When Sunlight Fails to Disinfect:

Understanding the Perverse Effects of Disclosing Conflicts of Interest

DAYLIAN M. CAIN

GEORGE LOEWENSTEIN

DON A. MOORE

Daylian M. Cain (daylian.cain@yale.edu) is assistant professor of organizational behavior, Yale School of Management, New Haven, CT 06520. George Loewenstein (gl20@andrew.cmu.edu) is Herbert A. Simon Professor of Economics and Psychology, Carnegie Mellon University, Pittsburgh, PA 15213. Don A. Moore (dmoore@cmu.edu) is Professor of organizational behavior, Carnegie Mellon University, Pittsburgh, PA 15213. Correspondence: Daylian Cain.
The paper is based on the first author's dissertation. The authors acknowledge the helpful input of John Deighton, Stephen Nowlis, and the reviewers at JCR. In addition, the authors thank Max Bazerman, Robyn Dawes, Baruch Fischhoff, Keith Murnighan, and Madan Pillutla for helpful comments. Finally, the authors thank Andrew Meyer, Sam Swift, Selina Tirtajana, and Rebecca Lane for help with data collection and preparation.

ABSTRACT

Disclosure is often proposed as a remedy for conflicts of interest, but it can backfire, hurting those whom it is intended to protect. Building on our prior research, we introduce a conceptual model of disclosure's effects on advisors and advice recipients that helps to explain when and why it backfires. Studies 1 and 2 examine psychological mechanisms (strategic exaggeration, moral licensing) by which disclosure can lead advisors to give more-biased advice. Study 3 shows that disclosure backfires when advice recipients who receive disclosure fail to sufficiently discount and thus fail to mitigate the adverse effects of disclosure on advisor bias. Study 4 identifies one remedy for inadequate discounting of biased advice: explicitly and simultaneously contrasting biased advice to unbiased advice.

INTRODUCTION

Imagine that your neighbor tells you about an "amazing" household cleaner she has discovered. In the same conversation, however, she discloses that she is part of a word-of-mouth affiliate program sponsored by the manufacturer, and that she receives coupons when her friends purchase the product (Berner 2006). She then quickly reassures you that the cleaner really is one of her favorite household products. How does your neighbor's disclosure that she has a conflict of interest affect your response to her advice? And how, if at all, does her disclosure change what she says about the product? This paper addresses these questions with a conceptual model and four lab experiments. The paper concludes with a policy-oriented discussion of challenges to regulation.

Literature and Background

Conflicts of interest occur when individuals' personal interests diverge from their professional or moral responsibilities to others. Although some reformers seek to manage or eliminate conflicts of interest in specific domains, the most common policy response to conflicts of interest is to disclose them. Rules requiring mandatory disclosure of conflicts of interest are ubiquitous (Harris and Souder 2004). For example, nearly all of Title IV of the Sarbanes-Oxley Act of 2001 is aimed at disclosure (e.g., of information on special-purpose accounting entities) by public corporations, and the health reform bill of 2010 includes "sunshine" provisions that mandate the disclosure of payments to physicians by pharmaceutical companies and medical device manufacturers.

Supporters of disclosure argue that transparency improves market efficiency, increases welfare, and protects the public by reducing information gaps between conflicted advisors and recipients of their advice (Healy and Palepu 2000; Gunderson 1997; Dye 2001; Verrecchia 2001). In the political realm, former U.S. senator Philip Hart has said that disclosure of political contributions reveals "the possibility of . . . conflict, leaving it to the voter to decide whether the conflict has influenced the official acts of the congressman or senator" (Hart 1975).

Existing models of conflicted advice-giving in economics (Holmstrom 1977; Crawford and Sobel 1982) assume that conflicts of interest are public knowledge, and these models do not explicitly focus on their disclosure. However, both of these models imply that advisees would be better off if they knew about their advisors' conflicts of interests. Milgrom and Roberts (1986) also use an economic model to analyze advice-giving, but their model assumes that advisees can freely verify the accuracy of any advice given to them, so advisors who wish to mislead do best by withholding information. For example, a car salesman might tell the buyer that a car meets some manufacturing standard (a fact that is easily verified) but neglect to report that this standard is out of date. As Milgrom and Roberts acknowledge, however, perfect verifiability is not all that common in daily life. The current paper focuses on intermediate and more-realistic cases in which advice cannot be perfectly verified. For example, very few patients will ever learn if their physician prescribed a drug (perhaps due to a conflict of interest) that was more expensive and less effective than available alternatives.

In prior research (Cain, Loewenstein, and Moore 2005), we examined disclosure in an experimental setup in which "estimators" guessed the value of jars of coins and were paid according to the accuracy of their estimates. "Advisors" were given better information for evaluating coin-jar values and gave advice to the estimators. In different experimental conditions, advisors had incentives that were either aligned with or conflicted with estimators' incentives, and, when incentives conflicted, this was either disclosed to the estimators or not. We found that disclosure exacerbated the negative effects of the conflict of interest, producing more-biased advice and reducing estimators' payoffs. In this paper, we advance this line of research on several dimensions.

Most importantly, the earlier research did not provide clear evidence of the psychological mechanisms involved. The first two studies in this paper address this shortcoming, examining two possible mechanisms by which disclosure might bias advice. The Cain et al. (2005) study also suffered from methodological limitations. Most significantly, estimators and advisors tended to underestimate the value of the coin jars. By construction, the conflict of interest upwardly biased advice; therefore, even biased advice initially moved advisees closer to the truth (although the negative effects of disclosure were sufficiently strong that disclosure ultimately helped advisors and hurt advisees). In addition, although Cain et al. examined disclosure of a conflict of interest, it was not a fully factorial design that orthogonally manipulated disclosure and the presence or absence of a conflict of interest. The prior research also did not allow for reputation effects, since advisors were randomly re-paired with advisees after each round. Our third study replicates Cain et al.'s earlier study, eliminating these shortcomings, and does so in a more realistic, information-rich domain. Finally, moving beyond simply documenting the perverse effect of disclosure, the fourth study we report tests a potential intervention for increasing its effectiveness by increasing consumers' discounting of biased advice.

Conceptual Model

As the following conceptual model illustrates, disclosure helps advisees to the extent that they are able to estimate the impact of the disclosed conflict of interest (and the act of disclosure itself) on the advice they receive and, hence, correct for it. Disclosure can have adverse effects when these conditions are not met, which is often the case.

Figure 1 illustrates various possible effects of disclosure on advisee error in a situation in which an advisee is attempting to estimate some quantity (e.g., the market value of a house) and the advisor provides advice. In the situation depicted in the figure, the advisor has a conflict of interest such that he or she personally benefits from upwardly biasing the advisee. The figure illustrates several possible scenarios in which the advisor gives different advice, and the advisee (estimator) discounts that advice to a greater or lesser extent. The problem for the advisee is that disclosure can make advice worse, and discounting by the advisee is insufficient to offset this increased bias (let alone correct for the bias stemming from the conflict of interest itself).

Insert figure 1 about here

The x-axis of the figure represents the degree to which the advisor inflates the advice that he or she gives the advisee. The origin on the x-axis represents a situation in which advice is unbiased: The suggested value of the estimate ("Suggest") minus the actual value ("Actual") equals zero. The leftmost vertical (dotted) line represents a situation in which there is an undisclosed conflict of interest that motivates the advisor to provide upwardly biased advice. Finally, the right-most vertical (dotted) line represents a situation in which there is a disclosed conflict. Below, we explain why the disclosure line is likely to lie to the right of the line without disclosure—that is, why advice is likely to be more biased with disclosure than without it.

The three upward-sloping curved lines represent possible advisee responses to advice. The top-most line reflects a situation in which there is no disclosure, so the advisee does not suspect that the advisor has a conflict of interest. In this situation, the advisee is likely to take the advice verbatim (i.e., the response line runs at 45 degrees diagonally). Once advice becomes very extreme, however, the advisee begins to discount it, depicted by the declining slope. The point at which the response line peaks represents the advice that would maximize the advisor's payoff. Research on the effect of implausibly extreme advice (Mussweiler and Strack 2001) raises the possibility that there is no maximum—that estimates keep rising as advice gets more extreme. That research finds that the more extreme a random suggestion (the "anchor," or mental starting point), the more extreme the resulting estimate. In the domain of advice, however, it seems likely that at some point advice would be discounted so severely that the discounting would offset the effects of the greater exaggeration. In the case of a real estate agent interacting with a homebuyer, for example, it seems unlikely that it would be profitable for the agent to propose a price of $10 million for a house in the quarter-million range. The middle and bottom lines depict greater discounting of advice by advisees, as might occur if the advisors' interest in obtaining a high estimate was disclosed.

Finally, the y-axis represents the net effect of the bias in the advice given and any discounting of that advice. With an undisclosed conflict, advice is biased but not discounted by the advisee (unless it is very extreme), resulting in a commensurately biased estimate (point B on the y-axis). Whether disclosure ultimately results in more or less bias in the advisee's estimate depends on the degree of discounting (i.e., which of the two lower response lines best represent advisee responses). The middle advisee response line illustrates a situation in which discounting is inadequate, creating greater bias in the advisee's estimate (point A). The lowest advisee response line depicts a situation in which discounting is much greater, resulting in less bias in the advisee's estimate (point C). As we will argue, actual discounting caused by disclosure tends to be less than optimal, resulting in case A more frequently than C.

Suppose, however, that a conflict is disclosed but the advisor does not succumb to the conflict and instead offers unbiased advice, while the advisee discounts substantially in response to the disclosure. In this situation, disclosure will lead the advisee to underestimate the true value (point D). This effect shows how conflicts of interest can undermine the credibility (and hence usefulness) of advice from an advisor who provides accurate advice despite having a conflict of interest.

In sum, as the figure illustrates, whether disclosure hurts or helps the advisee depends on the net impact of disclosure on two competing effects: (1) bias in the advisor's suggestion and (2) discounting by the advisee. Next, we examine each effect in detail, and then, in study 3, we put the pieces together and empirically test disclosure's net effect on consumers of advice.

Psychological Mechanisms

How much bias will the conflicted advisor intentionally add? Two considerations are especially relevant which could be influenced by disclosure: strategic considerations and moral considerations (Cain et al. 2005). The first consideration is simple: What advice will maximize the advisee's (estimator's) estimate and, hence, the advisor's payoff? Turning back to figure 1, this is the point at which the advisor believes the advisee's response curve will peak. This same response curve appears as the top (dotted) curve in figure 2. Below that, figure 2 presents three possible ways in which the advisee's response curve might change in response to disclosure. In all three lines, disclosure leads to greater discounting across the board, as signified by their shift downward. When disclosure of the advisor's conflict of interest will lead the advisor to show greater restraint, discounting will decrease and the response curve will shift to the left (Church and Kuang 2009). If disclosure leads to exaggeration, discounting increases and the peak shifts to the right.

Insert figure 2 about here

If disclosure causes the peak of the response curve to shift rightward (or causes advisors to anticipate that it does), then a purely self-interested advisor might inflate advice further after offering disclosure. We call this *strategic exaggeration*. Car sellers often inflate their asking prices initially in anticipation of the buyer haggling downward. Similarly, conflicted advisors in general might offer more-biased advice to make up for the expected discounting when their conflict is disclosed.

The opposite effect can also occur. The advisor might expect an advisee who is aware of a conflict to be extra skeptical of extreme advice, leading the peak of the response curve to shift left. To avoid this, self-interested advisors might attempt to counteract the increased mistrust that disclosure brings by reigning in advice so that it looks realistic. We call this *strategic restraint*.

