (3). Substituting in P_A and P_B, and noting that $d_A(\alpha)$ is the quantile function $G^{-1}(\alpha)$, in equilibrium

$$\alpha = \frac{\pi_A a}{\pi_B b + d_A(\alpha)} - 1, \, \beta = 0 \tag{7}$$

[13] If the symmetry conditions are violated then communication can be more complicated or break down in our simplified two-state game. In a richer state space, comparative cheap talk of the form in this paper remains credible and influential for arbitrary asymmetries (Chakraborty and Harbaugh, 2010; Chung, 2012).

where α and β now represent the fraction of expert types with a pure strategy of lying toward A and B respectively.

Notice that d_A is continuous and strictly increasing in α since G is continuous and strictly increasing in α, so the α that solves (7) is unique for $\beta = 0$. Also note that if (6) holds then $\pi_A P_A > \pi_B P_B$ so the expert strictly prefers to tell the truth when A is better, so it must be that $\beta = 0$. This leaves the only alternative equilibrium possibility as having $\alpha = 0$ and $\beta > 0$, but this requires $\pi_A a < \pi_B b$. Hence the equilibrium is unique. Imposing our same symmetry conditions as for pure cheap talk, we have the following for costly cheap talk due to lying aversion.[14]

Proposition 2 *For incentives and values satisfying (5), the unique costly cheap talk equilibrium has acceptance and lying rates given by (3) and (7).*

Lying aversion has the qualitative effect of selecting the simple type of cheap talk equilibrium analyzed in Proposition 1, and the quantitative effect of reducing the amount of lying. If the mass of the lying cost distribution is sufficiently concentrated near zero, then lying and acceptance rates are arbitrarily close to the pure cheap talk case. As lying aversion increases in the sense that the new distribution G stochastically dominates the original distribution, the marginal type $d_A(\alpha)$ that satisfies the equilibrium condition (6) becomes larger, so lying decreases and P_A increases accordingly. For instance, if G is uniform on $[0, \overline{d}]$, then $d_A(\alpha) = \alpha\overline{d}$, so the equilibrium lying rate α is decreasing in $\overline{d}$ from (7). For our purposes we will take G as given and test comparative static predictions for changes in the expert's incentives and the decision maker's values.

3 Testable Implications

We now assume that condition (5) holds and consider testable predictions for how expert incentives and decision maker values affect behavior. In the text we use Proposition 1 to provide intuitive proofs for the pure cheap talk case, and in the Appendix we use Proposition 2 to show that the same testable implications extend to our main case of lying aversion. We will then use the four parameterizations shown in Table 1 to experimentally test the predictions. The experimental procedures and results are discussed in Section 4. As seen in Table 1, the

[14] Lying aversion reduces α so it relaxes the symmetry conditions. For $\pi_A a \geq \pi_B b$, the recommendation condition from (7) is $d_A(1) \geq \frac{\pi_A a}{2} - \pi_B b$. For $a \geq b$ this also satisfies the acceptance condition. For $a < b$, the acceptance condition substituting in (7) is $d_A(\frac{a}{b}) \geq \frac{\pi_A ab}{a+b} - \pi_B b$. Hence the equilibrium in Proposition 2 holds for any asymmetries if lying aversion if $\overline{d}$ is sufficiently large.

parameter values are set so that, except for the Symmetric Baseline treatment, condition (5) holds weakly so the expert would always lie if lying costs were zero, but some expert types do not lie if there is any lying aversion at all.

	Choose A	Choose B	Choose C
A Better	$.8, 1$	$.8, 0$	$0, U[0,1]$
B Better	$.8, 0$	$.8, 1$	$0, U[0,1]$

$\pi_A = \pi_B = .8; a = b = 1$

Symmetric Baseline

	Choose A	Choose B	Choose C
A Better	$1, 1$	$.5, 0$	$0, U[0,1]$
B Better	$1, 0$	$.5, 1$	$0, U[0,1]$

$\pi_A = 1, \pi_B = .5; a = b = 1$

Asymmetric Incentives

	Choose A	Choose B	Choose C
A Better	$.8, 1$	$.8, 0$	$0, U[0,1]$
B Better	$.8, 0$	$.8, .5$	$0, U[0,1]$

$\pi_A = \pi_B = .8; a = 1, b = .5$

Asymmetric Values

	Choose A	Choose B	Choose C
A Better	$.8/1, 1$	$.8/.5, 0$	$0, U[0,1]$
B Better	$.8/1, 0$	$.8/.5, 1$	$0, U[0,1]$

$\pi_A = \pi_B = .8$ or $\pi_A = 1, \pi_B = .5; a = b = 1$

Opaque Incentives

Table 1: Expert and Decision Maker Payoffs

3.1 Symmetric Baseline

To see how communication can be persuasive, and to provide a baseline for testing the effects of asymmetries, suppose $\pi_A = \pi_B = \pi$ and $a = b = v$ for some $\pi > 0$ and $v \in (0, 1]$. From (4) there is a cheap talk equilibrium with $\alpha = \beta = 0$ so there is no lying and the recommended action is definitely better. In this equilibrium the expected value of either recommended action is v, so the decision maker follows the recommendation if the value of the outside option C is less. Therefore the decision maker takes the recommended action with probability $\Pr[v_C \le v] = v$. Without a recommendation the expected value of either action is $v/2$, so with probability $\Pr[v_C \le v/2] = v/2$ the decision maker will take one of these actions rather than the outside option. Hence a cheap talk recommendation doubles the probability that one of the two actions is accepted and thereby doubles the expert's expected payoff. As shown in the Appendix, if payoffs and/or values are higher for one action than the other communication is still persuasive as long as (5) is satisfied.[15]

In the Symmetric Baseline (SB) treatment of the experiment we set $\pi_A = \pi_B = .8$ and $a = b = 1$, so as seen in Table 1 the expert receives .8 if either A or B is chosen regardless of which is really better, while the decision maker receives 1 from choosing the better action and 0 from choosing the worse action. In this symmetric case there is a one-to-one relation between the acceptance probability and the expert's payoffs, so we test whether the probability that the expert's recommendation for A or B is accepted, P^{SB}, is higher with communication than the theoretical probability that A or B is chosen without communication.[16]

Persuasiveness Hypothesis: *Communication increases the probability that the decision maker chooses A or B rather than the outside option, $P^{SB} > 1/2$, thereby benefiting the expert.*

3.2 Asymmetric Incentives

When the expert has a financial incentive to recommend one choice over another, the incentive to lie in favor of that choice should lead decision makers to be suspicious, and hence less likely to follow a recommendation for that choice. Such discounting then reduces the gains from lying, so in equilibrium only a fraction of experts lie. We are interested in how this equilibrium interaction between discounting by the decision maker and the probability of lying by the expert changes as incentives change.

Since A is favored overall, $\pi_A a \ge \pi_B b$, an increase in π_A raises the asymmetry in favor of A and hence raises the probability α of lying in favor of A from (4), which implies a lower probability $P_A = a/(1+\alpha)$ that a recommendation for A is accepted. Similarly, a decrease in π_B that also increases the asymmetry in favor of A leads to an increase in α and a fall in P_A. As long as A remains favored as the incentives change, the probability of a lie in favor of B remains at $\beta = 0$, and the probability of accepting a B recommendation remains at $P_B = b$.

[15] Sufficient symmetry is necessary for persuasiveness. In a logit discrete choice model Chung (2012) finds that expert preferences switch from quasiconvexity to quasiconcavity for sufficient asymmetry, implying that the expert becomes worse off from any communication (Chakraborty and Harbaugh, 2010).

[16] If decision makers are risk averse, or are competitive and want to hurt the expert, the theoretical acceptance probability will be lower than 1/2, which biases our test against support of the hypothesis. However, the opposite holds if they are risk loving or altruistic.

In our experiment we test this discounting hypothesis in the Asymmetric Incentives (AI) treatment by setting $\pi_A = 1$ and $\pi_B = .5$. With this increase in π_A and decrease in π_B relative to the SB treatment, we predict more lying toward A than B in the AI treatment itself, $\alpha^{AI} > \beta^{AI}$, and more lying toward A than in the SB treatment, $\alpha^{AI} > \alpha^{SB}$. The chance P_A^{AI} that the expert follows the expert's A recommendation is predicted to fall accordingly.

Discounting Hypothesis: *A higher expert incentive for A or lower expert incentive for B increases the probability of a lie that A is better,* $\alpha^{AI} > \beta^{AI}$ *and* $\alpha^{AI} > \alpha^{SB}$, *resulting in a lower probability that a recommendation for A is accepted,* $P_A^{AI} < P_B^{AI}$ *and* $P_A^{AI} < P_A^{SB}$.

3.3 Asymmetric Values

Che, Dessein and Kartik (2013) show how an expert often has an incentive to "pander" by recommending an action that the decision maker is already leaning toward, with a resulting loss in credibility that hurts both sides. To capture a simplified version of pandering, we consider the case where the distribution is asymmetric with $a \neq b$ so one of the actions has a higher expected value to the decision maker.[17] For instance, a consumer is known to favor the design of one product over another. Since a truthful recommendation would raise the expected value more for one action, the expert has a pandering bias to falsely claim it is better.

Pushing A is more beneficial to the expert if the recommendation is believed, but the decision maker is more suspicious of such a recommendation so in equilibrium the expert does not always lie in favor of A. As seen from (4), an increase in a or decrease in b strictly increases α and has no effect on $\beta = 0$. From (3), an increase in a has a positive direct effect on P_A that is partially counteracted by the lower credibility from the rise in α,[18] and there is no effect on P_B. For a decrease in b, the effect on P_A via α is negative and the direct effect on P_B is also negative, so both acceptance rates fall.

In the Asymmetric Values (AV) treatment we set $a = 1$ and $b = .5$. The decrease in b relative to the SB treatment should more lying toward A and a fall in acceptance rates.

