

Chapman University

Chapman University Digital Commons

ESI Working Papers

Economic Science Institute

11-13-2020

Keeping a Clean Reputation: More Evidence on the Perverse Effects of Disclosure

Cary Deck

Chapman University, deck@chapman.edu

J. Dustin Tracy

Chapman University, tracy@chapman.edu

Follow this and additional works at: https://digitalcommons.chapman.edu/esi_working_papers



Part of the [Econometrics Commons](#), [Economic Theory Commons](#), and the [Other Economics Commons](#)

Recommended Citation

Deck, C. & Tracy, J. D. (2020). Keeping a clean reputation: More evidence on the perverse effects of disclosure. *ESI Working Paper 20-21*. https://digitalcommons.chapman.edu/esi_working_papers/316/

This Article is brought to you for free and open access by the Economic Science Institute at Chapman University Digital Commons. It has been accepted for inclusion in ESI Working Papers by an authorized administrator of Chapman University Digital Commons. For more information, please contact laughtin@chapman.edu.

Keeping a Clean Reputation: More Evidence on the Perverse Effects of Disclosure

Comments

ESI Working Paper 20-21

Previously titled "Competition and Coming Clean".

Keeping a Clean Reputation: More Evidence on the Perverse Effects of Disclosure

Cary A. Deck and J. Dustin Tracy

November 13, 2020

Abstract

When a principal relies on the advice of an agent, a conflict of interest can encourage the agent to provide advice that is biased in a self-serving manner. Conventional wisdom suggests that such behavior can be minimized with disclosure requirements. However, disclosure has been shown to exacerbate self-serving bias and actually lead to greater harm for the principal in one-shot interactions. In many naturally occurring market settings, agents form reputations, a mechanism that diminishes the incentive to provide biased advice. In this paper, we test for bias in the advice agents provide when faced with reputation concerns and examine the impact of disclosure in such an environment. Based on controlled laboratory experiments, we find little evidence of self-serving bias in the absence of disclosure when agents form reputations as principals use that information in selecting agents. However, we find that introducing disclosure leads agents to provide self-serving biased advice.

Keywords: Conflict of Interest, Principal-Agent, Disclosure, Market Competition

JEL-Codes: *D82, G14, G18, I18, K12, L14, M55*

1 Introduction

Lawyers, financial advisors, doctors and other agents are said to have conflicts of interest if their personal motivation is at odds with their professional responsibilities to serve their clients (Boatright, 2000). While there are various policies in place regarding when disclosure of a conflict of interest is required, survey literature reveals such conflicts are common and undermine trust in professionals. For example, Scholl and Hung (2018) find there is generally low trust in financial professionals. They report that while 51% of consumers would prefer their adviser not have a conflict of interest, many consumers lacked awareness of a financial advisers' responsibilities or their compensation structure. Ninety-four percent of doctors have a financial relationship with a pharmaceutical company according to Campbell et al. (2007). They report that most (83%) doctors received gifts including meals from pharmaceutical companies and many received funding for continuing education or conference attendance (35%) or payment for speaking or research activity (28%). Interestingly, Mainous (1995) finds that patients viewed personal gifts to physician less favorably than office use gifts, believing the former could impact the cost of care. Further, patients who believe physicians accept gifts from the pharmaceutical industry trust doctors significantly less than those patients who do not believe physicians accept gifts (Grande et al., 2012). Unsurprisingly, physicians believe that gifts from the pharmaceutical industry are less influential and more acceptable compared to patient attitudes (Gibbons et al., 1998).

The propensity for experts to behave in a manner that conflicts with the interest of the principal they represent has been established in economics through both lab and field experiments (see Kerschbamer and Sutter (2017) for a review of this literature). The conventional wisdom is that if a principal is aware of an agent's conflict of interest then the principal can take appropriate counteraction. As stated by Cain et al.

(2005, p. 2), “Common sense suggests that recipients of advice will benefit from being more fully informed when they are made aware of an advisor’s conflict of interest.” This notion has led to widespread requirements for principals to disclose conflicts of interest, particularly in markets for “credence goods” where the agent is not able to determine the quality of service provided even after the fact.¹ While there is general support for the disclosure of conflicts of interest, there is some debate about the potential impact. For example, even though a majority of patients report wanting to know if their doctor received a gift over \$100 from a pharmaceutical company, a majority also say their trust in their doctor would be undermined if they knew their doctor had received such a gift (Green et al., 2012). Rose et al. (2019) find that disclosing physicians’ financial relations within the healthcare industry improved patient awareness of those relationships, but actually had no impact on the likelihood of attending an appointment or stated trust and confidence in physicians. In Massachusetts, a disclosure law reduced prescriptions for both branded and generic antidepressants, antipsychotics and statins (Guo and Sriram, 2017). However, the 2012 Federal sunshine law had little impact on *average* payments from pharmaceutical companies to prescribers and instead led to higher payments to fewer doctors (Guo et al., 2017). Similarly, Chen et al. (2019) find that in states with laws requiring posted public disclosure, payments to physicians who accepted less than \$100 decreased. However, payments to physicians who accepted more than \$100 increased.