In contrast to these strategic considerations, moral considerations unambiguously cause advisors to increase exaggeration. In our model, maximizing advisor payoffs comes at the expense of the advisees, so advisors have a choice of how helpful versus hurtful their advice will be on any given response line. Even in one-shot dictator games (Forsythe et al. 1994), research has long shown that many people will share resources and show self-restraint towards anonymous others (Camerer 2003), especially when it is common knowledge that the recipient expects such benevolence (Dana, Cain, and Dawes 2006). Likewise, research on cheating behavior shows that people do not tend to cheat as much as they can get away with, only to the extent that they can rationalize to themselves (Mazar, Amir, and Ariely 2008). So, we predict that, especially without disclosure, advisors will show self-restraint and not give maximally biased advice.

When the welfare of others is a consideration, disclosure might reduce moral concerns. Prior research has suggested that when people demonstrate ethical behavior, they often become more likely to subsequently exhibit ethical lapses (Jordan, Mullen, and Murnighan 2008; Zhong, Liljenquist, and Cain 2009). For example, people who are given an opportunity to demonstrate

their own lack of prejudice are more likely to subsequently display discriminatory behavior (Monin and Miller 2001). Likewise, after a conflict of interest has been disclosed, advisors may feel that advisees have been warned and that advisors are "morally licensed" to provide biased advice.

Healy (2002) argues that "disclosure [often has] the effect of detaching the problem of honesty and bias from anybody in particular." Rules that mandate disclosure of conflicts of interest often make people feel that the outcomes of their actions, so long as they are at least minimally compliant, are the responsibility of the regulators. Disclosure of a conflict of interest can also reduce the perceived immorality of giving biased advice by signaling that bias is widespread and therefore less aberrant (Schultz et al. 2007). If advice recipients' expectations affect advisor behavior (Dana, Cain, and Dawes 2006), then the lowered expectations for honesty that come with disclosure might allow an advisor to rationalize providing biased advice because that it exactly what the advisee expects, or should expect, to receive.

Moral restraint can be represented by the discrepancy between the advice offered and the advice that the advisor thinks would result in a maximum payoff. As depicted in figure 3, moral licensing is the reduction of this restraint that is caused by disclosure. Studies 1 and 2 examine these moral and strategic mechanisms.

Insert figure 3 about here

STUDY 1: STRATEGIC EXAGGERATION AND STRATEGIC RESTRAINT

We designed our first study to address the question of whether and why advisors expect disclosure to shift the peak of the advisee response curve to the left or right.

Method

Three hundred sixty-four people were recruited by emailing an alumni list of Carnegie Mellon University. These participants completed an online survey for pay, with 1 in 50 participants winning a $50 Amazon gift card. The first part of the stimulus materials are presented in Appendix A. The survey asked people to imagine giving advice to another person (the "estimator"), who was trying to estimate how many jellybeans were in a jar that was depicted in a photo. Participants were all given a (hypothetical) conflict of interest: "You will be paid according to how much the estimator *overestimates* the number of jellybeans in the jar. *The higher the estimator's estimate (compared to the actual value), the more you get paid.*" Participants were also told this: "The true number of jellybeans in the jar is **2,400**. The estimator knows that you have better information than he or she has, but does not know that you know the true number. The estimator is merely told, 'There are *thousands* of jellybeans in the jar.'"

To minimize advisors' moral considerations for the estimators, estimators were to be imagined as being paid a flat rate for participation, not for accuracy. This was a 2 x 2 design; the first factor (disclosure vs. nondisclosure) varied within participants and the second factor (order

of the first factor) was counterbalanced between participants. In the nondisclosure condition, advisors were instructed, "Imagine that the estimator *does not* know about your payment incentive (that you earn more if he or she gives a higher estimate)." In the disclosure condition, advisors were instructed, "Imagine that the estimator *does* know about your payment incentive (that you will earn more if he or she gives a higher estimate)."

After each scenario, advisors were asked, "If your only goal was earning the highest payment for yourself, how would your suggestions across the two [disclosure vs. nondisclosure] scenarios compare . . ." and were then offered options of indicating that they would give a higher, lower, or same suggestion when the estimator knew about the incentive compared to when the estimator did not know. Next, advisors were prompted to give a numerical response indicating the exact advice that they would give in that scenario. After going through both the disclosure and nondisclosure scenarios, all advisors were asked to explain in their own words why their advice changed between conditions (if it did). Finally, all advisors (even those who did not change their advice across scenarios) were asked to select among possible reasons (the order of the first three reasons was counterbalanced) that best explained a change in advice in the presence of disclosure.

(1) "I think I should exaggerate less because the estimator might be suspicious of an excessively high number."
(2) "I think I should exaggerate more because the estimator, knowing I was biased, would adjust my advice downwards."
(3) "I think I should exaggerate more because the estimator, knowing I was biased, would reduce the weight he or she put on my advice."
(4) "Other."

Listing multiple interpretations of strategic exaggeration (#2, #3) allows that it might be over-selected, but the prior questions (and whether the advice actually increases or decreases) already are informative.

Results and Discussion

Seventeen participants were eliminated because either their first responses or the differences between their first and second responses were outliers by more than three standard deviations. Of the 347 participants who remained, 81 (23%) gave lower advice with disclosure, 104 (30%) gave higher advice with disclosure, and 162 (47%) gave the same advice across conditions. We note that a plurality of advisors said they would give the same advice with or without disclosure. Average advice (number of jellybeans) did not differ significantly between subjects if we compare the first response of those who got the disclosure condition first (M = 4017, SD = 1928) versus the first response of those who got the nondisclosure condition first (M = 4333, SD = 2092). There were also no significant differences within subject. The mean difference in advice (disclosure minus nondisclosure, within subject) was 13 jellybeans (SD = 1479) higher for nondisclosure when the nondisclosure condition came first, and was 1022 jellybeans (SD = 4661) higher for nondisclosure when the disclosure condition came first.

When presented with the above-four possible reasons for changing advice between disclosure conditions, 222 people gave a response: 34% chose reason #1 (exaggerate less, estimator might be suspicious), 29% chose reason #2 (exaggerate more, estimator will adjust downwards), 10% chose reason #3 (exaggerate more, estimator will reduce weight on advice), and the rest chose "other." This implies that 34% voted for strategic restraint and 39% voted for some form of strategic exaggeration.

Freehand responses (explaining the participants' actual thought processes) were coded by two design-blind research assistants, first coding independently, then communicating to settle all but 10 irreconcilable differences of opinion in coding (these 10 were thrown out). Ignoring 151 cells that were left blank and 19 that remained coded merely as "other," 167 freehand responses remained. Of these, 35% were coded as strategic restraint and 50% were coded as strategic exaggeration. The last 15% were divided into small groups in five miscellaneous categories; for example, the largest of these (with eight responses, or roughly 5%) suggested that, with disclosure, the estimator might help satisfy the advisor's (now known) interest in eliciting a high estimate by responding to high advice with a likewise high estimate.

In sum, study 1 suggests that disclosure is likely to cause some advisors to exaggerate their advice further; however, others are likely to rein in their advice, instead. Moreover, the

plurality of subjects stated that disclosure would not shift the profit-maximizing advice in either direction, at least when moral concerns were diminished. The upshot is that we require another mechanism to explain why disclosure might systematically worsen advice. Furthermore, the variance in the advisor's responses to disclosure illustrates another reason it is difficult for advice recipients to know how to adjust advice when it is disclosed as conflicted.

STUDY 2: MORAL LICENSING

Study 2 examined how disclosure effected advisors' perception of the morality of giving biased advice. Moral licensing suggests that offering biased advice will seem more morally acceptable with disclosure. While strategic reasons pull advice in both directions when disclosure is given, moral concerns tend to cause disclosure to exacerbate bias.

Method

One hundred seven people were recruited through a website run by Yale University and completed an online survey for pay. Participants were paid by lottery: Each had a 1 in 3 chance of winning a gift card worth $10 at Amazon.com. Personal identifiers and email addresses were separated from the data prior to analysis. Study 2 employed stimulus materials similar to those in study 1, except here, we merely told advisors that the jar contained between 1900 and 2900 jellybeans.

All participants were asked just two questions (the order counterbalanced between participants) on how they would rate the ethicality of suggesting "a number above 2900 (in hopes that the estimator overestimates the number of jellybeans)." One question specified disclosure ("The estimator is aware of your $50 incentive"), and one specified nondisclosure ("The estimator is unaware of your $50 incentive"). This made for a 2 (disclosure vs. nondisclosure) x 2 (question order) design. The first factor varied within participants, and the second factor varied between participants. Answers were ethicality ratings on a seven-item scale, ranging from "extremely ethical" to "extremely unethical."

Results and Discussion

There were no main or interaction effects of question order. Offering biased advice was rated as "somewhat unethical" without disclosure, whether this was asked first ($M = 5.4$, $SD = 1.37$) or last ($M = 5.38$, $SD = 1.7$), but with disclosure, it was rated as "somewhat ethical," whether this was asked first ($M = 3.58$, $SD = 1.8$) or last ($M = 3.72$, $SD = 1.63$), and, collapsing across question order, this disclosure difference was significant (paired $t[105] = 5.89$, $p < .001$). These results support the idea that disclosure lessens moral reluctance to providing biased advice.

The Advisee's Response

If advisees are aware that the act of disclosure can distort advice, they can, ostensibly, adjust for this distortion in addition to compensating for the conflict of interest. There are,

however, reasons to anticipate that advisees are unlikely to discount optimally. Research shows that people often "anchor" their perceptions on the information they receive initially and then make inadequate adjustments, even if they learn that the initial information was inaccurate or irrelevant (Tversky and Kahneman 1974; Chapman and Bornstein 1996; Hastie, Schkade, and Payne 1999; Skurnik et al. 2005). Moreover, the effect of anchoring may be especially prominent when it comes to conflicts of interest. People who are informed that an advisor has a conflict of interest are unlikely to know what to do about it. How much *should* one discount conflicted advice? Such a judgment will depend on a wide range of subsidiary judgments, including the ethicality of the advisor, whether the advisor is a "restrainer" or an "exaggerator," the cost of getting a second opinion, one's knowledge of the subject, and one's relationship with the advisor. Uncertainty about how to respond is likely to add noise to (but also decrease) advisees' discounting-response to advice.

STUDY 3: AN INTEGRATIVE STUDY OF THE EFFECTS OF DISCLOSURE

Study 3 examines the effect of disclosure of a conflict of interest on both the provider and recipient of advice, thus enabling us to test the model and the above-described effects. We paired "estimators" with "advisors" who had better information. Some advisors were given incentives that were aligned with those of the estimators, and some were given an incentive to bias the estimators. Crossed with the presence or absence of such a conflict of interest, we also manipulated whether incentives were or were not disclosed. We then examined the impact of disclosure on both parties.

As in our earlier research (Cain et al. 2005), we will make three predictions to replicate in this superior design. (1) Estimators' estimates will be less reliant on advisors' advice with disclosure than without disclosure, but, as the middle response line in figure 1 and the above discussion suggests, this discounting will not be sufficient. (2) As in figure 1, advisors with conflicts of interest will give more-biased advice under conditions with disclosure than without disclosure. (3) Estimators' estimates will be (a) higher and (b) more dispersed, and therefore less accurate with disclosure of conflicts of interest than without their disclosure, which will (c) lead to lower payoffs for estimators. Because the advisors' payoff schedules in this study varied between experimental conditions (even within disclosure conditions), it does not make sense to compare the effect of condition on advisor payoffs.

Our full factorial design includes a condition in which advisors honestly disclose that their interests are aligned with the estimator's. We expect the estimators in this condition to behave much as those in the nondisclosure conditions, implying that estimators' default assumption is that advisors have their best interests at heart unless they learn otherwise. Finally, this design allows us to test whether the perverse effects of disclosure we document will be robust to (limited) experience and feedback, even when advisor and estimator gain experience with one another and reputation formation is possible.

Method

Two hundred sixty-one members of the Carnegie Mellon University community were recruited with a promise of "$8-$15 per hour with an average of $10" and were run in groups of

six to 10. Individuals were randomly assigned to either the role of advisor or estimator. These roles were maintained throughout the experiment. Each estimator was randomly paired with an advisor in a pairing that lasted all four rounds of the exercise.