Pandering Hypothesis: *A higher relative decision maker value for A increases the probability of a lie that A is better,* $\alpha^{AV} > \beta^{AV}$ *and* $\alpha^{AV} > \alpha^{SB}$, *and a lower decision maker value for B decreases the probability that a recommendation for either action is accepted,* $P_A^{AV} < P_A^{SB}$ *and* $P_B^{AV} < P_B^{SB}$.

[17] Our approach differs from Che, Dessein and Kartik (2013) in that we assume the expert's preferences are state-independent and the decision maker's outside option is uncertain, but the same intuition holds.

[18] For the limiting pure cheap talk case in the sufficiently symmetric parameter range, substituting in (4) gives $P_A = b\frac{\pi_B}{\pi_A}$ so the effects cancel out.

3.4 Extension – Opaque Incentives

We now extend the model to allow the decision maker to be uncertain about the expert's incentives so that the expert might be biased or not.[19] For instance, a store might offer both its own brand and a national brand, and a customer is uncertain whether a salesperson benefits more from selling the store brand. We assume that with equal chance the expert is either biased toward A so $\pi_A > \pi_B$ or unbiased so $\pi_A = \pi_B = \pi_u$ for some $\pi_u > 0$. The decision maker only knows that the expert might be biased toward A. To focus on the incentive issue and simplify the analysis suppose $a = b = v$ for some $v \in (0, 1]$.

For the expert type who is biased toward A, recommending A is less suspicious than if their bias was known, so we might expect the biased expert to respond by lying even more. For the unbiased expert who has equal incentives for A or B, we might expect that the decision maker's suspicion of an A recommendation would make recommending A less attractive than if the expert's impartiality was known, so the unbiased expert might recommend B even when A is better. Checking, let the false claim probabilities be α_b and β_b for a biased type and α_u and β_u for an unbiased type, and consider a candidate equilibrium where the biased expert always lies in favor of A but never lies in favor of B, $\alpha_b = 1$, $\beta_b = 0$, and the unbiased expert never lies in favor of A but always lies in favor of B, $\alpha_u = 0$, $\beta_u = 1$. This implies that half the time the recommended action is good and half the time it is bad so the decision maker loses nothing from following the recommended action rather the opposite action, implying $P_A = P_B = v/2$. Given these acceptance rates the biased expert has a strict incentive to lie, and the unbiased expert is just indifferent so these strategies are a pure cheap talk equilibrium.[20]

As verified in the Appendix, with lying aversion the same pattern must hold that biased experts lie in favor of A and unbiased experts lie in favor of B, with corresponding effects on P_A and P_B. We test these predictions with the Opaque Incentives (OI) treatment, in which each expert is assigned with equal chance to have the same unbiased incentives as in the SB treatment, or the same biased incentives as in the AI treatment. We predict more lying by biased types toward A than in the AI treatment, and now expect lying by unbiased types toward B when none was predicted in the AI and SB treatments.

Regarding acceptance rates, since we predict lying in favor of both A and B in the OI treatment, and no lying in the SB treatment, lack of transparency is predicted to reduce

[19] We do not test the case where instead the expert is uncertain over the decision maker's values for A and B. Chakraborty and Harbaugh (2014) analyze this case in a random coefficients discrete choice model.

[20] Pure cheap talk is uninformative here because the types are equally likely. If one type is more likely and not too biased, the gains from lying or not can be equalized for that type with less than complete lying. Most relatedly, see Chakraborty and Harbaugh (2010, online Appendix).

acceptance rates for A and B relative to the SB treatment. And since there is lying in favor of A but not B in the AI treatment, the acceptance rate for B is expected to be lower in the AI treatment. The effect on the acceptance rate for A is ambiguous since the credibility of an A recommendation is enhanced by the presence of unbiased experts who do not lie in favor of A, but it is undermined by the stronger incentive of biased experts to lie in favor of A.

Noting that the main insight is similar to that of the political correctness theory of Morris (2001), we have the following hypothesis.

Political Correctness Hypothesis: *For symmetric values, if the expert is equally likely to be biased toward A or unbiased, then (i) biased and unbiased experts lie in opposite directions,* $\alpha_b^{OI} > \beta_b^{OI}$ *and* $\alpha_u^{OI} < \beta_u^{OI}$*, and are more likely to lie than if the expert is known to be biased or known to be unbiased,* $\alpha_b^{OI} > \alpha^{AI}$,$\alpha^{SB}$ *and* $\beta_u^{OI} > \beta^{AI}$,$\beta^{SB}$*, and (ii) the probability that a recommendation for either A or B is accepted is lower than if the expert is known to be unbiased,* $P_A^{OI} < P_A^{SB}$ *and* $P_B^{OI} < P_B^{SB}$*, and the probability that a recommendation for B is accepted is lower than if the expert is known to be biased,* $P_B^{OI} < P_B^{AI}$.

4 Experiment

4.1 Experimental Design

Following the exact model introduced in Section 3, we conduct an experiment using the parameter values for the different treatments shown in Table 1. In each treatment a "consultant" (expert) has two projects, A and B, one of which is better for the "client" (decision maker). The probability that a project is the better one is 1/2 independently each round. The client has an alternative project C with a value that is drawn from a uniform distribution independently each round. The consultant knows the realized values of projects A and B and the client knows the realized value of the alternative project C. After learning the realized values of projects A and B the consultant sends one of two messages to the client, either "I recommend Project A" or "I recommend Project B". The client learns the realized value of the alternative project C and sees the client's message. The client then chooses project A, B, or C and the decision and the outcome are displayed.

We conduct four sessions with 20 subjects each drawn from undergraduate classes at the Kelley School of Business at Indiana University. We follow a within-subject design in which the same subjects are exposed to all four treatments in the sequence SB-AI-OI-AV. This sequence minimizes the potential for confounding treatment effects with learning effects or experimenter demand effects since from each treatment to the next subjects do not have an incentive to either just follow their previous behavior or just switch to the opposite behavior. In the first

SB treatment experts have no incentive to lie, in the second AI treatment experts have an incentive to lie if B is better, in the third OI treatment with equal chance an expert is biased and has an incentive to lie if B is better, or is unbiased and has an incentive to lie if A is better, and in the final AV treatment all experts have an incentive to lie if A is better.[21]

The experiments are conducted on computer terminals at the university's Interdisciplinary Experimental Lab using z-Tree software (Fischbacher, 2007). In each session the subjects are randomly assigned to be one of 10 consultants and 10 clients. Each of the four treatments (or "series" as described to subjects) in a session lasts 10 rounds and in each treatment every client is matched anonymously with every consultant once and only once. At the end of the experiment one round from each treatment is randomly chosen as the round that the subjects are paid for to reduce earnings effects. Subjects learn which rounds were chosen and are paid privately in cash. No record of their actions mapped to their identity is maintained. The monetary payoffs are set 10 times larger in US dollar terms than in Table 1, but there is a 1 in 10 chance of each round being chosen, so the expected values for each round are the same as in Table 1. The average earnings in the 90-minute experiment are, excluding a $5 show-up payment, $22.02 for consultants and $27.78 for clients. The detailed procedures and verbal instructions are in Appendix B. Sample screen shots are in Appendix C.

4.2 Results and Analysis

In each of the identical four sessions there are four treatments, each with ten rounds. Since there are 20 subjects in a session, each round has 10 pairs of recommendations and decisions. To allow for some learning, we base the statistical analysis on behavior in the last 5 rounds of each treatment.[22] In these rounds there are 50 recommendations and 50 decisions in each session for each treatment, but these data points are not independent since players make multiple decisions and learn from interactions with other players. To allow for such interdependence, for our statistical analysis we follow the most conservative approach of analyzing behavior at the session level rather than subject level or subject-round level. In particular we treat

[21] As discussed below, the evidence for one of our secondary predictions for the OI treatment may still be confounded by learning effects. Given the complexity of the OI treatment, an alternative between-subject design would require either a large number of rounds in which reputation effects would be hard to eliminate, or some practice rounds with and without expert bias for the subjects to better understand the game. The within-subject design keeps track of behavior from such "practice rounds" in the simpler SB and AI treatments and allows it to be compared with behavior in the more complicated OI and AV treatments.

[22] The significance results are qualitatively the same if we base the tests on behavior frequencies over all 10 rounds, with one exception discussed in Footnote 26. Due to a coding error, data for the first round of the Asymmetric Values Treatment was not recorded so only 9 rounds of data are available for the treatment.

the frequency of a behavior in the last five rounds of a treatment for each session as the unit of analysis, and used paired tests to reflect the within-subject design. Since there are only four independent data points for a treatment, one for each session, statistical significance requires consistent differences in behavior frequencies across treatments within each of the four sessions. These frequencies are reported in Table 2. We do not formally model learning but for comparison in Figures 1 and 2 we indicate frequencies for the first five rounds by dashed gray lines and frequencies for the last five rounds by solid gray lines. We limit our discussion of learning to cases where there are large or unexpected differences between the first five and last five rounds.