In a creative experiment, Cain et al. (2005) show that disclosure can exacerbate the problem introduced by a conflict of interest rather than reducing it. In their study, agents provided estimates of the value of a jar of coins to principals. Principals were paid based on the accuracy of their own guess as to the value of the coins in the jar, but agents who had a conflict of interest were paid based on how large the principal’s estimate was. Cain et al. (2005) find that in the absence of disclosure, agents provide upwardly biased estimates. This bias leads principals to substantially overestimate the value of the coins. When the conflict of interest was disclosed to principals, the upward bias exhibited by agents did not decrease, but rather was dramatically *increased*. While the principals correctly anticipated agents would provide upwardly biased estimates, the principals badly underestimated the magnitude of the bias and thus the net effect of disclosure was actually harmful to the principals. Cain et al. (2011) report four additional experiments examining the robustness of Cain et al. (2005) and find the result is not unique to jars of coins and that it persists even with between round feedback. Cain et al. (2011) also show that agents perceive bias behavior as being less morally unacceptable with disclosure.²

While the counter-intuitive finding of Cain et al. (2005) and Cain et al. (2011) suggest that it may be optimal to not disclose conflicts of interest, the considered settings are ones in which there are no other pressures on the agents. Becker (1968) argues that when dishonest behavior has little to no cost it is expected to be prevalent; so, the observed bias in Cain et al. (2005) in the absence of disclosure is unsurprising. Church and Kuang (2009) show that principals being able to directly sanction agents, then outcomes are improved. However, the sanctions introduced by Church and Kuang (2009) are meant to model repercussions like malpractice claims against doctors and thus are very severe. In many circumstances where principals rely upon expert advice the opportunity for such dramatic sanctions may not be possible. For example, a doctor may prescribe a medication produced by a company whose representative regularly gives the doctor gifts over a different medication that is more likely to be beneficial to the patient. But if the prescribed medication has an FDA indication for the diagnosed ailment, then a lawsuit against the doctor is unlikely to be successful. However, many everyday circumstances involve reputation concerns and the fear of lost future opportunities make dishonesty costly for agents.

The expected effect of maintaining a reputation is straightforward in cases where the principal can easily assess the performance of an agent *ex post*, but for credence goods, such as selecting a doctor, financial

¹In some sense, the good used in our experiment can be viewed as an “Experience Good” since the principals eventually learn the true value of that good. However, the distinction of whether experience leads to knowledge of the true value of a good is not the key to assessing if one made the correct decision in our setting. Instead, what is relevant in our setting is the relationship between the value of the selected good and the value of the non-selected good. As our principals only receive a noisy signal of the value of the item not selected, it is difficult to ascertain if one made the best choice and thus our setting has the fundamental characteristic associated with a credence good.

²Cain et al. (2011) find that principals do discount agents with conflicts of interest when there is a second agent providing advice that is known to not have a conflict.

adviser, or lawyer, the effect is less obvious since assessing the agent is difficult. Still, Darby and Karni (1973) discuss how market forces and reputation can limit fraud for credence goods and Wolinsky (1993) presents a theoretical proof of how reputation and consumer search can discipline experts. However, the empirical support for the effectiveness of reputations is less clear. For example, Pope (2009) finds patients select hospitals based on rankings, but not underlying quality measures of care. Hannan et al. (1994) argue providing rankings improves quality and decreases mortality rates for coronary artery bypass graft surgery. But, Mukamel and Mushlin (1998) find the rankings increased market share for low mortality surgeons and hospitals leading to increased prices for the more highly demanded providers. Dranove et al. (2003) claim that rankings reduce consumer welfare, at least in the short term, as doctors attempting to improve their reputations elect to operate on healthier patients, a pattern that leads to increased costs, which results in no improvement in the outcomes of the healthier patients, but worse outcomes for less healthy patients. Outside of healthcare, Carl (2008) finds that disclosure of payments *increases* an agent’s credibility when engaged in word-of-mouth product promotion, a potential channel through which disclosure could be harmful. It is also possible that disclosure actually creates confusion, thus generating another path by which disclosure could be harmful. For home mortgages, Ben-shahar and Schneider (2011) argue that people may not know how to use disclosed information and that they can only pay attention to so many things so that it is better for the principals if they are focused on a limited number of key features of the decision.