The task for estimators was to estimate actual sale prices of four pieces of local real estate: homes close to campus that had been listed for sale in a Yahoo! real estate database (http://realestate.yahoo.com/) and eventually sold for prices that were publicly available online. There were four houses, which varied in presentation order by session. The homes, labeled by their street addresses, had sold for the following amounts: House #5392 = $200,384; House #7518 = $186,250; House #5248 = $175,000; House #5301 = $199,900. Estimators were given an information packet on each house and as much time to examine it as they needed. Information packets contained an exterior photo of the house, a map of its location, and basic information about the property (number of bedrooms, number of bathrooms, total number of rooms, number of floors, year of construction, exterior construction material, roof material, heating system type, square footage, lot size, cooling system, fireplaces, and garage spaces). Sample stimulus materials (of House #5392) are included in appendix B.

Prior to seeing this information, estimators were given advice from advisors who knew that they had all the information the estimators had plus information the estimators did not have about recent sale prices, tax-assessed values of comparable neighborhood homes, and the tax-assessed value of the home in question. Advisors wrote their suggested sale prices on an "advisor's report" that was transmitted by the experimenter to the estimator. Appendix C contains a sample advisor's report. Each advisor's report had a space for the estimator to write his or her own estimate directly under the advisor's suggested sale price. Estimators were informed that they would get one report per round, each time from the same advisor.

There were 126 participants in the role of advisor and 135 in the role of estimator. Whenever there were an odd number of participants in a session, one advisor's report was randomly selected for duplication (nine advisors' reports were duplicated this way). Our analysis of advisors does not include duplicated advice, but our analysis of estimators does, since an estimator's reaction to even duplicated advice is informative. After examining their advisor's report and their own information packet, estimators wrote down their estimate of the selling price of the property. All procedures were explained to participants before the experiment began.

Estimators were always paid more for accurate estimates. Advisors' pay depended on the condition to which they were assigned. There were four conditions:

1. "Accurate–undisclosed" advisors were paid the same as their estimators. Advisor incentives were not mentioned to the estimators, who were told that they would be paid more the closer their estimate was to the actual sale price (see appendix D).
2. "Accurate–disclosed" advisors, as was disclosed on the advisor's report, were paid more when their estimators were accurate. Advisors in this condition were instructed to print the following disclosure, neatly and exactly, immediately under their suggestions: "As an advisor, I am required to inform you that I am paid based on how *accurate* your estimate of the property sale price is relative to the actual sale price."
3. "High–undisclosed" advisors were paid based on how high their estimators' estimates were compared to the actual sale price (see appendix E).

4. "High–disclosed" advisors were also paid based on how high their estimators' estimates were compared to the actual sale price, and were required to provide the estimator with a handwritten disclosure: "As an advisor, I am required to inform you that I am paid based on how *high* your estimate of the property sale price is relative to the actual sale price." Advisors in the high–undisclosed and the high–disclosed conditions were paid at the same rate (see appendix E).

After completing the advisor's report, all advisors were asked to make their own best estimates of the sale price, and were additionally rewarded based on the accuracy of these personal estimates (see appendix F).

Participants were told that neither estimators nor advisors would receive feedback about their actual payoffs or about actual sale prices during the first two (of four) rounds. At the beginning of each of the last two rounds, however, each advisor was shown the estimate of the estimator to whom his or her advice was given in the previous round, and the actual sale price of the house in question was publicly announced to everyone. In other words, at the beginning of round 3, everyone received feedback on what occurred in round 2 (and at the beginning of round 4, feedback on round 3). Since participants received payoff schedules, feedback allowed both advisors and estimators to calculate their own payoffs from the previous round before continuing on to the next round. Additionally, each advisor saw a copy of the estimator's instructions and thus could also use this feedback to calculate the payoffs of the estimator with whom the advisor was paired.

At the experiment's conclusion, one of the four rounds was randomly selected for computing actual payoffs. Participants received a $7.50 base payment plus any money earned in the payoff round, paid in cash. Participants were encouraged to ask clarifying questions (there were few), which were answered by simple yes or no answers. After being paid, participants were debriefed and dismissed.

Results and Discussion

The two disclosure conditions (#2 – *accurate* and #4 – *high*) gave the advisors very different objectives, so this created four distinct conditions. Except where otherwise mentioned, we subjected our results to one-way analysis of variance (ANOVA) with planned contrasts. For all analyses, the dependent variable was some type of participant response (e.g., suggestions) averaged together across the four rounds. Planned contrasts compared the two accurate conditions against the two high conditions (to test the effect of incentives: accurate vs. high), and compared the high–undisclosed condition against the high–disclosed condition (to test the effect of disclosing vs. not disclosing a conflict of interest). We also verified that the two accurate conditions were not significantly different in any of our tests.

Advisors' Personal Estimates

In the two accurate-condition groups, advisors were relatively accurate in their estimates, suggesting that these stimuli were superior to that used in our prior research, in which underestimation even in accurate conditions corrupted the results. Actual sale prices ranged from

roughly $175,000 to just over $200,000, with advisors' mean personal estimates ranging from roughly $165,000 to $220,000.

House #	#5248	#7518	#5301	#5392
Actual sale price	$175,000	$186,250	$199,900	$200,384
Advisor estimates	$165,618	$202,022	$223,008	$220,391

Three of the house prices were, on average, overestimated, and one was underestimated. Mean estimates were roughly correlated with actual prices.

Advisors' Suggestions

Table 1 presents mean advisor personal estimates for all conditions (row 1), suggestions (row 2), the difference between advisor's suggestions and actual sale prices (row 3), and the difference between advisors' suggestions and their personal estimates (row 4). The table shows that advisors' suggestions (row 2) and personal estimates (row 1) were similar in the two accurate conditions. Advisors in the accurate conditions had no incentive to bias their advice, except possibly to counteract any anticipated estimator bias. Disclosure did not change this fact, and there were no significant differences (in any of our comparisons) between the accurate and accurate–disclosed conditions.

Insert table 1 about here

Suggestions were inflated in the high (conflict of interest) conditions, especially in the high–disclosed condition. The mean suggestion given across houses was $204,331 in the accurate–undisclosed condition and $204,640 in the accurate–disclosed condition, but $236,138 in the high–undisclosed condition and $255,394 in the high–disclosed condition. Figure 4, which displays mean advisors' suggestions broken down by condition and round, shows that this basic pattern occurred in every round.

Insert figure 4 about here

Planned comparisons show, as expected, that advisors gave sale-price suggestions that were higher (more than $40,000 higher, on average) in the conflict-of-interest conditions than in the accurate conditions ($t[122] = 5.98$, $p < .001$). Rows 3 and 4 of table 1 provide two additional measures of advisors' propensity to exaggerate house prices in the four conditions, one relative to actual sale prices, and the other relative to advisors' own estimates. Planned contrasts again showed that conflicted advisors exaggerated more than nonconflicted advisors for both measures

(t[122] = 5.87, $p < .001$, relative to actual prices; t[118] = 6.09, $p < .001$, relative to personal estimates).

Planned contrasts involving all three measures of exaggeration also reveal that disclosure led to greater distortion of advice when there was a conflict of interest. In the high–undisclosed condition, advisors gave suggestions that averaged $31,351 greater than their own estimates; but with disclosure, their suggestions averaged $51,562 above their own estimates. These values are significantly different from one another (t[122] = 2.24, $p < .05$). The third row displays a similar pattern of exaggeration; advisors in the high–disclosed condition also gave suggestions that were more inflated relative to actual prices than did advisors in the high–undisclosed condition (t[128] = 2.13, $p < .05$). Finally, as evident in the second row of the table, suggestions provided by advisors were more inflated in the high–disclosed condition than in the high–undisclosed condition (t[118] = 2.5, $p < .05$). These results support prediction 1.

Discounting of Advice by Estimators

Table 2 summarizes results for estimators' estimates. As the first row of table 2 shows, estimators discounted advisors' advice in the two conflict-of-interest conditions. Discounting, as measured by subtracting estimators' estimates from the suggestions they received, was greater in the two conflict-of-interest conditions than in the accurate conditions, as revealed by a planned contrast (t[102] = 5.9, $p < .001$; unequal variances accounted for). Table 2 shows that estimators in the accurate-disclosed condition behave much like those in the accurate-undisclosed condition.

The difference in discounting between the high–disclosed and high–undisclosed conditions was marginally significant, according to a planned contrast (t[75] = 1.81, $p = .07$. However, as the following analysis shows (supporting prediction 2), this marginally increased discounting was insufficient to compensate for the increased bias offered by advisors in the high–disclosed condition.

Estimators' Estimates

The second row of table 2 presents mean estimator estimates in the four conditions. The results reveal a significant effect of advisor incentives, (F[1, 134] = 6.1, $p < .01$). One can see (by comparing columns 1 and 2 to 3 and 4) that estimates were over $20,000 higher in the two conflict conditions than in the two accurate conditions. Mean estimates in the high–disclosed condition were over $25,000 higher than in either accurate condition, and this difference was significant by planned contrasts (t[131] = 3.67, $p < .001$, compared to those in the accurate condition; t[131] = 3.18, $p < .01$, compared to those in the accurate–disclosed condition). Thus, disclosure did not compensate for the damaging effects of conflicts of interest. In fact, as row 2 of table 2 shows, in the conflict of interest conditions, mean estimator estimates were more than $8,000 *higher* with disclosure than without disclosure, although the difference was not statistically significant (t[131] = 1.21, $p = .23$), contrary to prediction 3a. Also, contrary to 3b, estimator estimates were not significantly higher in variance with disclosure than without disclosure in the conflict-of-interest conditions, by Levene's test ($p = .31$). The estimates and standard deviations are listed in table 2.

Insert table 2 about here

The Bottom Line

Table 3 and figure 5 summarize payoffs in the four experimental conditions. The basic pattern of results does reveal a perverse effect of disclosure of conflict of interest. With the given setup, the combination of (nonsignificantly) higher mean estimates and (nonsignificantly) higher variance in these estimates made estimators earn over one-third less money per house when conflicts of interest were disclosed than when they were not, and this difference was significant ($t[131] = 2.2$, $p < .05$), supporting prediction 3c. The key finding is that disclosure hurt the financial outcomes of those it was supposed to protect.

Insert table 3 about here

Insert figure 5 about here

Effects of Feedback

In the last two rounds, estimators and advisors were told the actual sale price for the house from the previous round. This is not an ideal setup to examine the impact of feedback, since feedback about sale prices was confounded with simple experience (Koch and Schmidt 2009). That said, neither experience with the task nor feedback decreased disclosure's biasing effect. The only effect of feedback on discounting was found in the high–undisclosed condition, where one round of feedback marginally increased discounting ($t[32] = 1.98$, $p = .056$, by paired-samples t-test of round 2 vs. round 3).

Looking at the overestimation of sale prices, we examined the effect of feedback using a 4 (condition) x 2 (feedback) x 2 (round) ANOVA with repeated measures on the last two factors. Neither feedback's main effect nor any interaction with the other variables was statistically significant. It could nevertheless be possible that more feedback and more rounds would have taught estimators how their judgments were being affected by the advice they received, and perhaps disclosure would have enhanced such learning. Note, however, that few real-life house purchasers receive anywhere near the quality and amount of feedback received by the participants in our experiment. In fact, we would guess that unless they take active steps to obtain such information, most house buyers are unlikely to receive any feedback about the prices of houses they almost purchased ultimately sold for. Moreover, any benefits of providing feedback to consumers must be tempered by the commensurate benefits that such feedback confers to those who wish to manipulate consumers; after all, advisors learn from feedback, too.