	Var.	Ses. 1	Ses. 2	Ses. 3	Ses. 4	Overall
Symmetric Baseline Treatment						
Recommend A When B Better	α^{SB}	.14	.13	.11	.07	.11
Recommend B When A Better	β^{SB}	.14	.15	.04	.04	.10
A Recommendation Accepted	P_A^{SB}	.63	.77	.92	.96	.81
B Recommendation Accepted	P_B^{SB}	.39	.71	.88	1.00	.76
A or B Recommendation Accepted	P^{SB}	.52	.74	.90	.98	.79
Asymmetric Incentives Treatment						
Recommend A When B Better	α^{AI}	.35	.31	.52	.53	.41
Recommend B When A Better	β^{AI}	.21	.24	.04	.06	.13
A Recommendation Accepted	P_A^{AI}	.36	.32	.53	.53	.45
B Recommendation Accepted	P_B^{AI}	.68	.80	.71	1.00	.77
Opaque Incentives Treatment						
Biased: Recommend A When B Better	α_b^{OI}	.45	.45	.58	.78	.56
Biased: Recommend B When A Better	β_b^{OI}	.36	.07	.00	.31	.19
Unbiased: Recommend A When B Better	α_u^{OI}	.33	.15	.00	.00	.13
Unbiased: Recommend B When A Better	β_u^{OI}	.38	.58	.58	.50	.51
A Recommendation Accepted	P_A^{OI}	.58	.28	.68	.54	.52
B Recommendation Accepted	P_B^{OI}	.54	.64	.80	.92	.72
Asymmetric Values Treatment						
Recommend A When B Better	α^{AV}	.46	.68	.63	.70	.62
Recommend B When A Better	β^{AV}	.00	.05	.00	.09	.03
A Recommendation Accepted	P_A^{AV}	.46	.40	.68	.75	.58
B Recommendation Accepted	P_B^{AV}	.62	.30	.44	.40	.45

Table 2: Lying and Acceptance Frequencies in Rounds 6-10 of Each Treatment

For the SB treatment the experts have no incentive to lie so from the Persuasiveness Hypothesis we expect that decision makers will, on average, follow expert predictions whereas they would only do so half the time without communication. Focusing on the last 5 rounds of the treatment, Table 2 shows that experts lie about 11% of the time and decision makers follow the expert's recommendation about 79% of the time. The "Expert Lies" bar in Figure 1(a) shows the frequency that the expert recommends A given that B is better (α^{SB}) and the frequency that the expert recommends B given that A is better (β^{SB}). Figure 2(a) shows that the decision maker is less likely to follow the expert's recommendation as the outside option becomes more attractive.

Regarding the statistical significance of these results, since we treat the data generating process as being at the session level for each treatment, we are interested in whether the overall rate of acceptance for A and B is significantly above 1/2 in the four treatments. From Table 2, we see that this frequency P^{SB}, which is a weighted average of P_A^{SB} and P_B^{SB}, is above 1/2 in each session.[23] For any possible distribution the probability that this would occur if the hypothesis were false is less than $(1/2)^4 = 1/16$, which is the p-value for the one-sided Wilcoxon signed rank test as seen in Table 3. The one-sided t-test (which, in contrast with the subsequent tests, is a one sample test so it is not paired) indicates the difference is significant at the 5% confidence level. We report t-test results for completeness, though clearly the normality assumption is not valid in our context.

In the AI treatment where action A is more incentivized, the Discounting Hypothesis is that decision makers are less likely to accept a recommendation for A since experts are more likely to falsely claim A is better when in fact B is better. Consistent with the predictions for false claim rates in this hypothesis, from Table 2 and as seen in Figure 1(b) we find that experts falsely claim A to be better 41% of the time and falsely claim B to be better 13% of the time. From Table 2 we see that false claims are higher for A than B ($\alpha^{AI} > \beta^{AI}$) in each session for the treatment, which for any distribution would occur if the hypothesis were false with probability $(1/2)^4 = 1/16$ as seen in Table 3. We also see that false claims for A are always higher than in the SB treatment ($\alpha^{AI} > \alpha^{SB}$), which again implies a p-value of 1/16. Using a paired one-sided t-test both of these false claim differences are statistically significant at the 5% or better level. Going forward, we will refer to a difference as significant if the p-value for the Wilcoxon signed rank test reaches this lowest attainable level of $p = 1/16$ and

[23] Lower acceptance rates in the first two sessions may reflect the effects of one expert in each session who repeatedly lies for no material benefit, as discussed in Section 4.3. Consistent with such non-predicted lying affecting subsequent decision maker behavior, acceptance rates in the first five rounds of Sessions 1 and 2 are close to those of Sessions 3 and 4, but diverge in the last five rounds. Our use of session-level data prevents correlated error terms for subjects within a session from biasing inference.

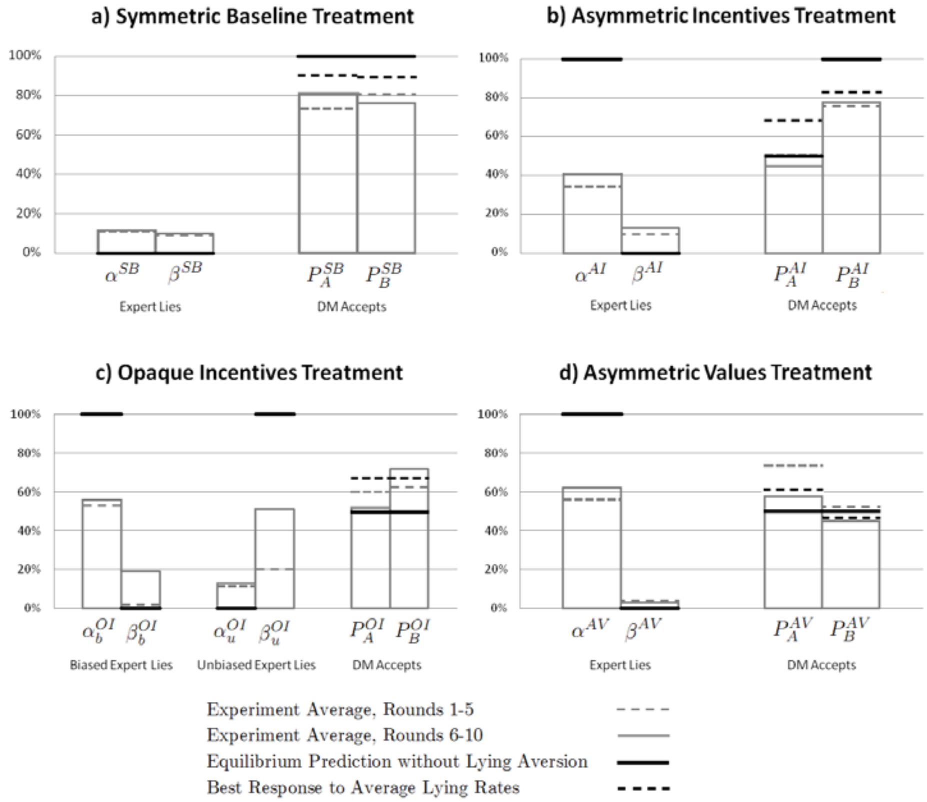


Figure 1: Lying and Acceptance Frequencies

the paired one-sided t-test also indicates significance at the 5% or better level.

Regarding acceptance rates for the AI treatment, as seen in Table 2 and in Figure 1(b) decision makers accept the recommendation 45% of the time when A is recommended and 77% of the time when B is recommended. From Table 2 acceptance rates in the treatment are lower for A than for B in each session ($P_A^{AI} < P_B^{AI}$), and also acceptance rates are lower for A in the treatment than they are in the SB treatment ($P_A^{AI} < P_A^{SB}$). As seen in Table 3, both of these differences are significant. Looking at Figure 2(b), decision makers are unlikely to accept A recommendations when the outside option v_C is favorable.

For the OI treatment, from the Political Correctness Hypothesis we expect that a biased expert type is more likely to falsely claim A is better and, of particular interest, an unbiased expert type is more likely to falsely claim B is better. Consistent with the prediction, from the summary data in Table 2 and as seen in Figure 1(c) we find that biased experts falsely claim A is better 56% of the time and falsely claim B is better 19% of the time, while unbiased experts falsely claim A is better 13% of the time and falsely claim B is better 51% of the time.

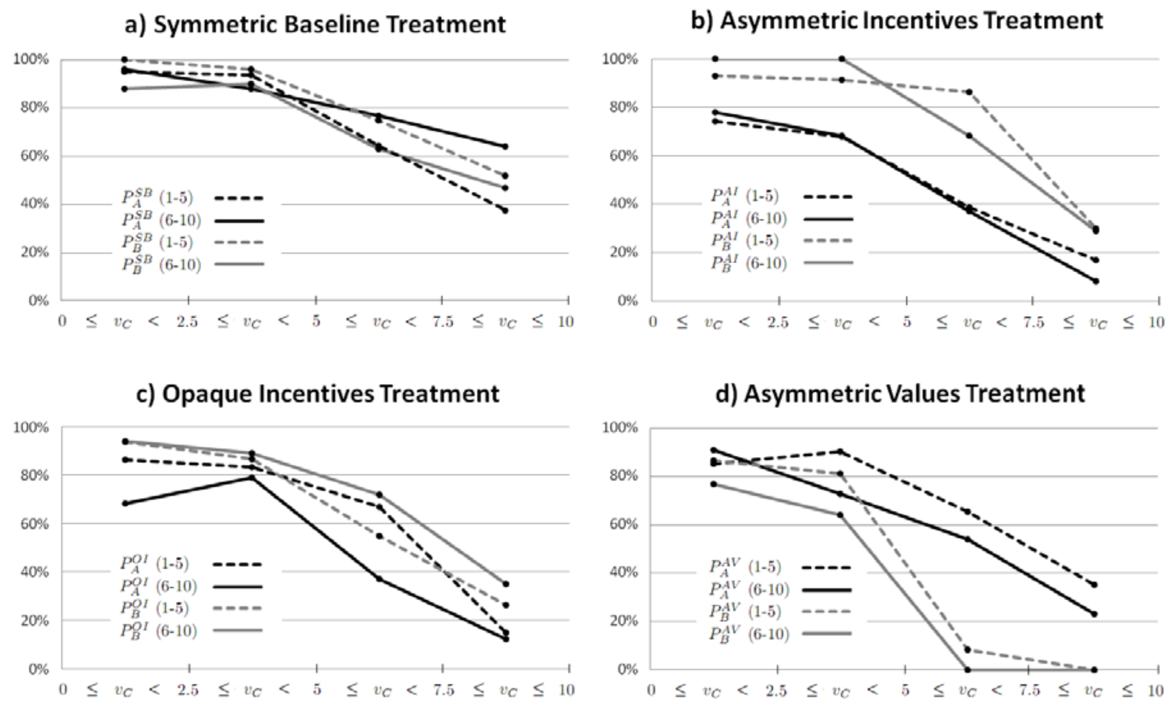


Figure 2: Acceptance Frequencies by Value of Outside Option

Notice from the dashed gray line in Figure 1(c) that unbiased experts do not immediately appear to recognize the benefits of lying toward B, but learn strongly over the course of the experiment. From Table 3, both of the false claim differences for the last five rounds are significant ($\alpha_b^{OI} > \beta_b^{OI}$ and $\alpha_u^{OI} < \beta_u^{OI}$). If we compare behavior with that in the SB treatment, we find that in each session in the OI treatment biased experts are more likely to falsely claim A is better and unbiased experts are more likely to falsely claim B is better ($\alpha_b^{OI} > \alpha^{SB}$ and $\beta_u^{OI} > \beta^{SB}$), and both differences are significant. Comparing behavior with that in the AI treatment, we find that biased experts are more likely to falsely claim A is better and unbiased types are more likely to falsely claim B is better ($\alpha_b^{OI} > \alpha^{AI}$ and $\beta_u^{OI} > \beta^{AI}$), and that the differences are significant. However, note that biased experts may be continuing behavior in the OI treatment that they learned in the AI treatment, so this higher rate of lying might reflect learning instead.[24]

[24] Arguing for a learning interpretation, the extra incentive to lie in the Opaque Incentive treatment is weak for biased experts (rather than for unbiased experts whose political correctness behavior is the focus of this treatment). Arguing against, from Figure 1(b-c) there is a bigger jump in lying between the treatments than there is between the first and second half of each treatment.