If reputation is sufficient to discipline agents, it is unclear if disclosure is socially desirable as the upside benefit is limited and disclosure could actually introduce self-serving bias of assessments. Of course, if agents exhibited biased assessments even in the presence of reputation concerns, then it is possible that disclosure could be beneficial, as typically assumed, or harmful, as suggested by some. In this paper we use a controlled laboratory experiment inspired by Cain et al. (2005) to examine the impact of disclosure on self-serving bias in assessments by agents in a credence good setting where agents have reputation concerns. In our task, principals are forced to select between two options of unknown value (jars with indeterminate compositions of coins), but have to rely on an agent’s assessment of the value of those options. The principal receives a benefit based on the value of the selected option and a noisy signal of the value of the non-selected option, both of which are independent of their agent’s behavior. Principals can rate agents after their interaction and can use the ratings provided by other principals when selecting an agent. Agents have a financial incentive to encourage a principal to select a particular option (jar of coins), but the agent also has an incentive for principals to choose to solicit the agent’s assessments. There are two treatments in the experiment: one in which the agent’s conflict of interest is not disclosed to the principal and one in which it is disclosed. We measure bias as a systematic overstatement by the agent of the value of the conflicted option in comparison to the other option. Thus, our use of the term bias coincides with both the statistical definition and the colloquial one. As a prelude to the results, when there is no disclosure we do not observe systematically biased assessments of values. But, like Cain et al. (2005), when disclosure is introduced we find that agents do provide systematically biased assessments of value and principals are unable to fully incorporate this fact into their decision making.

2 Experimental Design and Hypotheses

In this section we lay out the experimental design including the task and procedures. We also present a simple theoretical model that serves as the basis for our hypothesis.

2.1 Main Experimental Task and Treatments

Our experiment varies whether or not the agents’ conflict of interest is disclosed in an environment in which an agent’s future earning can be diminished by an unfavorable reputation. In the main experimental task, a principal (“investor” in the subject interface) must select between two options labeled A and B. The options are jars of coins and the principal’s value for an option equals the value of the coins in the associated jar. However, the principal cannot observe the jars directly and instead is forced to rely upon an assessment of the value of each option provided by an agent (“expert” in the subject interface). The agent has a direct incentive

to encourage the principal to select option A, but also has an incentive to maintain a good reputation in order to be selected by other principals in the future as the game is played over multiple rounds. The rationale for using jars of coins is that the agent always has plausible deniability, even to the researcher, that the reported estimates of value match the agent’s actual beliefs.³

A session of the experiment involved 12 subjects. Half of the subjects were randomly assigned the role of principal and half were assigned the role of agent. Once assigned, subjects retained their roles throughout the entire experiment, which lasted for multiple blocks of distinct rounds.

Table 1 lists the sequence of events within each round of the experiment. In Stage 0, subjects are randomly assigned to a group consisting of two agents and two principals. Additionally, the value of options A and B were assigned. In Stage 1, the agents observed two images of jars of coins (labeled “Jar A” and “Jar B”).⁴ Figure 1 provides a sample of one of the images shown to agents in the experiment. After observing the images, each agent would provide an estimate for the value of the coins in each jar. In Stage 2, principals see the mean number of stars each agent in the group has received over the course of the previous rounds in the current block. In the Disclosure Treatment, agents were also informed that the expert would earn an extra \$0.50 if the principal chooses Jar A. In the Baseline, no such disclosure was made although the agent’s still had the same conflict of interest. The presence or absence of the disclosure statement is the only difference between the two treatments.⁵ The principals then selected an agent, E . An agent earned \$0.75 for each principal that selected the agent. In Stage 3, principals observed \hat{v}_A^E and \hat{v}_B^E which are the estimated values for Jar A and Jar B provided by the agent selected in Stage 2. Principals then select a jar. The principal’s agent earns an additional \$0.50 if the principal selects Jar A. In Stage 4, the value of the selected option, v_J^E , was revealed to the principal. Each principal also received a noisy signal of the true value of the jar that was not selected. The noisy signal was equal to $v_J^E + e_{it}$ where $e_{it} \sim \mathcal{U}[-\$2.00, \$2.00]$. Principals then rated their agent on a scale of 1 to 5 stars with 1 star being the worst and 5 stars being the best. Each agent was informed of their own payoff, the number of principals who selected the agent, and the number of principals who selected the agent and subsequently selected Jar A. Agents were never informed of the true value of Jars A and B nor were they informed of their own reputation.⁶