Insert figure 6 about here

Figure 6 provides a summary of our overall findings and can be compared to our earlier conceptual model (figure 1). Figure 6 is simplified by assuming that the response functions (both with and without disclosure) are straight lines that project from the no-bias point. The x-axis depicts advisor suggestions in the four conditions (normalized by actual house prices, which involved subtracting the actual home price from advisors' suggestions). The y-axis depicts estimators' estimates (again normalized by actual prices, estimate minus actual). The regression lines show the relationship between the suggestions that estimators received and the estimates they made. Flatter slopes indicate greater discounting of advice. The intersecting (horizontal and vertical) lines cross the regression lines where the mean for that condition lies: Horizontal lines represent the mean "estimate minus actual" (bias in estimates) for each condition, and vertical lines represent the mean "suggestion minus actual" (bias in suggestions) for each condition. The vertical lines' intersections with the x-axis show that advisor suggestions were higher in the two high conditions than in the two accurate conditions and higher in the high–disclosed condition than in the high–undisclosed condition. The graph also shows that advisor suggestions were discounted to a greater degree in the high–disclosed condition than in the high–undisclosed condition, but that this discounting was insufficient to offset the greater exaggeration by advisors in the high–disclosed condition. Hence (as the horizontal lines show), the estimates are higher relative to actual prices in the high–disclosed condition than in the high–undisclosed condition. So, in this study, as depicted in the middle response line (which illustrates disclosure with some discounting) in figure 1's conceptual model, the impact of disclosure on the exaggeration of advice outweighed the discounting effect.

How to Enhance Discounting

Looking back at our original model in figure 1, the question remains: How can we enhance the intended discounting effect of disclosure so that advisees respond to conflicted advice as they are depicted on the lowest (optimal) response line in our model, instead of the middle (perverse) response line? The limited discounting in our model suggests that conflicted advice is difficult for advisees to evaluate. Study 3 manipulated disclosure between subjects; therefore, participants to whom the conflict of interest was disclosed had no referent against which to compare their situation or the advice they received. Lacking any point of comparison, participants may not have appreciated the significance of the disclosure. Disclosures in the real world—outside the lab—are likely to share this feature. For example, if a doctor discloses receiving payments from the manufacturer of a drug he or she is prescribing, the patient has little basis for evaluating the severity of the conflict of interest or how it may have affected the doctor. Research on joint–separate preference reversals (Hsee et al. 1999) shows that the weight placed on attributes that are difficult to evaluate tends to be higher in joint evaluation (where the evaluator sees stimuli with the attribute next to the same stimuli without that attribute). Based on this insight, we hypothesized that advisees might discount biased advice more if an explicit contrast was drawn between biased advice disclosed as conflicted and unbiased advice known to be unconflicted; this explicit comparison makes salient the degree to which a conflict of interest may be skewing the advice.

STUDY 4: ENHANCING DISCOUNTING OF CONFLICTED ADVICE

Method

One hundred ninety-two passengers on a ferry from Connecticut to Long Island participated by completing a survey (for $2 pay). The instructions read as follows:

> In this survey, you will imagine yourself receiving advice, and you will respond with how likely you are to follow that advice. . . . Suppose you are currently suffering from an inflammatory condition of the thyroid. You have consulted several doctors, all local endocrinologists. The doctors are aware of the standard treatment(s) for your condition and that you are seeking multiple opinions. The doctors may disagree on what is the best treatment, but they insist that they make their recommendations with only your health in mind. You are to respond with how likely you are to take Dr. Glenn's advice (which will follow on the next page), given whatever you are told and whatever context you would imagine is most likely.

What followed varied according to the eight conditions, explained below (and fully shown in appendix G). Participants received (a) only one piece of advice, always from "Dr. Glenn," or (b) two pieces of advice, the first always from Dr. Glenn and the second always from "Dr. Andrew," who always gave advice divergent from Dr. Glenn. The advice from Dr. Glenn either (a) conformed to the standard treatment or (b) diverged from the standard treatment; abnormal prescriptions were always "three months more than the normal dosage." And, while the main instructions indicated that the doctors claimed to only have their patients' health in mind, each piece of advice came with a disclosure of either (a) no financial ties to the manufacturer of the drug, or (b) a potential conflict of interest in the form of the doctor having consulting ties to the manufacturer of the drug. All of these manipulations were combined to create eight conditions that permitted us to perform the planned contrasts that we present in the results.

We predict that biased advice will not be significantly discounted, even when (i) it is known to be abnormal advice, and (ii) even when that abnormal advice is disclosed as coming from a conflicted advisor, *unless* (iii) it is explicitly contrasted with normal advice and (iv) the normal advice is known as coming from an unconflicted advisor. These caveats are listed below when Dr. Glenn's advice satisfies them; *only in condition four* does his advice satisfy all caveats.

1.	AP-COI	abnormal prescription ("AP"), conflict of interest ("COI"), one doctor [i, ii]
2.	AP-NC	abnormal prescription, no conflict of interest ("NC"), one doctor [i]
3.	AP-COI/NP-COI	two doctors, the first suggesting abnormal prescription and having conflict of interest, the second suggesting normal prescription ("NP") and also having conflict of interest (i, ii, iii)
4.	AP-COI/NP-NC	two doctors, the first suggesting abnormal prescription and having conflict of interest, the second suggesting normal prescription but having no conflict of interest (all four: i, ii, iii, iv)

5.	AP-NC/NP-COI	two doctors, the first suggesting abnormal prescription but having no conflict of interest, the second suggesting normal prescription and having conflict of interest (i, iii)
6.	AP-NC/NP-NC	two doctors, the first suggesting abnormal prescription but having no conflict of interest, the second suggesting normal prescription and having no conflict of interest (i, iii, iv)
7.	NP-NC	one doctor, suggesting normal prescription and having no conflict of interest
8.	NP-COI	one doctor, suggesting normal prescription but having conflict of interest

Then, all participants were asked to rate the first (i.e., Dr. Glenn's) advice as follows: "Please rate how likely you are to take Dr. Glenn's advice in this situation." These ratings were elicited with a seven-point scale, which ranged from "certainly not" to "certainly." Our prediction implies that only the fourth condition is an ideal condition for participants to (perhaps correctly) worry about Dr. Glenn's advice. Our prediction also implies that, without satisfying all four caveats, even abnormal advice will seem quite agreeable despite being disclosed as coming from an advisor with a conflict of interest (as in condition 1). In this way, disclosure might fail to encourage advisees to search out a second opinion if the first opinion still seems sufficiently agreeable.

Results and Discussion

For all analyses, the dependent variable was the participants' responses on the seven-point scale, which asked them to rate how likely they would be to take the first doctor's advice. The results are summarized in table 4 and figure 7.

We made two initial comparisons by examining Dr. Glenn's advice in those conditions where it came alone (conditions 1, 2, 7, and 8). First, we wanted to know if participants were more likely to accept explicitly "normal" advice (one-month dosage) over explicitly "abnormal" advice (four-month dosage), regardless of conflicts of interest. To answer this question, we aggregated responses from conditions 7 and 8 (to summarize responses to normal advice), and we aggregated responses from conditions 1 and 2 (to summarize responses to abnormal advice), and we then ran a *t*-test to compare the two aggregations against each other. Indeed, participants preferred normal advice: Respondents were more likely to say they would accept normal advice (mean rating of 5.48, corresponding to between "probably" and "very likely") than abnormal advice (mean rating of 4.62; between "possibly" and "probably"), and this difference was significant ($t[1, 94] = 3.1$, $p < .01$). Table 4 also shows how these normal-versus-abnormal advice conditions compare without aggregating.

Insert table 4 about here

Second, we wanted to know if participants had a general preference for doctors who had no conflicts of interest, regardless of the advice that doctor gave. To answer this, we ran a *t*-test similar to the one just described by aggregating conditions 7 and 2 (where Dr. Glenn has no

conflict of interest) and comparing that aggregate against conditions 1 and 8 combined (where Dr. Glenn's conflict of interest is disclosed). Indeed, participants preferred no conflicts: No-conflict conditions had a combined mean of 5.52, while conflict-of-interest conditions had a significantly lower combined mean of 4.58 ($t[1, 94] = 3.48$, $p < .001$) (see table 4).

Our main prediction was that participants would discount the conflicted advice most in condition four. To examine this, we subjected our results to a one-way ANOVA that used planned contrasts to compare each of the eight conditions against each other. Planned contrasts employed Games-Howell post-hoc tests, which do not assume equal variance (results appear in table 4). We first contrasted condition 1 (AP-COI: Dr. Glenn suggesting an abnormally high dosage, with a disclosed conflict of interest) against condition 4 (AP-COI/NP-NC: same advice and disclosure from Dr. Glenn, but this time contrasted with normal advice from Dr. Andrew, here known to have no conflicts of interest). Indeed, an unbiased second opinion reduced the rating of Dr. Glenn's advice. Condition 1 showed a mean of 4.38, which translated on the response scale to between "possibly" and "probably," while condition 4 showed a lower mean of 2.79, or between "very unlikely" and "improbably" ($p < .01$). When Dr. Glenn was the only source of advice, however, disclosure alone had little effect. For example, there is no significant difference in participant's willingness to take Dr. Glenn's abnormal advice between condition 1 (where a conflict of interest is disclosed, mean result: 4.38) and any of the conditions where Dr. Glenn gave the same abnormal advice but is described as having no conflict of interest (conditions 2, 5, and 6, with means 4.88, 4.33, and 3.58 respectively).

Insert figure 7 about here

Second opinions often do not arrive automatically with the first opinion, but are often available, at some cost. Two issues arise here. First, if people are uncertain how much a conflict of interest might have biased any particular piece of advice, they might not see the value of obtaining a second opinion. Second, even if people see the value of a getting a second opinion, doing so will probably take time, effort, and money. At the same time as it corroborates the benefits of obtaining an unbiased opinion (Robertson 2010), the lack of discounting shown in study 4 (in conditions where there was a single conflicted doctor) suggests that disclosure alone might not move people to get a second opinion, even when they know the first opinion is biased. In combination, these results suggest that it might be worthwhile to provide incentives or even a requirement to obtain second opinions; at least they suggest that disclosure often requires a particular set of circumstances to succeed as a discounting cue.

A POLICY-ORIENTED GENERAL DISCUSSION

In combination, these four studies suggest that disclosure is not a panacea for problems created by conflicts of interest. In fact, this research shows how disclosure can hurt exactly the people it is intended to protect. Moreover, in the real world, additional factors are likely to come into play and further exacerbate the perverse effects shown in these studies. For example, when a disclosure is made in person, the advice recipient may trust the advisor *more* as a result of the

disclosure: those who disclose their conflicts of interests may appear to be "forthcoming." Disclosure can also change the nature of interactions between parties, making it more uncomfortable for an advice recipient to reject advice. Since most people view succumbing to a conflict of interest as a matter of corruption rather than unconscious bias (Bazerman, Loewenstein, and Moore 2002; Cain and Detksy 2008), failing to heed an advisor could express that the advisee thinks the advisor is morally corrupt (Sah, Loewenstein, and Cain 2009). For example, if a doctor suggests to a patient that she enroll in an experimental drug trial, and then discloses that he gets $5,000 if a patient enrolls, the patient might feel pressured to enroll so as not to seem to doubt the doctor's integrity.