For acceptance rates in the OI treatment the results are more mixed. Theory predicts that decision makers should be more suspicious of an A recommendation than in the SB treatment ($P_A^{OI} < P_A^{SB}$) and the differences are significant. And theory predicts that decision makers should be more suspicious of a B recommendation than in either the SB or AI treatment ($P_B^{OI} < P_B^{SB}$ and $P_B^{OI} < P_B^{AI}$), but neither difference is significant. Hence it seems that decision makers are correctly suspicious of an A recommendation, and that unbiased experts correctly anticipate this and falsely claim that B is better, but it does not appear that decision makers fully anticipate such lying by unbiased experts.[25] However, the power of our test is limited by the small sample size, so even if decision makers did fully anticipate lying toward B it might not be significant in our test. From the gray dashed line in Figure 1(c) decision makers do not appear to become more suspicious of B recommendations over the course of the experiment, but instead are more likely to accept B recommendations in later rounds.[26] This pattern is also seen in Figure 2(c) where decision makers with a favorable outside option become less rather than more suspicious of B recommendations over time. We did not use a roll-reversal mechanism to speed up learning, so ten rounds might simply have been insufficient opportunity for decision makers to successfully learn to best respond.

Finally, for the AV treatment, from the Pandering Hypothesis we expect that experts are more likely to falsely claim that A is better than to falsely claim that B is better ($\alpha^{AV} > \beta^{AV}$). Consistent with the prediction, from the data in Table 2 and as seen in Figure 1(d) we find a difference of 62% vs. 3% in the summary data, which from Table 3 is statistically significant. Comparing the frequency of false claims for A in this treatment with that in the SB treatment, the difference of 62% vs. 11% is also significant ($\alpha^{AV} > \alpha^{SB}$).[27] Regarding acceptance rates in the AV treatment, consistent with theory the acceptance rates for both A and B fall relative to the SB treatment ($P_A^{AV} < P_A^{SB}$ and $P_B^{AV} < P_B^{SB}$), but as seen in Table 3 only the fall for A is significant. Moreover, as seen from Figure 1(d) and in Figure 2(d), decision makers take

[25] Such behavior may reflect level-k thinking in which subjects do not fully consider the entire chain of strategic interactions (e.g., Stahl and Wilson, 1995; Nagel, 1995; Crawford, 2003).

[26] Since behavior is initially closer to the equilibrium prediction, the difference in session means for P_B^{OI} and P_B^{SB} is significant if we include the first 5 rounds rather than restricting the analysis to the final 5 rounds.

[27] Subjects who were randomly assigned in the OI treatment to be unbiased rather than biased were more likely to lie (toward A) in the AV treatment. After recognizing by the end of the OI treatment that they benefit by lying toward B, they may have thought more carefully about the AV treatment and recognized they should lie toward A. It may also reflect an experimenter demand effect whereby subjects anticipate that between sessions their behavior is expected to change and do so accordingly. However, we did not see such a reversal between the AI and OI treatments, and the OI randomization ensures that there is no predictable pattern across treatments in the direction of lying incentives.

some time to become suspicious of A recommendations.[28]

	Hypothesis	Signed rank test	Paired t-test
Persuasiveness Hypothesis			
A and B Acceptance Rate SB Treatment	$P^{SB} > 1/2$	.063	.032
Discounting Hypothesis			
A vs. B Lying Rates AI Treatment	$\alpha^{AI} > \beta^{AI}$	.063	.037
A Lying Rates AI vs. SB Treatments	$\alpha^{AI} > \alpha^{SB}$	.063	.010
A vs. B Acceptance Rates AI Treatment	$P_A^{AI} < P_B^{AI}$	.063	.007
A Acceptance Rates AI vs. SB Treatments	$P_A^{AI} < P_A^{SB}$	.063	.001
Political Correctness Hypothesis			
A vs. B Lying Rates Biased OI Treatment	$\alpha_b^{OI} > \beta_b^{OI}$	.063	.017
A vs. B Lying Rates Unbiased OI Treatment	$\alpha_u^{OI} < \beta_u^{OI}$	.063	.022
A Lying Rates Biased OI vs. SB Treatments	$\alpha_b^{OI} > \alpha^{SB}$	.063	.008
B Lying Rates Unbiased OI vs. SB Treatments	$\beta_u^{OI} > \beta^{SB}$	.063	.004
A Lying Rates Biased OI vs. AI Treatments	$\alpha_b^{OI} > \alpha^{AI}$	.063	.018
B Lying Rates Unbiased OI vs. AI Treatments	$\beta_u^{OI} > \beta^{AI}$	.063	.008
A Acceptance Rates OI vs. SB Treatments	$P_A^{OI} < P_A^{SB}$	.063	.027
B Acceptance Rates OI vs. SB Treatments	$P_B^{OI} < P_B^{SB}$	.438	.372
B Acceptance Rates OI vs. AI Treatments	$P_B^{OI} < P_B^{AI}$	.188	.137
Pandering Hypothesis			
A vs. B Lying Rates AV Treatment	$\alpha^{AV} > \beta^{AV}$	.063	.000
A Lying Rates AV vs. SB Treatments	$\alpha^{AV} > \alpha^{SB}$	.063	.002
A Acceptance Rates AV vs. SB Treatments	$P_A^{AV} < P_A^{SB}$	.063	.005
B Acceptance Rates AV vs. SB Treatments	$P_B^{AV} < P_B^{SB}$	.125	.096

Table 3: p-values for One-Sided Hypothesis Tests, $n = 4$

The above results are for paired tests which typically generate lower $p-$values for within-subject designs since they control for subject-specific factors.[29] However, in our case the treatment effects are so strong relative to any such factors that in almost every case the unpaired tests have equal or lower $p-$values than the paired tests. As seen in Table 2, for

[28] Acceptance probabilities are higher for A than B. While B is never a good choice when $v_C > 5$, A can still be a good choice for v_C not too large if there is some chance that the expert is not lying.

[29] Since we treat each session with 20 subjects as one data point, subject-specific factors are largely averaged out, but there may also be session-specific factors such as the evolution of play due to strategic interactions.

all but one hypothesis for false claim rates (α_b^{OI} vs. α^{AI}) all four rates for one treatment are ranked above all four rates for the other treatment, which is the strongest possible ranking and implies $p = .014$ for the unpaired Wilcoxon (Mann-Whitney) rank order test. The differences in acceptance rates also all remain significant with equal or lower p-values in unpaired tests, except that there is still no significant tendency in the OI treatment for decision makers to appropriately discount recommendations in the opposite direction of biased experts.

4.3 Expert Lying Aversion

The model assumes heterogeneous lying aversion, but as shown in Proposition 2 the comparative static predictions of the model do not depend on the exact distribution of lying costs. The experiment is not designed to estimate the distribution of lying costs since it does not include treatments that systematically vary the incentive to lie. Figure 3 shows the lying frequencies for the 40 different experts, 10 in each of the four sessions. The experts are ordered according to the frequency of lying in the predicted direction, i.e., the percentage of times that the expert lies in the direction for which in equilibrium there is a material incentive to lie. This is a weighted average of $\alpha^{AI}, \alpha_b^{OI}, \beta_u^{OI}$, and α^{AV}, where the weights depend on the actual number of opportunities to lie which vary randomly based on whether A or B was in fact the better action. To reduce noise, data for all 10 rounds is shown.

Looking at the Figure, there is substantial variation in how frequently experts lie in the predicted direction. Of the 40 experts, 7 lie at least 80% of the time, 25 lie between 20% and 80% of the time, and 8 lie less than 20% of the time or not all. With pure cheap talk the equilibrium and the best response is to always lie in these cases, so the pattern appears consistent with our assumption from the literature that a substantial fraction of experts are averse to but not categorically opposed to lying. However, our model is not designed to test this assumption and it is not possible to confirm that experts have heterogenous lying costs. If, for instance, each expert has the same lying cost then in the resulting mixed strategy equilibrium there would be heterogeneity in realized lying behavior, which with anonymity does not preclude some players from consistently lying more than others.

Lying aversion in our experiment could reflect a true preference against lying, but may also be a reduced form for other considerations such as an altruistic concern for the payoff to the decision maker (Hurkens and Kartik, 2009). Note that one expert in each of Sessions 1 and 2 lies consistently in a non-predicted direction which hurts the decision maker without benefiting the expert materially.[30] Lying frequencies in a direction for which there is no

[30] Such lying might reflect competitive or “nasty” preferences (Abbink and Sadrieh, 2009) or might just reflect expert confusion about the best strategy.