Stage 5 determines if a block continues for an additional round or not. For Rounds 1 through 4 of each block, once Stage 4 was completed the process returned to Stage 0 and repeated. However, starting in the 5th round of a block, at Stage 5 the process returned to Stage 0 with a 75% chance and the block ended with a 25% chance. Thus, the horizon of the block was indefinite although subjects knew the stochastic nature of the termination. Importantly, reputations did not carry over between blocks. Thus, a new block offered an agent a chance to start over. This was done out of concern that an agent who formed a bad reputation early might never have the chance to improve it in later rounds since only selected agents are rated. A single set of realizations was drawn and used to determine the duration of each block for all sessions in both treatments. Ultimately, there were four blocks lasting a combined 29 rounds. Agents were paid their cumulative earnings for two randomly chosen blocks whereas principals received the value of the jar selected in one of the 29 rounds chosen at random.

2.2 Additional Experimental Details

The jars that were used contained mixtures of quarters, dimes, nickels and pennies. The actual values of the contents ranged from a low of \$13.32 to a high of \$22.43 with a mean value of \$17.77. Jars contained between 175 and 466 coins each. In the instructions it was demonstrated that volume was not a good indicator of value of coins. The intention was that each image would be used as both Option A and Option B during the course of the experiment although not in the same round to control for jar-specific characteristics. However,

³This process is similar to Cain et al. (2005) who used jars of coins for the same reason.

⁴All subjects were informed in the instructions that every jar would contain between \$10 and \$25.

⁵In the first round of a block there were no previous ratings to display. For this reason, in the first round of a block each principal was assigned to a unique agent. This process ensured that an average rating could be calculated for all principals in all subsequent rounds of a block.

⁶These two design choices are meant to mimic situations like health care where the doctor cannot know the true value of the outcome to the patient and where reputation is often informal.

Table 1: Sequence of Events within a Round

Stage	Active Role	Event	Payoffs	
			Agent	Principal
0 Set Up	Nature	Form Groups of 2 Principals and 2 Agents; Assign Jars A & B for Agents $I \in \{1, 2\}$ values, $\{v_A^1, v_B^1, v_A^2, v_B^2\}^a \in (\$10, \$25)$		
1 Estimate	Agents	Observe images of Jars A & B; Estimate the value of each jar $\{\hat{v}_A^1, \hat{v}_B^1, \hat{v}_A^2, \hat{v}_B^2\}$		
2 Select Agent	Principals	[NOT in Baseline: See disclosure statements] Observe Agents' $\bar{r}^I(\star)$ (except round 1) ^b Select Agent $E = \{1, 2\}$ (except round 1) ^b	+75¢(for E) ^c	
3 Select Jar	Principals	Observe selected Agent's estimates, \hat{v}_A^E and \hat{v}_B^E ; Select Jar $J \in \{A, B\}$	If $J=A$ +50¢(for E) ^c	
4 Feedback	Principals	Receive signals, v_J^E and $v_{-J}^E + e_{it}$, ^d Rate selected Agent $r_t^E \in \{1, 2, 3, 4, 5\}(\star)$		+ v_J^E
	Agents	See aggregate selections and own payouts		
5 Continue	Nature	Determine if block continues; ^e if yes, $\bar{r}^I(\star)$ updated and go to Stage 0; if no, $\bar{r}^I(\star)$ reset and block ends		

^a $v_A^1 = v_B^2$ and $v_A^2 = v_B^1$.
^b In Round 1, there were no ratings to show, and one Agent is assigned to each Principal ensuring ratings in future rounds.
^c The Agent who is not selected earns no payment.
^d $e_{it} \sim \mathcal{U}(-\$2, +\$2)$.
^e Starting in the 5th round, the block continued for another round with a 75% probability.

due to a programming error, some images were treated as Option A twice or as Option B twice for some subjects. Therefore, the analysis presented in the next section is limited to cases where an image was used both for Option A and Option B over the course of the experiment.⁷

The experiment was implemented via oTree (Chen et al., 2016) and conducted at the Economic Science Institute at Chapman University. A copy of the instructions is included in Appendix A.1. Subjects were recruited from the lab's standing pool of volunteers and none of the subjects had participated in any related study. Each session lasted up to 90 minutes and subjects earned an average of \$22.90 including a \$7 participation payment. For both the Baseline and the Disclosure Treatment, three sessions were run involving a total of 72 subjects.