Why is the call for disclosure so popular despite how it can backfire? One possible explanation is that most people are simply not aware of disclosure's pitfalls. At first glance, disclosure seems like a sensible remedy to a situation in which one party possesses an otherwise hidden incentive to mislead another party. A more cynical explanation would play on the "Chicago Theory of Regulation" (Stigler 1971; Peltzman 1976; Becker 1983), which posits that regulation typically exists not for the general benefit of society but for the benefit of the regulated groups. These entities might be aware of the ineffectiveness of disclosure but accept it because it *benefits them*. For example, even though consumer advocates fought hard for warning labels on cigarette packages, the tobacco industry has defended itself against litigation since then by citing the warning labels as evidence that consumers knew the risks. "What was intended as a burden on tobacco became a shield instead" (Action on Smoking and Health 2001). Moreover, even the *regulators* may be attracted to disclosure if they see it as absolving them of responsibility for protecting consumers by ostensibly empowering consumers to protect themselves. Disclosure may also be perceived as the lesser of evils for those who might otherwise face more substantive regulation. For example, pharmaceutical firms are often strong proponents of disclosure laws, since it is better for them (and for researchers who receive their funding) if researchers must *disclose* financial ties to the industry rather than actually having to *sever* them. This all suggests that disclosure may be problematic for more reasons than those identified by the experiments reported above.

The most effective antidote for the problems caused by conflicts of interest is not to disclose them, but to eliminate them. Physicians, for example, could (and, we believe, should) be prohibited from accepting gifts from pharmaceutical companies. Investment banks could be barred from providing buy/sell recommendations on the stocks of companies whose issues they underwrite. Bond-rating firms could be paid by those who use the information they generate rather than by the companies whose bonds they rate. Even if disclosure does no direct harm (e.g., if it does not morally or strategically license bias, etc.), it can have a pernicious effect if it substitutes for more-effective regulations, thereby morally licensing policy makers to not take more substantive measures to deal with conflicts.

Granted, eliminating conflicts of interest could be prohibitively costly in some cases. Reducing conflicts of interest in physicians' treatment recommendations, for example, might mean that patients always receive their diagnoses and treatments from different people, which could greatly increase the cost of medical care (although the reduction of conflicts of interest would have countervailing effects). Whether the benefits outweigh the costs needs to be judged on a case-by-case basis. The point of this paper is that the benefits of disclosure are easy to overestimate.

It would be a mistake, however, to conclude that disclosure is always counterproductive, as some recent laboratory research illustrates (Church and Kuang 2009; Koch and Schmidt 2009). Research on practical examples of disclosure, summarized in *Full Disclosure* (Fung, Graham, and Weil 2007), also shows that disclosure can have real beneficial effects. For example, following a spate of highly publicized SUV rollovers, regulations that required auto manufacturers to publicly disclose rollover ratings led to significant and rapid changes in auto design, resulting in a general decrease in the rollover risk for SUVs. Disclosure is likely to be helpful when information is disclosed in an easily digestible form (or is made available to intermediaries, e.g., ratings companies, who process it for consumers) and when it is clear how one should respond to the disclosed information. The rollover ratings met both criteria: The ratings were represented simply as one to five stars, making it easy for consumers to compare—i.e., evaluate jointly—the relative rollover risks of various SUVs. Even when information isn't presented in such a simple form, disclosure is likely to prove helpful when the recipients are savvy repeat-players who know what to do with the disclosed information, such as institutional investors, experienced attorneys, or managers in government agencies (Church and Kuang 2009; Malmendier and Shanthikumar 2008). Disclosure is much less likely to help individuals such as personal investors, purchasers of insurance, homebuyers, or patients, who are unlikely to possess the knowledge or experience to know how much they should discount advice or whether they should get a second opinion in a given conflict-of-interest situation (Malmendier and Shanthikumar 2007).

However, even when advisees pay little attention to disclosed information, it can have a "telltale heart" effect and cause advisors to exercise restraint. Furthermore, transparency through disclosure can be viewed as an inherently desirable feature of society, even if it sometimes leads to adverse consequences. Most of us want to know if someone has a motivation to mislead us. As long as disclosure is not viewed as an equal substitute for elimination of conflict, and as long as disclosed information is interpretable by those who receive it (and not written in fine-print legalese), the benefits of disclosure may outweigh the risks delineated in this paper.

In conclusion, we echo the sentiments of Suroweicki (2006) in concluding that transparency may be good, but objectivity is even better: Regulators should be focusing less on disclosing sources of bias, and more on insuring that objective information reaches the audience, if not in lieu of biased information, at least directly alongside it.

APPENDIX A
SAMPLE STIMULUS MATERIAL FOR STUDY 1

Imagine you are randomly and anonymously paired with another person on an estimation task. Your role is "**advisor**." The other person's role is "**estimator**." The estimator is shown the picture below and is asked to guess how many jellybeans are in the jar.

However, before the estimator guesses the number of jellybeans, you will give him or her a suggestion about how many jellybeans are in the jar. The true number of jellybeans in the jar is **2,400**. The estimator knows that you have better information than he or she has, but does not know that you know the true number. The estimator is merely told, "There are *thousands* of jellybeans in the jar."

The estimator will be paid $10 for participating, so his or her earnings will not depend on how accurate he or she is; but assume that the estimator will try to make as accurate an estimate as possible.

YOUR PAYMENT: You will be paid according to how much the estimator *overestimates* the number of jellybeans in the jar. *The higher the estimator's estimate (compared to the actual value), the more you get paid.*

[Instructions continue as explained in main text…]

APPENDIX B

SAMPLE STIMULUS MATERIAL FOR ADVISORS

Note: Estimators received this information also, except that the information highlighted in gray (Total Market Value, and TMV/Sale-price of comparable homes) was replaced by "Only advisors have this info."

APPRAISER INFO: 5392 Wilkins Ave.

(*Code:5392)*

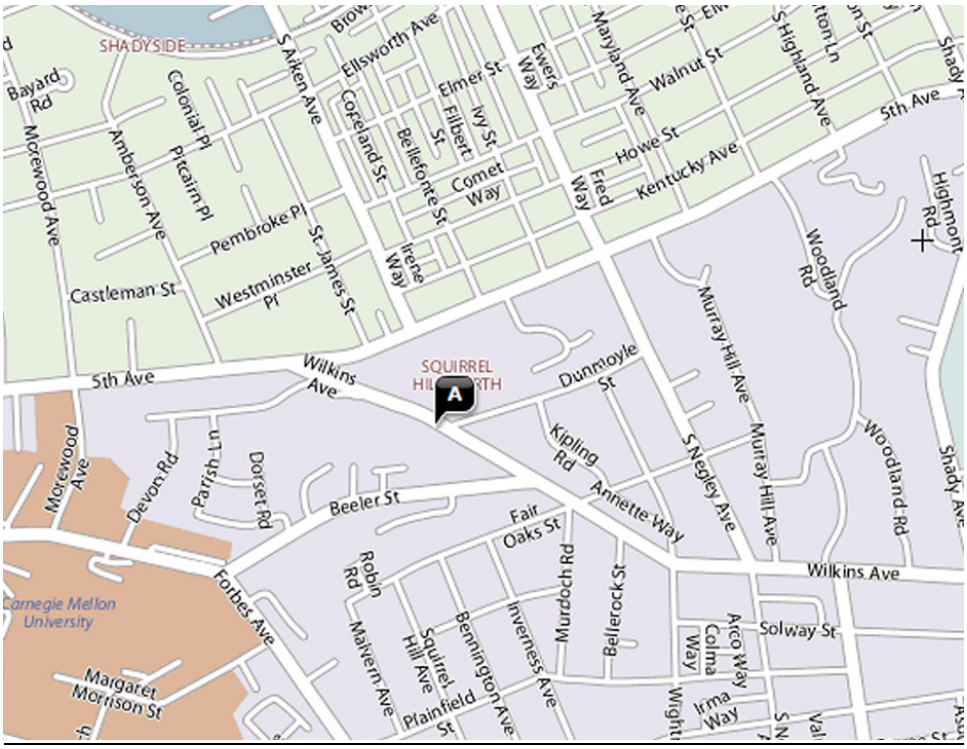
Sale Date:	3/20/2002
Sale Price:	**$???????**
Total Market Value	$238,200

LAND - PRIMARY SITE (6232 SQFT). 2 STY OLD STYLE HOUSE W/ PORCH FRAME - OPEN

Building Information

Total Rooms:	8		
Bedrooms:	4		
Stories:	2	Full Bathrooms:	2
Year Built:	1924	Half Bathrooms:	1
Exterior Finish:	Brick	Heating:	Central Heat
Roof:	Shingle	Cooling:	
Basement:	Full Basement	Fireplace(s):	1
		Garage:	0
Condition:	Good	Finished Living Area:	2160 Square Foot

Yahoo! Map of location:



Comparable Houses

Address	5262 BEELER	5136 BEELER
Year Built	1924	1929
Sale Price	$179,000	$215,000
Sale Date	07/31/2000	09/01/1999
Total M.V.	$185,400	$235,300

APPENDIX C

SAMPLE ADVISOR'S REPORT

[The disclosure note and other info shown in handwritten fonts were to be handprinted by advisors.]

Advisor's report and estimator's estimate

I have carefully examined the property information, along with its tax-assessed value and the sale-price of comparable houses. I suggest that it is worth:

For the property coded: 5392,
the suggested sale-price is:

$279,000

Advisor's participant code: S401

***Note:** As an advisor, I am required to inform you that I am paid based on how *high* your estimate of the property sale-price is relative to the actual sale-price.

To be completed by the estimator:

Please print neatly what you think was the property's sale-price at the time of sale:

$_____

Please enter your Participant Code in the blank below and raise your hand when you are finished.

Estimator participant's code:_____

APPENDIX D

ACCURATE ADVISOR PAYOFFS (SAME AS ESTIMATOR)

If advisor's personal estimate is within–	Advisor earns
$2,000 of sale price	$2.00
$5,000 of sale price	$1.75
$7,500 of sale price	$1.50
$10,000 of sale price	$1.25
$12,500 of sale price	$1.00
$15,000 of sale price	$0.75
$17,500 of sale price	$0.50
$20,000 of sale price	$0.25

APPENDIX E

CONFLICTED ADVISOR PAYOFFS

If estimate is above sale price by *at least*–	Advisor earns
$5,000	$1.00
$10,000	$1.10
$15,000	$1.30
$20,000	$1.60
$25,000	$2.00
$30,000	$2.50
$35,000	$3.10
$40,000	$3.80
$45,000	$4.60
$50,000	$5.50

APPENDIX F

ALL ADVISOR PAYOFFS FOR THEIR PERSONAL ESTIMATES

If advisor's personal estimate is within–	Advisor earns
$2,000 of sale price	$2.00
$5,000 of sale price	$1.75
$7,500 of sale price	$1.50
$10,000 of sale price	$1.25
$12,500 of sale price	$1.00
$15,000 of sale price	$0.75
$17,500 of sale price	$0.50
$20,000 of sale price	$0.25

APPENDIX G

SURVEYS FROM STUDY 4

Main Instructions: In this survey, you will imagine yourself receiving advice and you will respond with how likely you are to follow that advice… Suppose you are currently suffering from an inflammatory condition of the thyroid. You have consulted several doctors, all local endocrinologists. The doctors are aware of the standard treatment(s) for your condition and that you are seeking multiple opinions. The doctors may disagree on what is the best treatment, but they insist that they make their recommendations with only your health in mind. You are to respond with how likely you are to take Dr. Glenn's advice (which will follow on the next page), given whatever you are told and whatever context you would imagine is most likely.

[Survey Condition 1: AP-COI] [AP = Abnormal Prescription, COI = Conflict of Interest]
Dr. Glenn recommends that you take the standard oral medication for 4 months; although this is 3 months more than the normal dosage, Dr. Glenn explains that this recommendation is within safe tolerances. Dr. Glenn often does paid consulting for the drug manufacturer and is keenly aware of their product. Dr. Glenn explains that this dosage gives the drug its best chance for working on your particular condition.

[Survey Condition 2: AP-NC] [NC = No Conflict]
Dr. Glenn recommends that you take the standard oral medication for 4 months; although this is 3 months more than the normal dosage, Dr. Glenn explains that this recommendation is within safe tolerances. Dr. Glenn has no financial ties to the manufacturer of the drug and is keenly aware of their product. Dr. Glenn explains that this dosage gives the drug its best chance for working on your particular condition.