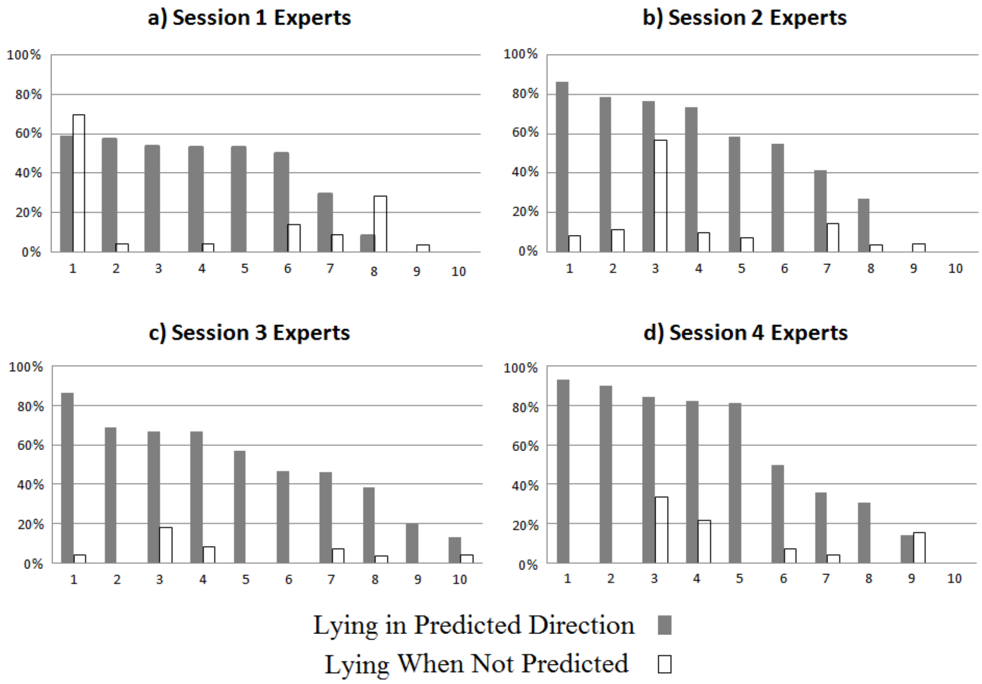


Figure 3: Overall Frequency of Lying by Each Expert

material incentive to lie are also shown in Figure 3. These frequencies are a weighted average of $\alpha^{SB}, \beta^{SB}, \beta^{AI}, \beta_b^{OI}, \alpha_u^{OI}$ and β^{AV}. Such lying may confuse decision makers and lead to further effects on play in the session, but our statistical analysis based on session-level data is unaffected by such interdependence between rounds.

4.4 Decision Maker Payoffs and Best Responses

The decision maker can take a sure payoff of v_C from the outside option or can accept the expert's recommendation and, depending on whether the expert lied or not, receiver either zero or 10 (or 5 for the B choice in the AV treatment). A risk neutral decision maker's best response acceptance rates from (1) are shown in the dark dashed lines in Figure 1 based on the empirical lying rates in each treatment. Expected (monetary) payoffs from deviating are small in the neighborhood of the best response and increase in the square of the size of the deviation. Actual acceptance rates are on average 8% below these best response rates, which may reflect risk aversion. Overall, if (risk neutral) decision makers best responded they would have made

\$30.06 on average (net of the show-up payment) but instead made \$27.42 on average.[31] For comparison, if they had always ignored the experts and chosen C they would have made \$20 on average,[32] and if the experts never lied and the decision makers best responded to such honesty they would have made \$37.50 on average.[33]

Cases where acceptance rates diverge the most from the general pattern may reflect failure by decision makers to properly account for expert behavior. First, in the AI treatment the lying rates of 41% for A and 13% for B imply from (1) that the expected value of A when recommended is $(1-.13)/(1-.13+.41)=.68$, implying a best response acceptance rate of 68% for A, but the actual rate is only 45%. Such strong discounting of A recommendations suggests that decision makers underestimate how much communication is possible from a biased expert. Second, in the OI treatment the combined lying rates for biased and unbiased experts of 33% for A and 35% for B imply from (1) that the expected value of B when recommended is $(1-.35)/(1-.35+.33)=.66$ so the best response acceptance rate is 66% for B, but the actual acceptance rate is 72%. Insufficient discounting of B recommendations, especially when risk aversion appears to induce excess discounting in every other situation, is consistent with the inference from Table 3 and from the learning patterns in Figures 1(c) and 2(c) that some decision makers may not recognize how opaque incentives give even unbiased experts an incentive to lie. However, as discussed above, more data would be required to draw any firm conclusions

5 Conclusion

This paper brings together insights from the literature on cheap talk recommendations into one simple, easily testable model. These insights were originally developed under varying assumptions, but they are sufficiently general that their main implications continue to hold. The model shows how recommendations through cheap talk can be persuasive even when the expert always wants the decision maker to take an action rather than no action, how asymmetries in incentives and values can distort communication but need not preclude it,

[31] These are based on the last five rounds of each treatment where the decision makers use the lying rates for their own session to calculate P_A and P_B. Actual payoffs, which allowed for an equal chance of any round (not just the last five) being chosen in each session as the payoff round, were \$27.92.

[32] In the pure cheap talk equilibrium with honest recommendations in the SB treatment and lying in the other treatments, decision makers would have made 10 + 7.50 + 7.50 + 6.75=\$31.75 and experts would have made (1/2)8 + (1/2)7.50 + (1/2)7.75 + (1/2)8.00=\$15.63.

[33] In this last case experts would have received $8+(10+5)/2+(10+5+8+8)/4+(8+8/2)/2=\$29.25$ on average rather than \$22.02, and would have suffered no utility loss from lying. Hence, as is common in cheap talk games, lack of trust hurts both the sender and receiver.

and how lack of transparency about the expert's incentives can lead even an expert with unbiased incentives to offer a biased recommendation. When experts are lying averse there is a unique equilibrium with testable comparative static predictions about expert and decision maker behavior that do not depend on the exact distribution of lying costs.

In the first experimental tests of these predictions from the literature, we find that for every hypothesis regarding expert behavior we can reject the null hypothesis of no change in the hypothesized direction. The false claim rates by experts change as predicted overall and in every session. Of particular interest is that when incentives are opaque we find that biased experts are more likely to lie, and that unbiased experts appear to recognize that they are more persuasive if they sometimes lie in order to avoid the recommendation favored by a biased expert. For the hypotheses regarding decision maker behavior, the acceptance rates also change as predicted and the changes are also statistically significant except for the opaque incentives treatment. Within the limited sample in the experiment, decision makers do not consistently discount the recommendation that is the opposite of that favored by the biased expert. Whether this represents a general failure of decision makers to understand how lack of transparency warps the incentives of even unbiased experts, or whether it is just a small sample artifact of our experiment, is an open question.

Since biased incentives make communication less reliable, and lack of transparency further undermines communication, these results provide theoretical and empirical support for policy measures that try to eliminate biases or at least make them more transparent. However, there are two important caveats. First, as shown by Inderst and Ottaviani (2012), disclosure requirements can lead to endogenous changes in incentives that affect the equilibrium of the subsequent communication game, and there may also be other endogenous changes in the payoff or information structure in response to disclosure requirements or other measures. Second, as seen from the theoretical results by Che, Dessein and Kartik (2013) that our experiment supports, an expert will often inefficiently pander to the decision maker even when the problems of biased incentives and lack of transparency about incentives are both solved.

6 Appendix A – Proofs

Proof of Persuasiveness Hypothesis: Since (5) holds by assumption the equilibrium is given by (3) and (7). Since half the time A is better, in which case there is no lying, and half the time B is better, in which case there is lying toward A with probability α, the probability that A is taken is $\frac{1}{2}P_A+\frac{1}{2}\alpha P_A = a/2$ and the probability that B is taken is $\frac{1}{2}(1-\alpha)P_B = \frac{1}{2}(1-\alpha)b$. Hence the overall probability that A or B is taken is $P = (a+b-\alpha b)/2$. From (7), α is strictly smaller with lying costs than with pure cheap talk, so substituting from (4),

$$P > \frac{a+b-(\pi_A a/\pi_B b-1)b}{2}. \tag{8}$$

Note that $(a+b-(\pi_A a/\pi_B b-1)b)/2-a/2 = \frac{1}{2\pi_B}(2\pi_B b-\pi_A a) \geq 0$ where the inequality follows from the recommendation condition in (5), so $P > a/2$. Also note that $(a+b-(\pi_A a/\pi_B b-1)b)/2-b/2 = \frac{1}{2\pi_B}(\pi_B a-\pi_A a+\pi_B b) \geq 0$ where the inequality follows from the acceptance condition in (5), so $P > b/2$.

Without communication, the decision maker will either take the outside option C or take the action A or B offering the highest expected payoff where $E[v_A] = a/2$ and $E[v_B] = b/2$. So without communication $P = \max\{a/2, b/2\}$, which is lower than with communication.