Guessing the correct value of the coins is challenging. As a demonstration, we encourage the reader to guess the value of the coins shown in Figure 1 before reading the footnote for this sentence.⁸ Therefore, a separate study was conducted to calibrate error in estimating the value of each jar of coins. We refer to this separate study as the Guess Experiment. Twenty four subjects participated in the Guess Experiment in a single session at the Economic Science Institute. This session lasted thirty minutes and relied on volunteers from the same subject pool as the main experiment although no one was allowed to participate in both the Guess Experiment and the main experiment. Subjects in the Guess Experiment observed the full set of images used in the main experiment and a quadratic scoring rule was used to incentivize truthful reporting of the value of the coins in each image.⁹ The average earnings in the Guess Experiment were \$14.58 including the \$7 participation payment.

⁷Results that include all estimates are not qualitatively different. See Appendix A.2.

⁸The jar shown in Figure 1 contains \$17.73.

⁹Subjects earned \$25 - \$2(guess - true value)² each round, and were paid their average earning over all rounds.

Figure 1: Sample Image



2.3 Theoretical Model

Levitt and List (2007) present a model in which the utility of agent i for an action has additively separate terms for the moral component, M_i , and the financial component, W_i . Specifically, they posit

$$U_i(a, v, n, s) = M_i(a, v, n, s) + W_i(a, v)$$

where a is the action, v is amount of money at stake, n capture societal norms or laws, and s is the scrutiny the action will receive. For an immoral action the first term will be negative. However, if the second term is greater than the first, the immoral action will bring agent i a net increase in utility and the agent will take the action.

In this model, an increase in scrutiny for a given norm increases the moral cost but it has no impact on the financial benefits (i.e., $\frac{\partial M_i}{\partial s} < 0$ and $\frac{\partial W_i}{\partial s} = 0$). Thus, disclosure to the principal that agent i earns a bonus if the principal chooses one option but not the other should increase the moral scrutiny of the estimates i reports, which would discourage immoral action by the agent. With the introduction of a reputation the financial term could also depend on both n and s as well (i.e., $v = f(n, s, \dots)$). In such a setting an immoral action leads to both greater moral costs and reduced financial benefit (i.e., $\frac{\partial M_i}{\partial s} < 0$ and $\frac{\partial W_i}{\partial v} \frac{\partial v}{\partial s} < 0$). This would make disclosure even more attractive for principals.

However Cain et al. (2005) speculate that disclosure may actually create a “moral license” to engage in otherwise socially unacceptable behavior. Effectively, this line of reasoning posits that since the principals have been warned, they should expect agents to engage in self-serving behavior and account for this when making their own decisions. This line of reasoning is similar to the old adage of *caveat emptor*. In terms of the model by Levitt and List, moral licensing alters the norms, n , such that the immorality of an action is reduced (i.e., $M_i(a, v, n_{disclosure}, s) > M_i(a, v, n_{nondisclosure}, s)$). If true, the net effect of such moral licensing is ambiguous and disclosure could lead to more or less self serving behavior. This ambiguity is true even when taking reputation into account as the financial harm from an immoral action could be reduced as the norm changes.

Based on this model, our experiment seeks to answer the two following questions which serve as the main

hypotheses to be tested. First, will agents exhibit a self-serving bias in the presence of reputation concerns? Second, will the disclosure of the agent’s conflict of interest lead to more self-serving bias on the part of agents?

3 Results

The results are presented in three subsections. The first subsection reports the estimates of jar values from the Guess Experiment as these serve as a reference point for estimates reported in the main experiment. The second subsection focuses on the behavior of agents in the Baseline and in the Disclosure Treatment, which is the primary focus of the paper and where our two main hypotheses are tested. The third subsection reports exploratory analysis of the behavior of principals in the Baseline and in the Disclosure Treatment.

3.1 Guess Experiment

The mean guess of jar value was \$14.61, while the mean of the true values was \$17.78. That is, the average error was -\$3.17. Figure 2 shows a boxplot of the estimates from the Guess Experiment for each jar, ordered by the actual value of the coins in the jar (and not the order in which subjects in either treatment observed the jars). The ♦s show the actual value of a given jar while the thick black line gives the median guess and the gray box shows the interquartile range of guesses. In only 3 out of 30 jars did the median guess not fall below the true value. In fact, for a sizable majority of the jars the true value exceed the 75th percentile of guesses.¹⁰ A sign test rejects that errors are unbiased (p -value < .001). This leads to Finding 1.

Finding 1: People systematically underestimate the value of jars.