[Survey Condition 3: AP-COI/NP-COI] [NP = Normal Prescription, / divides two pieces of advice]
Dr. Glenn recommends that you take the standard oral medication for 4 months; although this is 3 months more than the normal dosage, Dr. Glenn explains that this recommendation is within safe tolerances. Dr. Glenn often does paid consulting for the drug manufacturer and is keenly aware of their product. Dr. Glenn explains that this dosage gives the drug its best chance for working on your particular condition.
Dr. Andrew recommends that you take the standard oral medication for 1 month, which is the normal dosage. Dr. Andrew often does paid consulting for the drug manufacturer and is keenly aware of their product. Dr. Andrew explains that this dosage is sufficient to provide every opportunity for the drugs to work.

[Survey Condition 4: AP-COI/NP-NC]
Dr. Glenn recommends that you take the standard oral medication for 4 months; although this is 3 months more than the normal dosage, Dr. Glenn explains that this recommendation is within safe tolerances. Dr. Glenn often does paid consulting for the drug manufacturer and is keenly aware of their product. Dr. Glenn explains that this dosage gives the drug its best chance for working on your particular condition.
Dr. Andrew recommends that you take the standard oral medication for 1 month, which is the normal dosage. Dr. Andrew has no financial ties to the manufacturer of the drug and is keenly aware of their product. Dr. Andrew explains that this dosage is sufficient to provide every opportunity for the drugs to work.

[Survey Condition 5: AP-NC/NP-COI]
Dr. Glenn recommends that you take the standard oral medication for 4 months; although this is 3 months more than the normal dosage, Dr. Glenn explains that this recommendation is within safe tolerances. Dr. Glenn has no financial ties to the manufacturer of the drug and is keenly aware of their product. Dr. Glenn explains that this dosage gives the drug its best chance for working on your particular condition.
Dr. Andrew recommends that you take the standard oral medication for 1 month, which is the normal dosage. Dr. Andrew often does paid consulting for the drug manufacturer and is keenly aware of their product. Dr. Andrew explains that this dosage is sufficient to provide every opportunity for the drugs to work.

[Survey Condition 6: AP-NC/NP-NC]

Dr. Glenn recommends that you take the standard oral medication for 4 months; although this is 3 months more than the normal dosage, Dr. Glenn explains that this recommendation is within safe tolerances. Dr. Glenn has no financial ties to the manufacturer of the drug and is keenly aware of their product. Dr. Glenn explains that this dosage gives the drug its best chance for working on your particular condition.
Dr. Andrew recommends that you take the standard oral medication for 1 month, which is the normal dosage. Dr. Andrew has no financial ties to the manufacturer of the drug and is keenly aware of their product. Dr. Andrew explains that this dosage is sufficient to provide every opportunity for the drugs to work.

[Survey Condition 7: NP-NC]
Dr. Glenn recommends that you take the standard oral medication for 1 month, which is the normal dosage. Dr. Glenn has no financial ties to the manufacturer of the drug and is keenly aware of their product. Dr. Glenn explains that this dosage is sufficient to provide every opportunity for the drugs to work.

[Survey Condition 8: NP-COI]
Dr. Glenn recommends that you take the standard oral medication for 1 month, which is the normal dosage. Dr. Glenn often does paid consulting for the drug manufacturer and is keenly aware of their product. Dr. Glenn explains that this dosage is sufficient to provide every opportunity for the drugs to work.

[Main Response Query – NB: scale was flipped in half of the trials]

Please rate how likely you are to take Dr. Glenn's advice in this situation
(Circle the number that best applies, using the key below)
1: Certainly Not (0%)
2: Very Unlikely (1-20%)
3: Improbably (21-40%)
4: Possibly (41-60%)
5: Probably (61-80%)
6: Very Likely (81-99%)
7: Certainly (100%)

REFERENCES

Action on Smoking and Health (2001), "Tobacco Deal Has Potential for Even Greater Dirty Tricks," http://www.no-smoking.org/sept97/9-11-97-1.html.

Bazerman, Max H., George Loewenstein, and Don A. Moore (2002), "Why Good Accountants Do Bad Audits," *Harvard Business Review*, (November), 96-103.

Becker, Gary S. (1983), "A Theory of Competition among Pressure Groups for Political Influence," *Quarterly Journal of Economics*, 98, 371-400.

Berner, Robert (2006), "I Sold It through the Grapevine," http://www.businessweek.com/magazine/content/06_22/b3986060.htm.

Cain, Daylian M., George Loewenstein, and Don A. Moore (2005), "The Dirt on Coming Clean: Perverse Effects of Disclosing Conflicts of Interest," *The Journal of Legal Studies,* 34, 1-25.

Camerer, Colin F (2003), *Behavioral Game Theory: Experiments on Strategic Interaction,* Princeton: Princeton University Press.

Chapman, Gretchen B. and Brian H. Bornstein (1996), "The More You Ask for the More You Get: Anchoring in Personal Injury Verdicts," *Applied Cognitive Psychology*, 10, 519-40.

Church, Bryan K., and Xi Jason Kuang (2009), "Conflicts of Interest, Disclosure, and (Costly) Sanctions: Experimental Evidence," *The Journal of Legal Studies*, 38, 505-32.

Crawford, Vincent P. and John Sobel (1982), "Strategic Information Transmission," *Econometrica*, 50 (6), 1431-51.

Dana, Jason D., Daylian M. Cain, and Robyn M. Dawes (2006), "What You Don't Know Won't Hurt Me: Costly (But Quiet) Exit in Dictator Games," *Organizational Behavior and Human Decision Processes,* Vol. 100, No. 2, 193-201.

Dye, Ronald A. (2001), "An Evaluation of Essays on Disclosure and the Disclosure Literature in Accounting," *Journal of Accounting and Economics*, 32, 181-235.

Forsythe, Robert, Joel L. Horowitz, N. E. Savin, and Martin Sefton (1994), "Fairness in Simple Bargaining Experiments," *Games and Economic Behavior*, 6, 347-69.

Fung, Archon, Mary Graham, and David Weil (2007), *Full Disclosure: The Perils and Promise of Transparency*, New York: Cambridge University Press.

Gunderson, Martin (1997), "Eliminating Conflicts of Interest in Managed Care Organizations through Disclosure and Consent," *Journal of Law, Medicine, and Ethics*, 27, 5-15.

Harris, Jared and David Souder (2004), "Bad Apples or Bad Bushel? Ethics, Efficiency and Capital Market Integrity," *Business and Professional Ethics Journal*, 23 (1), 201-02.

Hart, Philip (1975), *Congressional Record* (Senate), 94th Cong., 1st Sess.

Hastie, Reid, David A. Schkade, and John W. Payne (1999), "Juror Judgments in Civil Cases: Effects of Plaintiff's Requests and Plaintiff's Identity on Punitive Damage Awards," *Law and Human Behavior*, 23, 445-70.

Healy, Kieran (2002), "Market Discipline and Organizational Responsibility," http://www.kieranhealy.org/blog/archives/000175.html.

Healy, Paul M. and Krishna Palepu (2000), "Information Asymmetry, Corporate Disclosure, and the Capital Markets: A Review of the Empirical Disclosure Literature," http://papers.ssrn.com/sol3/papers.cfm?abstract_id=258514.

Holmstrom, Bengt (1977), "On Incentives and Control in Organizations," Ph.D. dissertation, Graduate School of Business, Stanford University, CA 94305.

Hsee, Christopher K., George Loewenstein, Sally Blount, and Max H. Bazerman (1999), "Preference Reversals between Joint and Separate Evaluations of Options: A Review and Theoretical Analysis," *Psychological Bulletin*, 125, 576-90.

Jordan, Jennifer, Elizabeth Mullen, and J. Keith Murnighan (2008), *Moral Miscreants and Immoral Angels: A Model of Moral Equilibrium,* unpublished article.

Koch, Christopher and Carsten Schmidt (2009), "Disclosing Conflict of Interest–Does Experience and Reputation Matter?" *Accounting, Organizations, and Society*, 35, 95-107.

Malmendier, Ulrike and Devin Shanthikumar (2007), "Are Small Investors Naïve about Incentives?" *Journal of Financial Economics*, 85, 457-89.

_____(2008), "Do Security Analysts Speak in Two Tongues?" Unpublished article.

Mazar, Nina, On Amir and Dan Ariely (2008), "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance," *Journal of Marketing Research*, 45, 633-44.

Milgrom, Paul and John Roberts (1986), "Relying on the Information of Interested Parties," *RAND Journal of Economics*, 17 (Spring), 18-32,

Monin, Benoit and Dale T. Miller (2001), "Moral Credentials and the Expression of Prejudice," *Journal of Personality and Social Psychology,* 81 (1), 33-43.

Mussweiler, Thomas and Fritz Strack (2001), "Considering the Impossible: Explaining the Effects of Implausible Anchors." *Social Cognition,* 19, 145-60.

New York Stock Exchange (2002), "Disclosure and Reporting Requirements," http://www.nyse.com/pdfs/ im02-26Microsoft%20Word%20-%20Document%20in%2002-26.pdf.

Peltzman, Sam (1976), "Toward a More General Theory of Regulation," *The Journal of Law and Economics*, 19, 211-40.

Robertson, Christopher T. (2010), "Blind Expertise," *New York University Law Review*, 85, 174-257.

Sah, Sunita, George Loewenstein, and Daylian M. Cain (2009), "The Burden of Disclosure," working paper, Carnegie Mellon University, Pittsburgh, PA 15213.

Schultz, P. Wesley, Jessica M. Nolan, Robert B. Cialdini, Noah J. Goldstein, and Vladas Griskevicius (2007), "The Constructive, Destructive, and Reconstructive Power of Social Norms," *Psychological Science*, 18, 429-34.

Skurnik, Ian, Carolyn Yoon, Denise C. Park, and Norbert Schwarz (2005), "How Warnings Become Recommendations," *Journal of Consumer Research*, 31, 713-24.

Stigler, George (1971), "The Theory of Economic Regulation," *Bell Journal of Economics and Management Science*, 2, 3-21.

Surowiecki, James (2002), "The Talking Cure," *New Yorker*, December 9.

Tversky, Amos and Daniel Kahneman (1974), "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185, 1124-31.

Verrecchia, Robert E. (2001), "Essays on Disclosure," *Journal of Accounting and Economics,* 32, 97-180.

Zhong, Chen-Bo, Katie Liljenquist, and Daylian M. Cain (2009), "Moral Self-Regulation: Licensing and Compensation," in *Psychological Perspectives on Ethical Behavior and Decision Making*, ed. David De Cremer, Greenwich, CT: Information Age Publishing.

Table 1
Advisor Exaggeration of Sale Prices

	Accurate–undisclosed n = 26	Accurate–disclosed n = 23	High–undisclosed n = 36	High–disclosed n = 41	Effect of condition Acc, Acc-D, High, High-D	Effect of incentives Accurates vs. Highs	Effect of disclosure High vs. High-D
Advisor's personal estimate	$202,978 (7,715)	$200,529 (13,449)	$203,205 (11,505)	$203,939 (12,102)	NS	NS	NS
Advisor's suggestion	$204,331 (6,841)	$204,640 (8440)	$236,138 (36,071)	$255,394 (55,877)	$p < .001$	$p < .001$	$p < .05$
Advisor suggestion minus actual	$14,040 (7,299)	$14,685 (7,988)	$45, 788 (36,007)	$64,1412 (56,079)	$p < .001$	$p < .001$	$p < .05$
Advisor suggestion minus advisor personal estimate	$1,142 (7,126)	$3,840 (7,410)	$31, 351 (33,393)	$51,562 (52,628)	$p < .001$	$p < .001$	$p < .05$

(Standard deviations are in parentheses.)