Regarding expert payoffs, with communication from (7) if B is better then expert types with a lying cost lower than d_A lie and receive a higher payoff than $\pi_B b$, while expert types with a lying cost above d_A tell the truth and receive $\pi_B b$. If A is better the expert receives $\pi_A a$. So, since $\pi_A a > \pi_B b$ by assumption, the lowest payoff for any expert type is $\pi_B b$. Without communication the expert receives $\max\{\pi_A a/2, \pi_B b/2\} = \pi_A a/2$, so communication benefits every expert type weakly and some types strictly if $\pi_B b \geq \pi_A a/2$, as required by (5). ■

Proof of Discounting Hypothesis: Totally differentiating (7), $d\alpha/d\pi_A = a/(\pi_B b + d_A + \alpha d'_A) > 0$ and $d\alpha/d\pi_B = -\alpha b/(\pi_B b + d_A + \alpha d'_A) < 0$ where the derivative d'_A exists and is strictly positive by the assumption that G has full support with no mass points. So an increase in π_A or decrease in π_B implies there is strictly more lying and hence from (3) that P_A is strictly lower. ■

Proof of Pandering Hypothesis: Totally differentiating (7), $d\alpha/da = \pi_A/(\pi_B b + d_A + \alpha d'_A) > 0$ and $d\alpha/db = -\alpha\pi_B/(\pi_B b + d_A + \alpha d'_A) < 0$ so an increase in a or decrease in b implies there is strictly more lying toward A. For an increase in a this implies from (3) that, in addition to the direct positive effect on P_A, there is an indirect negative effect from the rise in α. The direct effect must strictly dominate since from (6) α only increases if P_A increases. For a decrease in b there is the strictly negative direct effect on P_B, and a strictly negative indirect effect on P_A from the rise in α. ■

Proof of Political Correctness Hypothesis: The expected values of each action conditional on a recommendation for that action are, using the assumption that $a = b = v > 0$,

$$\begin{aligned} E[v_A|m_A] &= \Pr[v_A > v_B|m_A]v = \frac{\frac{1}{2}(1-\beta_b)+\frac{1}{2}(1-\beta_u)}{\frac{1}{2}(1-\beta_b)+\frac{1}{2}(1-\beta_u)+\frac{1}{2}\alpha_b+\frac{1}{2}\alpha_u}v \\ E[v_B|m_B] &= \Pr[v_B > v_A|m_B]v = \frac{\frac{1}{2}(1-\alpha_b)+\frac{1}{2}(1-\alpha_u)}{\frac{1}{2}(1-\alpha_b)+\frac{1}{2}(1-\alpha_u)+\frac{1}{2}\beta_b+\frac{1}{2}\beta_u}v \qquad (9) \end{aligned}$$

and the expected values of the unrecommended actions are $E[v_B|m_A] = (1-\Pr[v_A > v_B|m_A])v$ and $E[v_A|m_B] = (1-\Pr[v_B > v_A|m_B])v$.

For each of the two actions, each of the two expert types can sometimes lie toward the action or never lie, so there are 16 cases. By elimination we want to show that only $\alpha_b > 0$, $\beta_b = 0$, $\alpha_u = 0$, $\beta_u > 0$ is possible in equilibrium.

If $\alpha_b > 0$ it must be that $\pi_A P_A > \pi_B P_B$, and if $\beta_b > 0$ it must be that $\pi_A P_A < \pi_B P_B$, so any case with $\alpha_b, \beta_b > 0$ is not possible. Similarly any case with $\alpha_u, \beta_u > 0$ is not possible.

If $\alpha_b = \alpha_u = 0$ then it must be that $\pi_A P_A \leq \pi_B P_B$ and $\pi_u P_A \leq \pi_u P_B$. From (9), $\Pr[v_A > v_B|m_A] = 1$ and $\Pr[v_B > v_A|m_B] = 1/(1+\beta_b/2+\beta_u/2) > 1/2$, so $E[v_A|m_A] > E[v_B|m_A]$ and $E[v_B|m_B] > E[v_A|m_B]$ and the acceptance probabilities are $P_A = v$ and $P_B = v/(1+\beta_b/2+\beta_u/2)$. This implies $P_A \geq P_B$, a contradiction with $\pi_A P_A \leq \pi_B P_B$ since $\pi_A > \pi_B$, so any case with $\alpha_b = \alpha_u = 0$ is not possible.

If $\beta_b = \beta_u = 0$ then it must be that $\pi_A P_A \geq \pi_B P_B$ and $\pi_u P_A \geq \pi_u P_B$. From (9), $\Pr[v_A > v_B|m_A] = 1/(1+\alpha_b/2+a_u/2)$ and $\Pr[v_B > v_A|m_B] = v$, so $E[v_A|m_A] > E[v_B|m_A]$ and $E[v_B|m_B] > E[v_A|m_B]$ and the acceptance probabilities are $P_A = v/(1+\alpha_b/2+a_u/2)$ and $P_B = v$. This implies $P_A < P_B$, a contradiction with $\pi_u P_A \geq \pi_u P_B$ since $\pi_u > 0$, so any case with $\beta_b = \beta_u = 0$ is not possible.

Therefore we are left with two cases, $\alpha_b = 0, \beta_b > 0, \alpha_u > 0, \beta_u = 0$ and $\alpha_b > 0, \beta_b = 0, \alpha_u = 0, \beta_u > 0$, both of which imply $E[v_A|m_A] > E[v_B|m_A]$ and $E[v_B|m_B] > E[v_A|m_B]$ so again $P_A = E[v_A|m_A]$ and $P_B = E[v_B|m_B]$. The former case implies that the financial benefit is higher for biased types to recommend B and for unbiased types to recommend A, $\pi_A P_A < \pi_B P_B$ and $\pi_u P_A > \pi_u P_B$, which cannot both hold for $\pi_A > \pi_B$, so only the latter case is possible.

The equilibrium indifference conditions for this case $\alpha_b > 0, \beta_b = 0, \alpha_u = 0, \beta_u > 0$ are

$$\pi_A P_A - d_A = \pi_B P_B \qquad \text{for } v_B > v_A \qquad (10)$$

$$\pi_u P_A = \pi_u P_B - d_B \qquad \text{for } v_A > v_B \qquad (11)$$

for the biased and unbiased types respectively where $d_A = G^{-1}(\alpha_b)$ and $d_B = G^{-1}(\beta_u)$.

Regarding existence of equilibria, for any β_u the LHS of (10) is larger than the RHS starting at $\alpha_b = 0$. From (9) note $\frac{d}{d\alpha_b}P_A < \frac{d}{d\alpha_b}P_B < 0$ so as α_b increases the LHS of (10) falls faster until by the assumption on the support of G it eventually is smaller than the RHS. Therefore for every $\beta_u \in [0,1]$ there is a unique $\alpha_b \in (0,1)$ that solves (10). Similarly, note also that $\frac{d}{d\beta_u}P_B < \frac{d}{d\beta_u}P_A < 0$ so an increase in β_u pushes the RHS of (11) down faster than the LHS for any given α_b, so starting at $\alpha_b = 0$ in which case $\beta_u = 0$ solves (11), for every $\alpha_b \in [0,1]$ there is a unique $\beta_u \in [0,1)$ that solves (11). Considering the implicit functions $\alpha_b(\beta_u)$ and $\beta_u(\alpha_b)$ so defined, they must cross at least once in the interior so an equilibrium with lying by both types exists. By the same properties $\frac{d}{d\alpha_b}P_A < \frac{d}{d\alpha_b}P_B < 0$ and $\frac{d}{d\beta_u}P_B < \frac{d}{d\beta_u}P_A < 0$, these functions are both increasing so there may be more than one such fixed point.

Despite the possibility of non-uniqueness we can still use the equilibrium requirement $\alpha_b > 0, \beta_b = 0, \alpha_u = 0, \beta_u > 0$ to compare equilibrium behavior with the case of transparent incentives. For an unbiased expert there is no lying when incentives are transparent, so there is more lying with opaque incentives. For a biased expert there is an incentive to lie in favor of A when incentives are transparent, so the question is whether opaque incentives increase this incentive. From (10), the incentive to lie increases if P_A increases and/or P_B decreases. P_B must decrease with opaque incentives since there is no lying in favor of B when the expert is known to favor A. Hence a decrease in P_A with opaque incentives is necessary for lying in favor of A to not increase. But from (9), a decrease in P_A requires $\alpha_b - 2\alpha + \alpha\beta_u > 0$ where α is the lying rate with transparent incentives. This in turn requires $\alpha_b > \alpha$, so in either case of an increase or decrease in P_A lying by biased experts increases.

Regarding acceptance probabilities, compared to when the expert is known to be unbiased, the increases in lying in both directions imply P_A and P_B both decrease. Compared to when the expert is known to be biased, P_B falls due to lying toward B by unbiased types, but the impact on P_A is ambiguous since biased types lie more toward A but with some chance the expert is unbiased and never lies toward A. ■

7 Appendix B – Instructions and Procedures

Subjects are led into the computer room and told to sit in front of any of 20 terminals. Instructions are provided on the terminals for the consultants and clients as the experiment progresses, and summaries are read aloud to ensure that information is common knowledge. A screenshot of the main introductory screen is in Appendix C and a summary is read aloud after subjects are seated.

After all subjects have clicked through the introductory instructions on their own terminals they see a screen which assigns them their roles for the rest of the experiment and provides instructions for the first Symmetric Baseline treatment (or "first series" as described to subjects). Screenshots of these instructions for both the consultant and client are included in Appendix C. The following summary is read aloud.

> **First Series [Symmetric Baseline Treatment]:** The consultant has two projects - Project A and Project B. One is a good project worth $10 to the client and the other is a bad project worth $0 to the client. Each round the computer randomly assigns one project to be good and tells the consultant. The client does not know which project is good and which is bad.
>
> The client has his/her own project - Project C. Each round Project C is randomly assigned by the computer to be worth any value between $0.00 and $10.00. Any such value is equally likely. The computer tells the client how much Project C is worth, but the consultant does not know.
>
> The consultant will give a recommendation to the client via the computer. The recommendation will be "I recommend Project A" or "I recommend Project B".
>
> After getting the recommendation, the client will make a decision. The client earns the value of the project that is chosen. If the client chooses Project A or Project B the consultant will earn $8 in that round. However, if the client chooses his/her own Project C instead, the consultant will earn $0. One round from this series will be randomly chosen at the end as the round you are actually paid for.

After subjects click through the instructions on their screen, in each round consultants see the values of A and B and make their recommendations, and clients see the recommendations and the values of V_C and make their choices. Screenshots for these recommendations and choices are also included in Appendix C. The realized values and choices are then revealed to each consultant-client pair at the end of each round. Subjects only see the realized values and choices for their own pairing in that round. After the 10 rounds are completed subjects see a summary of their actions and payoffs in each round. They then see a screen with instructions for the next Asymmetric Incentives treatment ("second series") and the following summary is read aloud.

> **Second Series [Asymmetric Incentives Treatment]:** Everything is the same as the first series, except if the client chooses Project A the consultant will earn $10 and if the client chooses Project B the consultant will earn $5.