Because people so badly underestimate the true value of the jars, it is possible that someone who was deliberately inflating their estimates could still systematically report estimates below the true value. For this reason, we use the estimates from the Guess Experiment as a basis of comparison for identifying agent bias in stated estimates of value in the main experiment. While it is clear from Figure 2 that there is considerable variation across individuals with respect to their estimates of the value of a given jar, if agents in the main experiment are not systemically manipulating their estimates then on average estimates should be equal across the two experiments.

3.2 Behavior of Agents

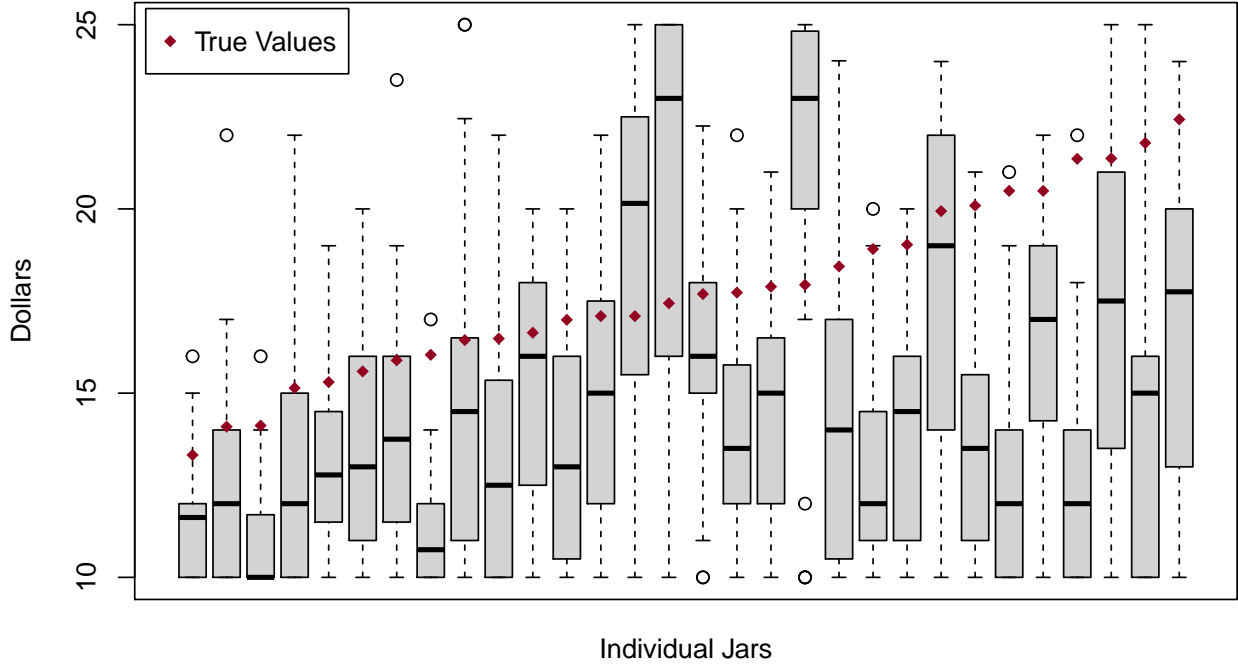
Table 2 reports the estimates made by agents for all jars by treatment, and by whether there was a bonus to the agent if the principal chose that jar (Jar A), or if there was no bonus associated with the jar (Jar B). The average actual value of \$17.81 in Table 2 differs from the \$17.78 figure reported in the previous subsection, because most although not all jars were observed twice by agents during the main experiment and the values in Table 2 account for the frequency with which the jar was observed.

Table 2: Means by Treatment

	Baseline	Disclosure	Actual Value
Estimate Jar A	16.15	16.28	17.81
Estimate Jar B	16.36	15.88	17.81

¹⁰Surowiecki (2005) discusses the “wisdom of crowds” arguing the average value of independent assessments tends to be an accurate forecast of an unknown value (e.g., average guess of the weight of an ox or the number of jellybeans in a jar tends to be close to the respective true values. Our subjects do not exhibit a “wisdom of crowds” for the value of coins in a jar.

Figure 2: Boxplots of Incentivized Estimates and Actual Values



Note: Subject guesses for each individual jar are displayed as separate boxplots. Plots are ordered by actual value of the jar. (Subjects did not see jars in value order.)