Table 2
Estimator Estimates

	Accurate–undisclosed n = 31	Accurate–disclosed n = 23	High–undisclosed n = 39	High–disclosed n = 42	Effect of condition Acc, Acc-D, High, High-D	Effect of incentives Accurates vs. Highs	Effect of disclosure High vs. High-D
Discounting: absolute value of (suggestion-estimator estimate)	$614 (10,228)	$805 (10,864)	$11,216 (25,983)	$25,609 (38,641)	$p < .001$	$p < .001$	$p = .07$
Estimator estimate	$202,529 (12,495)	$203,835 (13,038)	$221,209 (32,885)	$229,605 (43,613)	$p < .01$	$p < .01$	NS

(Standard deviations are in parentheses.)

Table 3
Simple Bottom Line: ANOVA on Estimator and Advisor Payoffs across All Rounds

	Accurate–undisclosed Est: n = 31 Adv: n = 26	Accurate–disclosed Est: n = 23 Adv: n = 23	High–undisclosed Est: n = 39 Adv: n = 36	High–disclosed Est: n = 42 Adv: n = 41	Effect of condition Acc, Acc-D, High, High-D	Effect of incentives Accurates vs. Highs	Effect of disclosure High vs. High-D
Estimator payoff	$1.86 (1.00)	$1.86 (1.00)	$1.37 (1.17)	$.87 (.92)	$p < .001$	$p < .001$	$p < .05$
Advisor payoff	$1.86 (.92)	$1.86 (1.00)	$2.67 (1.55)	$2.98 (1.75)	$p < .001$	$p < .001$	NS

(Standard deviations are in parentheses.)

Table 4
Study 4 Survey Results

	Mean (SD) See key†	1 AP-COI	2 AP-NC	3 AP-COI/ NP-COI	4 AP-COI/ NP-NC	5 AP-NC/ NP-COI	6 AP-NC/ NP-NC	7 NP-NC	8 NP-COI
1: AP-COI	4.38 (1.47)		NS	$p = .10$	**	NS	NS	***	NS
2: AP-NC	4.88 (1.4)	NS		**	***	NS	*	**	NS
3: AP-COI/NP-COI	3.25 (1.22)	$p = .10$	**		NS	$p = .10$	NS	***	**
4: AP-COI/NP-NC	2.79 (1.32)	**	***	NS		**	NS	***	***
5: AP-NC/NP-COI	4.33 (1.40)	NS	NS	$p = .10$	**		NS	***	NS
6: AP-NC/NP-NC	3.58 (1.38)	NS	*	NS	NS	NS		***	$p = .055$
7: NP-NC	6.17 (0.64)	***	**	***	***	***	***		**
8: NP-COI	4.79 (1.28)	NS	NS	**	***	NS	$p = .055$	**	

AP = Abnormal Prescription (4 months). NP = Normal Prescription (1 month). COI = Conflict. NC = No Conflict.

* $p < .05$ ** $p < .01$ *** $p < .001$ (All p-values generated by ANOVA, Games–Howell post hoc tests; NS = not significant)

†Key to MEAN:
1: Certainly Not (0%)
2: Very Unlikely (1-20%)
3: Improbably (21-40%)
4: Possibly (41-60%)
5: Probably (61-80%)
6: Very Likely (81-99%)
7: Certainly (100%)

FIGURE LEGEND

FIGURE 1
MAIN MODEL: THE EFFECTS OF DISCLOSURE

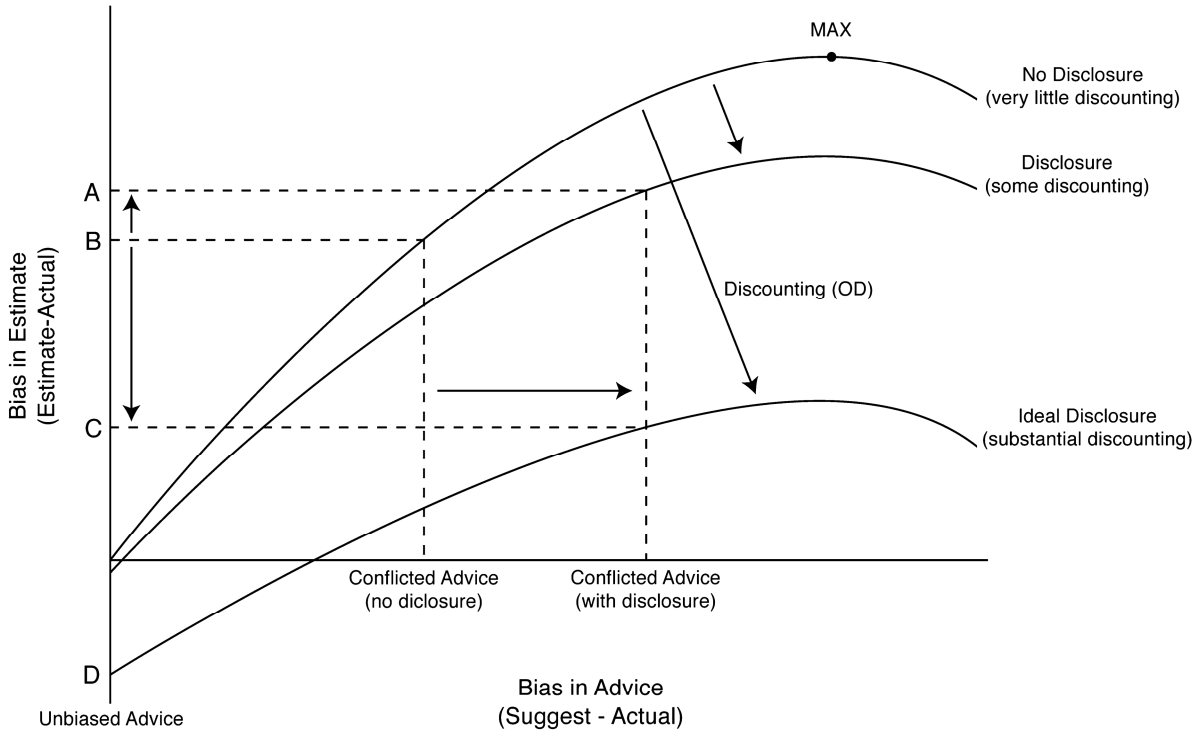


FIGURE 2
STRATEGIC RESTRAINT AND STRATEGIC EXAGGERATION

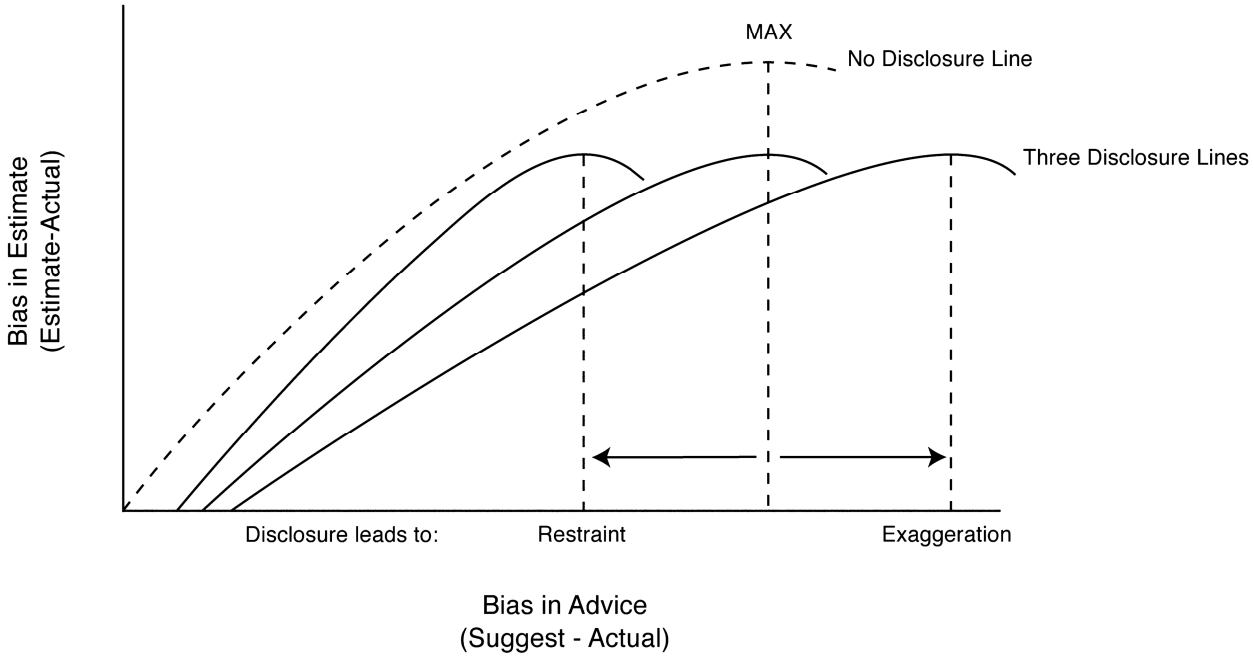


FIGURE 3
MORAL LICENSING

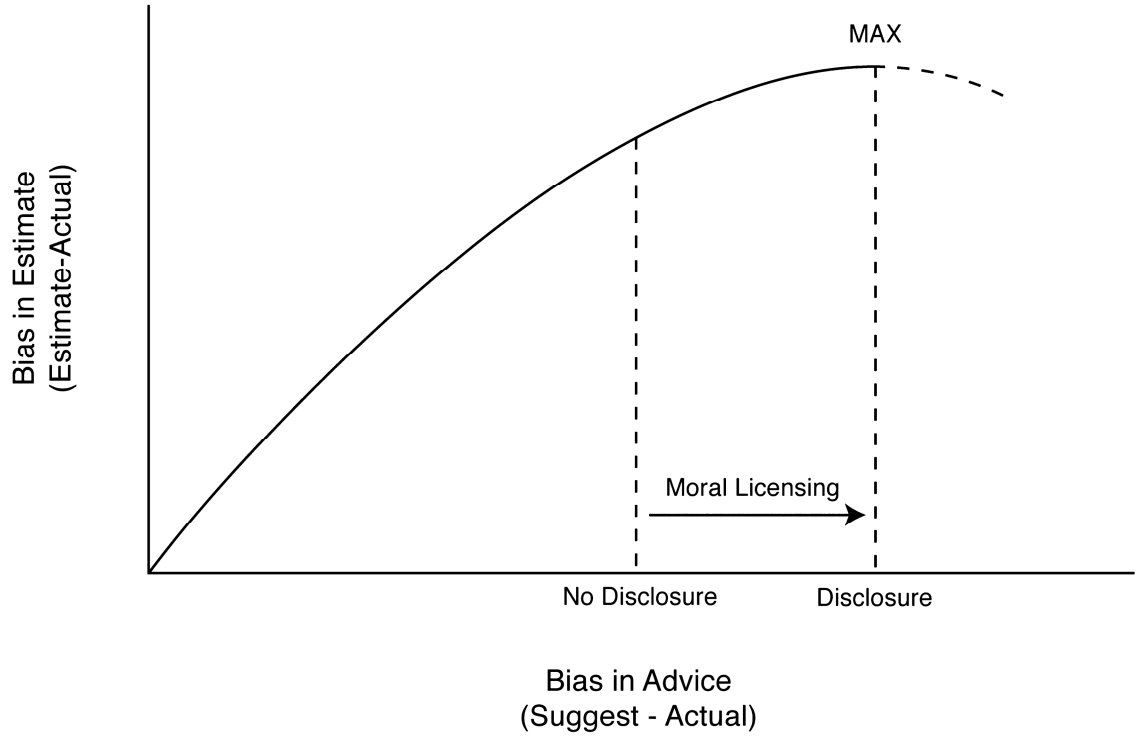
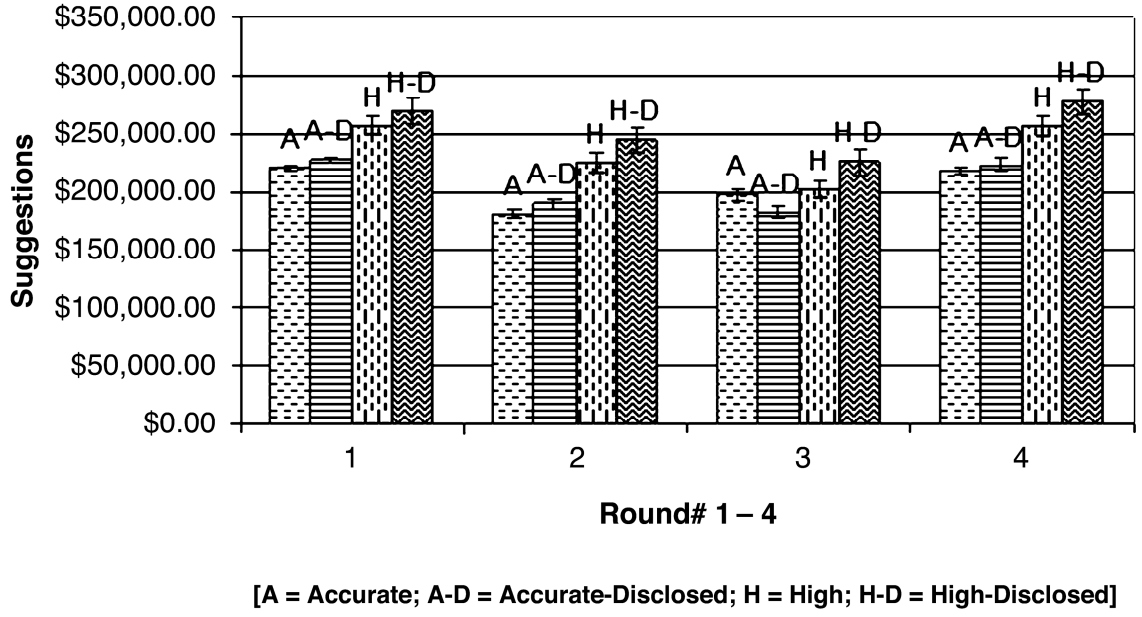