After the Asymmetric Incentives treatment is completed in the same manner as the Symmetric Baseline treatment, the Opaque Incentives treatment ("third series") begins. This is the most complicated treatment but its description is facilitated by its being a mix of the first two treatments. Each consultant is assigned the incentives of the first or the second treatment with equal chance. The assignment is block random and is for the duration of the treatment. The consultants know their incentives, but the clients do not know whether the consultant in any given round is biased or not. The following summary of the treatment is read aloud while subjects see a screen with the detailed description.

> **Third Series [Opaque Incentives Treatment]:** Everything is the same as before, except the Computer will randomly assign the consultant's payoff scheme. Half of the consultants will earn \$8 if either Project A or B is chosen, and half of consultants will earn \$10 if project A is chosen but only \$5 if project B is chosen. The consultant knows his/her payoff scheme but the client does not know which payoff scheme the Computer assigned the consultant.

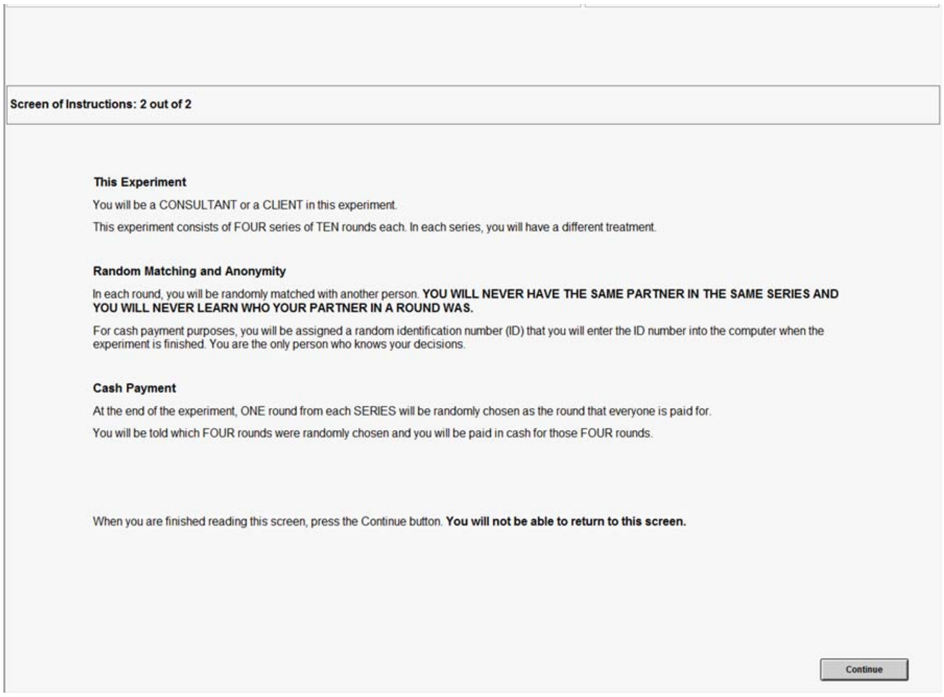
After the Opaque Incentives treatment is completed, the Asymmetric Values treatment ("fourth series") begins. The following summary of the treatment is read aloud while subjects see a screen with the detailed description.

> **Fourth Series [Asymmetric Values Treatment]:** Everything is the same as the first series, including the consultant's payoff scheme of earning \$8 if either Project A or B is chosen, except the value of Project B to the client if it is good is only \$5 instead of \$10. If project A is good its value to the client is still \$10. A bad project is still worth \$0 to the client.

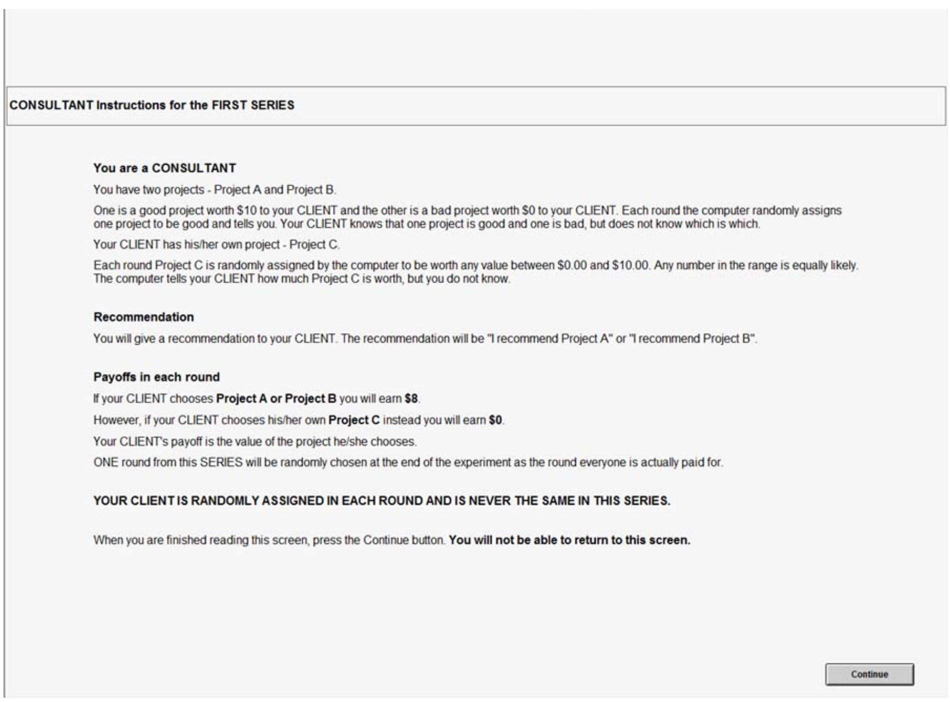
After the Asymmetric Values treatment is completed, subjects see a screen summarizing their actions and payoffs across all 40 rounds. On this screen they are also told what rounds have been randomly chosen as the rounds they will be paid for. They enter into the computer a number 1-20 that has been randomly placed next to the terminal and they are called out in sequence by this number to be paid by the experimenter. They do not see the payments for any other subjects, they sign for their payment next to their number without a printed name, and the record of their signature is preserved separately by the university's bookkeeping department. Other than the signature, no personally identifying records are maintained.

8 Appendix C – Sample Screenshots

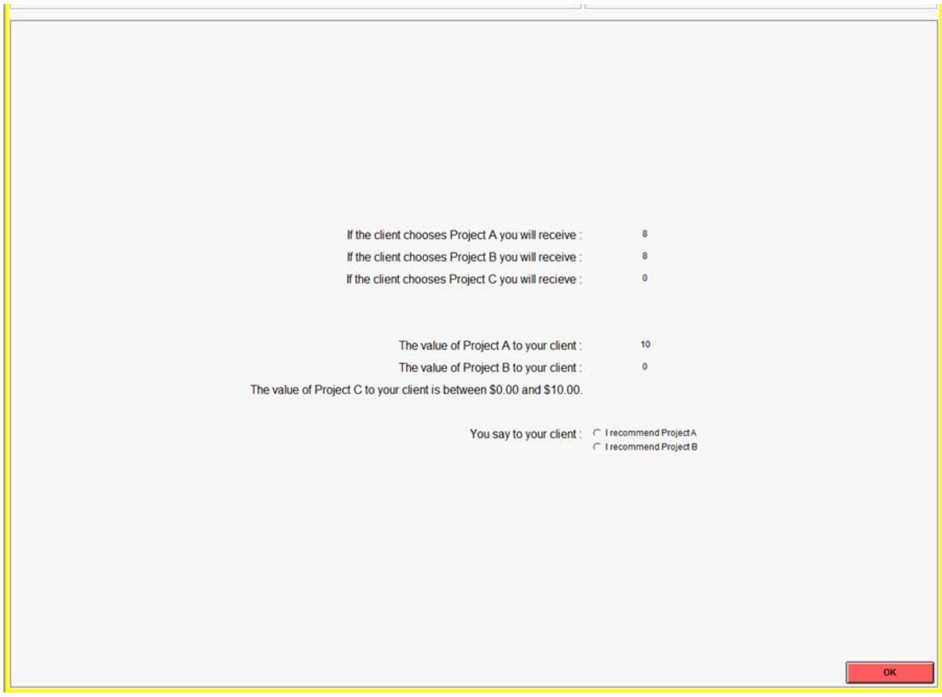
Main introductory screen:



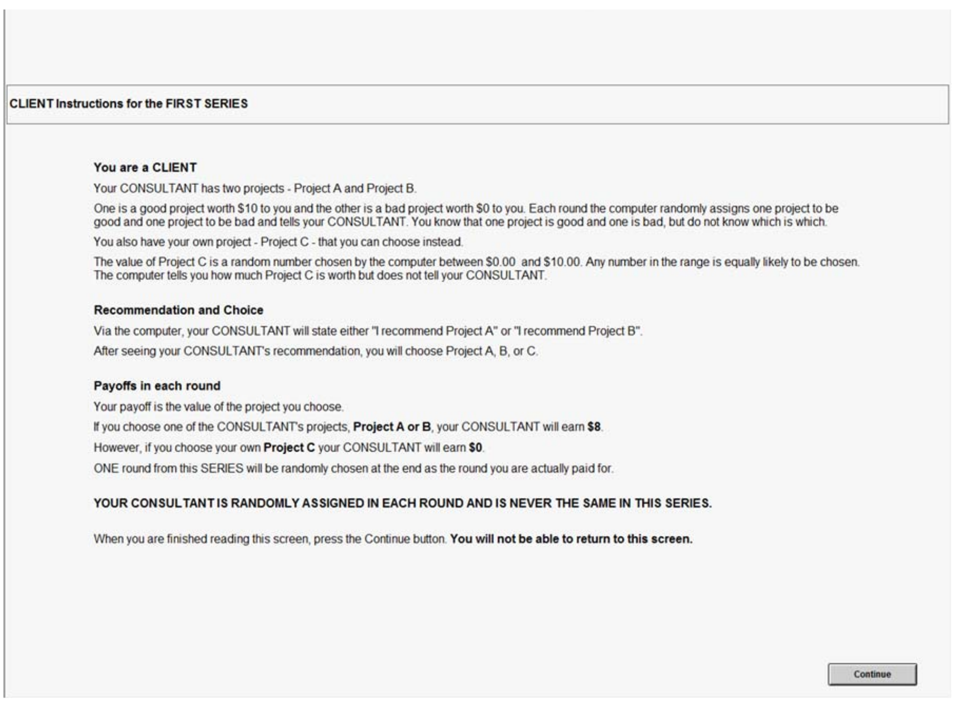
Consultant instructions for Symmetric Baseline treatment:



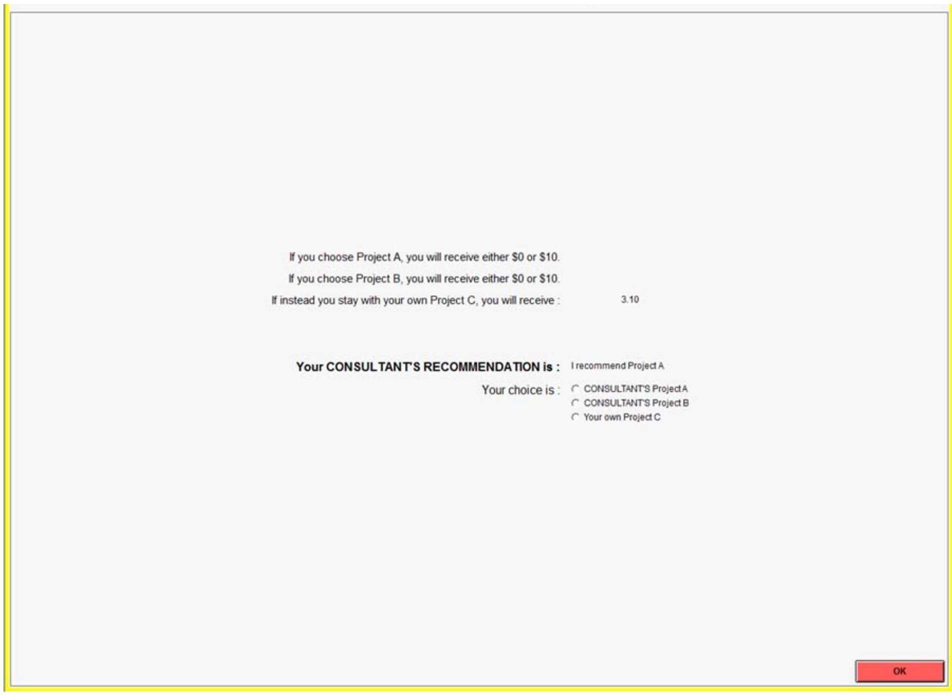
Consultant decision screen for Symmetric Baseline treatment:



Client instructions for Symmetric Baseline treatment:



Client decision screen for Symmetric Baseline treatment:



Client payoff screen for Symmetric Baseline treatment:

References

[1] Abbink, Klaus and Abdolkarim Sadrieh. 2009. The pleasure of being nasty. *Economics Letters* 105(3): 306–308.

[2] Abeler, Johannes, Daniele Nosenzo, and Collin Raymond. 2016. Preferences for truth-telling. Working paper.

[3] Armstrong, Mark and Jidong Zhou. 2011. Paying for prominence. *Economic Journal* 121(556): 368–395.

[4] Battaglini, Marco. 2002. Multiple referrals and multidimensional cheap talk. *Econometrica* 70(4): 1379–1401.

[5] Blume, Andreas, Douglas V. DeJong, Yong-Gwan Kim, and Geoffrey B. Sprinkle. 1998. Experimental evidence on the evolution of meaning of messages in sender-receiver games. *American Economic Review* 88(5): 1323–1340.

[6] Blume, Andreas, Douglas V. DeJong, Yong-Gwan Kim, and Geoffrey B. Sprinkle. 2001. Evolution of communication with partial common interest. *Games and Economic Behavior* 37: 79–120.

[7] Bolton, Patrick, Xavier Freixas, and Joel Shapiro. 2007. Conflicts of interest, information provision, and competition in the financial services industry. *Journal of Financial Economics* 85: 297–330.

[8] Cai, Hongbin, and Joseph Tao-Yi Wang. 2006. Overcommunication in strategic information transmission games. *Games and Economic Behavior* 56(1): 7–36.

[9] Cain, Daylian M., George Loewenstein, and Don A. Moore. 2005. The dirt on coming clean: perverse effects of disclosing conflicts of interest. *Journal of Legal Studies* 34(1): 1–25.

[10] Chakraborty, Archishman and Rick Harbaugh. 2007. Comparative cheap talk. *Journal of Economic Theory* 132(1): 70–94.

[11] Chakraborty, Archishman and Rick Harbaugh. 2010. Persuasion by cheap talk. *American Economic Review* 100(5): 2361–2382.

[12] Chakraborty, Archishman and Rick Harbaugh. 2014. Persuasive puffery. *Marketing Science* 33(3): 382–400.

[13] Che, Yeon-Koo, Wooter Dessein, and Navin Kartik. 2013. Pandering to persuade. *American Economic Review* 103(1): 47–79.

[14] Chung, Wonsuk. 2012. Product recommendations by a biased seller. Working paper.

[15] Crawford, Vincent. 2003. Lying for strategic advantage: Rational and boundedly rational misrepresentations of intentions. *American Economic Review* 93(1): 133–149.

[16] Crawford, Vincent P. and Joel Sobel. 1982. Strategic information transmission. *Econometrica* 50(6): 1431–1450.

[17] Darby, Michael R. and Edi Karni. 1973. Free competition and the optimal amount of fraud. *Journal of Law and Economics* 16(1): 67–88.

[18] De Jaegher, Kris and Marc Jegers. 2001. The physician-patient relationship as a game of strategic information transmission. *Health Economics* 10: 651–658.

[19] Dessein, Wouter. 2002. Authority and communication in organizations. *Review of Economic Studies*, 69(4): 811–838.

[20] Dickhaut, John, Kevin McCabe, and Arijit Mukherji. 1995. An experimental study of strategic information transmission. *Economic Theory* 6(3): 389–403.

[21] Dulleck, Uwe, Rudolf Kerschbamer, and Matthias Sutter. 2011. The economics of credence goods: an experiment on the role of liability, verifiability, reputation, and competition. *American Economic Review 101* (2): 526–555.

[22] Fischbacher, Urs. 2007. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10(2): 171–178.

[23] Gentzkow, Matthew, and Jesse M. Shapiro. 2006. Media bias and reputation. *Journal of Political Economy*, 114(2): 280–316.

[24] Gibson, Rajna, Carmen Tanner, and Alexander F. Wagner. 2013. Preferences for truthfulness: heterogeneity among and within individuals. *American Economic Review* 103(1): 532–548.

[25] Gneezy, Uri. 2005. Deception: the role of consequences. *American Economic Review* 95(1): 384–394.

[26] Hurkens, Sjaak and Navin Kartik. 2009. Would I lie to you? On social preferences and lying aversion. *Experimental Economics* 12(2): 180–192.

[27] Inderst, Roman and Marco Ottaviani. 2012. Competition through commissions and kickbacks. *American Economic Review* 102(2): 780–809.

[28] Kartik, Navin. 2009. Strategic communication with lying costs. *Review of Economic Studies* 76(4): 1359–1395.

[29] Kartik, Navin, Marco Ottaviani, and Francesco Squintani. 2007. Credulity, lies, and costly talk. *Journal of Economic Theory* 134(1): 93–116.

[30] Lafky, Jonathan and Alistair J. Wilson. 2016. Quality vs. quantity in information transmission: theory and evidence. Working paper.

[31] Lai, Ernest K., Wooyoung Lim, and Joseph Tao-Yi Wang. 2015. An experimental analysis of multidimensional cheap talk. *Games and Economic Behavior* 91: 114–144.

[32] Li, Ming and Kristóf Madarász. 2008. When mandatory disclosure hurts: expert advice and conflicting interests. *Journal of Economic Theory* 139(1): 47–74.

[33] Milgrom, Paul R. 1981. Good news and bad news: representation theorems and applications. *Bell Journal of Economics* 12(2): 380–391.

[34] Morgan, John and Phillip C. Stocken. 2003. An analysis of stock recommendations. *RAND Journal of Economics* 34(1): 183–203.

[35] Morris, Stephen. 2001. Political correctness. *Journal of Political Economy* 109(2): 231–265.

[36] Nagel, Rosemarie. 1995. Unraveling in guessing games: an experimental study. *American Economic Review* 85(5): 1313–1326.

[37] Pitchik, Carolyn and Andrew Schotter. 1987. Honesty in a model of strategic information transmission. *American Economic Review* 77(5): 1032–1036.99

[38] Sánchez-Pagés, Santiago and Marc Vorsatz. 2007. An experimental study of truth-telling in a sender–receiver game. *Games and Economic Behavior* 61(1): 86–112.

[39] Sobel, Joel. 1985. A theory of credibility. *Review of Economic Studies* 52(4): 557–573.

[40] Sobel, Joel. 2016. Lying and deception in games. Working paper.

[41] Stahl, Dale O. and Paul W. Wilson. 1995. On players' models of other players: theory and experimental evidence. *Games and Economic Behavior* 10(1): 218–254.

[42] Sutter, Matthias. 2009. Deception through telling the truth?! Experimental evidence from individuals and teams. *Economic Journal* 119(534): 47–60.

[43] Thaler, Richard H. and Cass R. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness.* Yale University Press.

[44] Vespa, Emanuel and Alistair Wilson. 2016. Communication with multiple senders: an experiment. *Quantitative Economics* 7(1):1–36.

[45] Wang, Joseph Tao-yi, Michael Spezio, and Colin F. Camerer. 2010. Pinocchio's pupil: using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games. *American Economic Review* 100(3): 984–1007.