Table 3 reports similar information to Table 2, but it is presented as deviations from the average estimates in the Guess Experiment.¹¹ Table 3 also reports the values from (two-sided) paired t-tests and Wilcoxon rank-sum tests comparing agent’s estimates when the jar is and is not bonus eligible (i.e., comparing estimates for a given image when that image is used as Jar A and Jar B, respectively). In the Baseline, the reported values for a given jar do not depend on whether the jar is Jar A or Jar B as the difference in estimates is not statistically different from zero. This result indicates that agents are not systematically favoring the bonus eligible jar in the Baseline. That is, when the conflict of interest is not disclosed there is no observed bias on the part of agents in the presence of reputation concerns. However, in the Disclosure Treatment, the difference is statistically significant with agents reporting higher estimates for a jar when it is the bonus earning Jar A than when it is the non-bonus earning Jar B. This indicates that disclosure of the conflict of interest introduces bias into agent estimation despite the presence of reputation concerns.

To further analyze agent behavior we rely upon the following regression equation:

$$\hat{v}_{ijt} = \beta_0 + \beta_1 Value_{jt} + \beta_2 Baseline \& Jar A_j + \beta_3 Disclosure \& Jar A_j + \beta_4 Disclosure \& Jar B_j + \varepsilon_{ijt}$$

where \hat{v}_{ijt} is agent i ’s reported estimate of the value of jar j in round t . $Value_{jt}$ is a variable that captures the value of the coins in the jar for which the agent is providing an estimate. For robustness, we consider both the case where Value is the actual value of the coins in the jar (using the variable *ActualValue*) and the case where Value is the mean guess of the value of the coins in the jar from the Guess Experiment (using the variable *MeanGuess*). The main independent variables of interest are the interaction variables: *Baseline &*

¹¹Attention is restricted to cases in which the agent observed a given jar as both Jar A with a bonus payment and as Jar B without the bonus. As discussed in the previous section, a programming error meant that not all agents observed all images as both Jar A and Jar B exactly once during the experiment.

Table 3: Deviations from Guess Treatment Estimates for Cases in which each Jar was both A and B

	Baseline	Disclosure
Estimate Jar A	1.77	2.14
Estimate Jar B	1.89	1.63
Difference	-0.11	0.52
t-test p-value	0.51	0.01
Numb. Jar Pairings	300	300
Wilcoxon-Test p-value	0.49	0.01

Jar A, *Disclosure & Jar A*, and *Disclosure & Jar B*, which are all binary variables that take the value 1 if the estimate is for the stated jar in the stated condition and take the value 0 otherwise. The omitted group is *Baseline & Jar B*, which serves as a reference point.

The regression estimation for agent reports is reported in Table 4. Specifications (1) and (3) use *ActualValue* while specifications (2) and (4) use *MeanGuess*. Specifications (1) and (2) do not include controls for the specific jars or round within a block nor do they cluster errors by agent, while specifications (3) and (4) do. We report all four specifications for the sake of robustness. The coefficient estimate for *Baseline & Jar A* is not statistically significant in any specification. In fact, it is nominally negative, indicating the mean Jar A reported estimates are, if anything, less than the mean Jar B reported estimates. This pattern provides further evidence that in the absence of disclosure agents do not provide self-serving biased estimates systematically favor Jar A. However, as shown in the lower panel, which reports the estimated difference between β_3 and β_4 , in the Disclosure Treatment agents report that a jar is worth 51.7¢ more when it is Jar A than when it is Jar B. In all four specifications, this difference is positive and at least marginally statistically significant. Thus, when the conflict of interest is disclosed to principals then agents do provide self-serving estimates of jar value. Together, the evidence in Table 3 and Table 4 support Finding 2.

Finding 2A: In the absence of disclosure, agents do not exhibit bias when there are reputation concerns.

Finding 2B: In the presence of reputation concerns, disclosure of an agent’s conflict of interest leads to biased and self-serving behavior.

In terms of the model by Levitt and List (2007) discussed in the previous section, the observed behavior indicates that agents believe disclosure creates a moral license to be self-serving and will reduce the scrutiny given to agent’s selfish action.

Table 4: Regression Estimates for Agents' Estimates of Jars

VARIABLES	(1)	(2)	(3)	(4)
Constant	9.576*** (0.786)	2.965*** (0.687)		
Value	0.349*** (0.036)		0.905*** (0.022)	
MeanGuess		0.913*** (0.038)		1.207*** (0.030)
Baseline & Jar A	-0.112 (0.253)	-0.112 (0.215)	-0.112 (0.191)	-0.112 (0.191)
Disclosure & Jar A	0.011 (0.662)	0.005 (0.622)	0.005 (0.660)	0.005 (0.660)
Disclosure & Jar B	-0.506 (0.662)	-0.513 (0.622)	-0.513 (0.534)	-0.513 (0.534)
Observations	1,200	1,200	1,200	1,200
Number of Unique IDs	36	36	36	36
Overall R^2	0.068	0.285	0.343	0.343
Jar # Controls	N	N	Y	Y
Round # Controls	N	N	Y	Y
Clustering	N	N	Y	Y
Jar A - Jar B for Disclosure	0.517	0.517	0.517	0.517
SE	0.253	0.215	0.313	0.313
$P > t $	0.041	0.016	0.099	0.099
Standard errors in parentheses		*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$		

3.3 Behavior of Principals

Principals make three sequential decisions each round: selecting an agent, choosing a jar based on the selected agent's estimates, and rating the agent. In this subsection we analyze these three activities in turn.