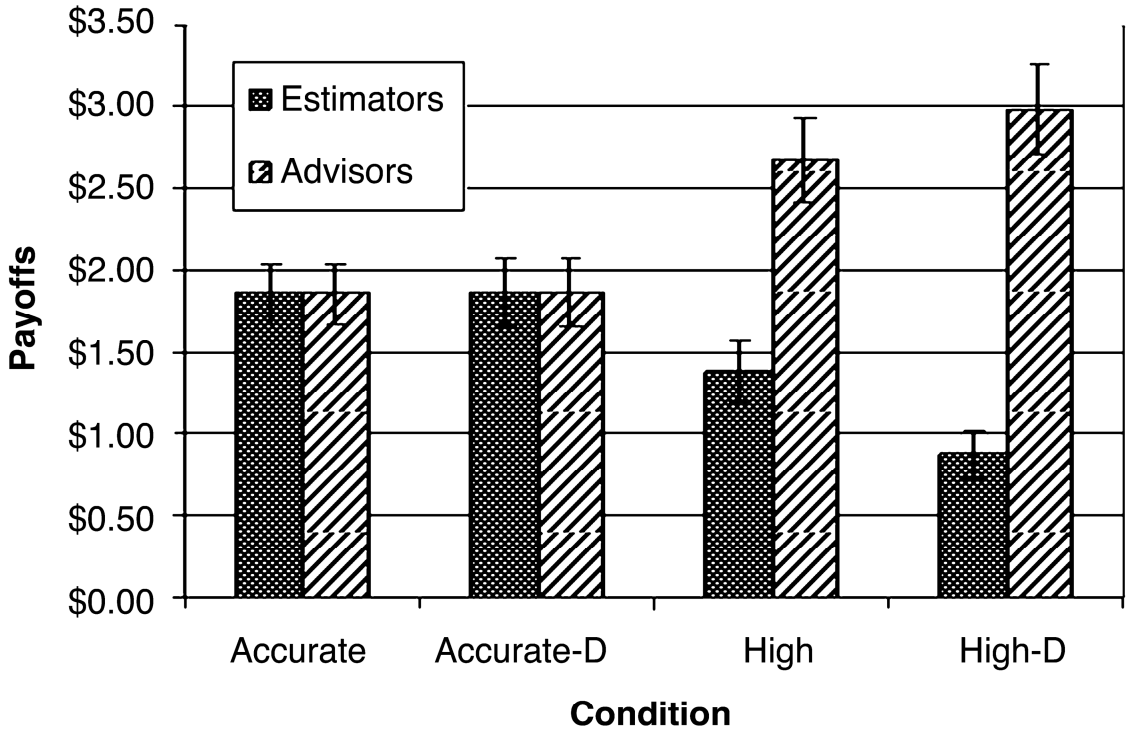
FIGURE 4
STUDY 3: ADVISORS' SUGGESTIONS ACROSS CONDITIONS AND ACROSS ROUNDS



(Error bars show standard errors.)

FIGURE 5

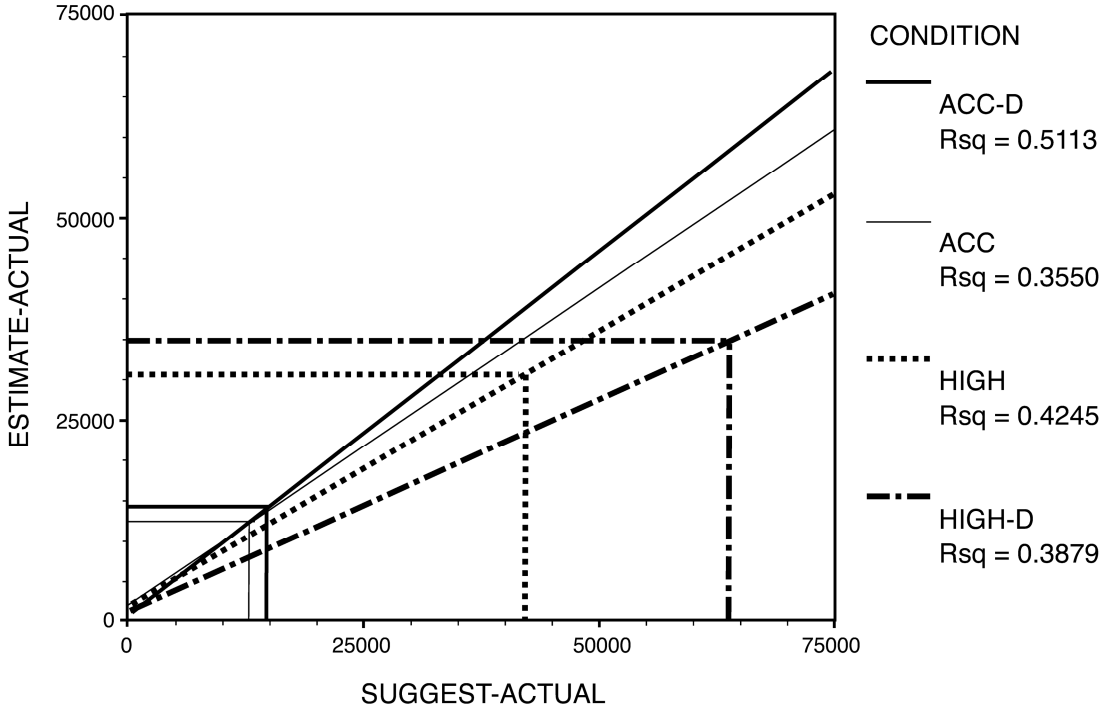
STUDY 3: PAYOFFS FOR THE TWO ROLES
IN THE FOUR EXPERIMENTAL CONDITIONS



(Error bars show standard errors.)

FIGURE 6

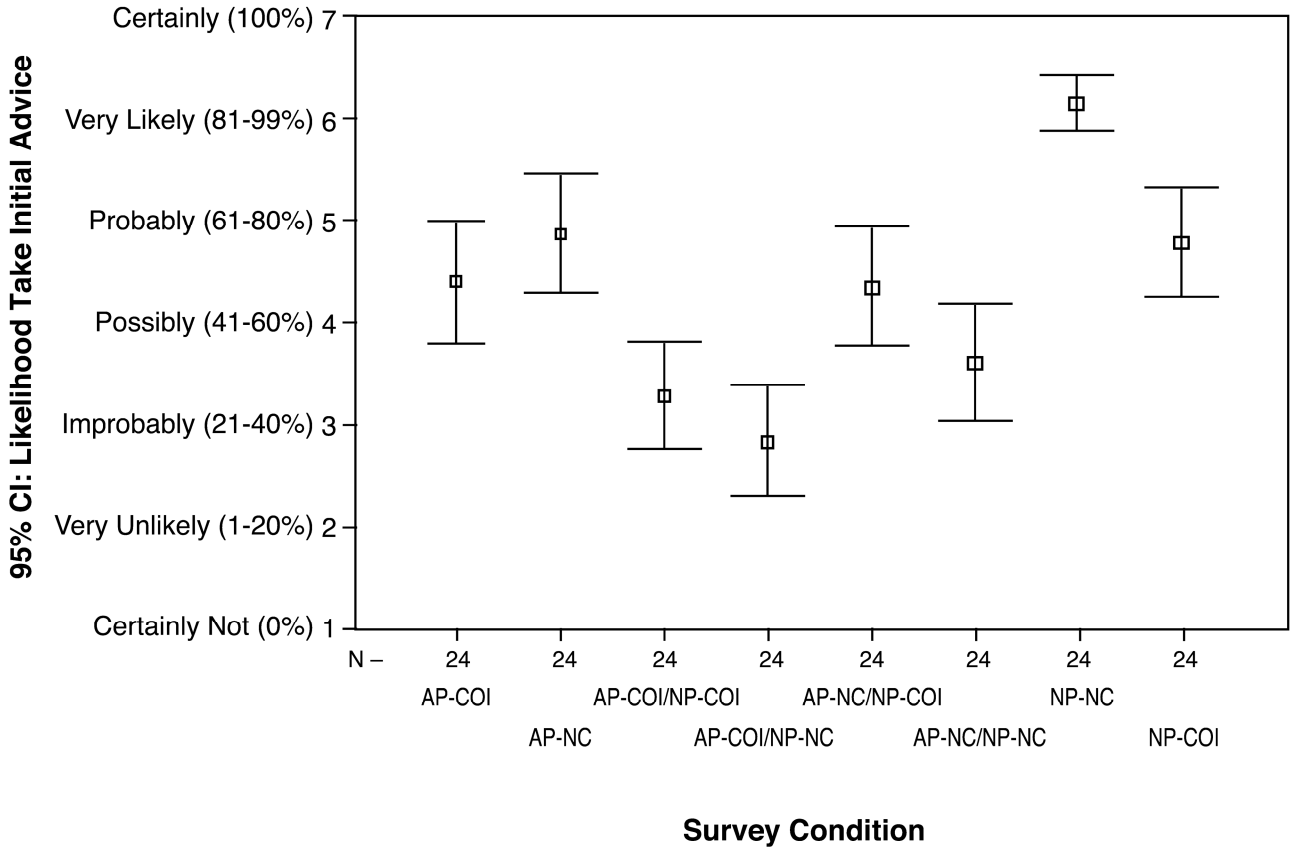
STUDY 3: OVERALL SUMMARY OF FINDINGS



(Horizontal lines represent the mean estimate-actual, or the bias in estimates, by condition; vertical lines represent the mean suggestion-actual, or the bias in suggestions, by condition.)

FIGURE 7

STUDY 4: SURVEY, OVERALL FINDINGS



AP = Abnormal Prescription (4 months). NP = Normal Prescription (1 month).
COI = Conflict. NC = No Conflict.

HEADINGS LIST