Selecting an Agent

To examine how reputation impacts a principal's selection of an agent, we estimate the following probit model:

$$Agent2_{it} = \Phi(\beta_0 + \beta_1 StarsDiff_{it} + \beta_2 Disclosure_i * StarsDiff_{it} + \varepsilon_{ijt})$$

where $Agent2_{it}$ takes the value 1 if principal i chose Agent 2 in round t and the value 0 if Agent 1 was chosen. The independent variable, $StarsDiff$, is the difference in the number of stars each agent has (i.e., Agent 2's rating - Agent 1's rating) for the two agents in the principal's group in the given round. $StarsDiff$ is interacted with a binary variable for the Disclosure Treatment to allow for the rating differential to matter differentially to principals across treatments. The regression clusters standard errors at the principal level and includes a random effect for each principal.

Table 5 reports marginal effects from the estimation. Specifications (1) and (2) differ in that specification (2) includes controls for the specific round within the block while specification (1) does not. The estimation shows that each additional star in an agent's rating over and above the rating of the other agent increases the likelihood of the higher rated being selected by about 32%. Interestingly, the coefficient estimate for the interaction term is indistinguishable from zero, meaning that principals do not evaluate reputations

differentially across treatments. That principals are more likely to select the agent with the higher rating indicates that reputation does matter in this setting and that agents are under pressure to maintain a high rating relative to the other agents. This is stated as Finding 3.

Finding 3: Principals rely on the rating system when selecting an agent and thus agents face reputation concerns in this setting, although the importance of reputation does not vary with disclosure of the agent’s conflict of interest.

Table 5: Estimated Marginal Effects of Likelihood of Choosing Agent 2

VARIABLES	(1)	(2)
Stars Agent 2 - Stars Agent 1	0.318*** (0.044)	0.318*** (0.043)
Disclosure * (Stars Agent 2 - Stars Agent)	-0.000 (0.087)	0.000 (0.086)
Observations	900	900
Round # Controls	N	Y
Log Likelihood	-349.2	-346.2
Standard errors in parentheses	*** p<0.01, ** p<0.05, * p<0.1	

Choosing a Jar

Similar to the approach we use for identifying which agent a principal selects, we rely on a probit model to determine what influences the jar a principal selects. Specifically, we rely upon the following specification:

$$JarB_{it} = \Phi(\beta_0 + \beta_1 EstimateDiff + \beta_2 Disclosure_i + \beta_3 Disclosure_i * EstimateDiff_{it} + \varepsilon_{it})$$

where $JarB_{it}$ indicates principal i chose Jar B in round t . The independent variable of interest is $EstimateDiff$, which is the difference in the selected agent’s estimate of Jar A minus the selected agent’s estimate of Jar B. The regression includes a random effect for each principal and clusters standard errors by principal. Table 6 reports estimated marginal effects from the probit regression. In specification (1) it is assumed that there is no difference in jar selection between treatments while specification (2) allows for the difference in jar selection to vary between treatments. The $Disclosure$ variable tests to see if the variation is significant. Specification (3) allows for a shift in jar selection due to the treatment as well as in interaction term between the treatment and $EstimateDiff$.

The interpretation of the marginal effect of $EstimateDiff$ is that for every dollar that the estimate of Jar A exceeds Jar B there is a 8% decrease in the chance that Jar B is picked. The negative sign on this coefficient (in all specifications) is what one would expect if principals believe agent estimations convey useful information. From specifications (2) and (3), principals in the Disclosure Treatment are actually *less* likely to choose Jar B by 10% as compared to the Baseline, *ceteris paribus*. This effect is the opposite of what one would observe if principals were wary of choosing Jar A in the Disclosure Treatment, but it is consistent with principals not fully appreciating the implications of the agent’s conflict of interest and having a desire to help the agents earn a larger profit. Specification (3) reveals that the impact of differences in the reported estimated values of the jars does not vary with treatment. This leads to Finding 4.

Finding 4: Subjects directly rely on the agents’ estimates of jar value to determine which option to select in both treatments. However, Disclosure actually increases the likelihood that the principal will select the option that benefits the agent *ceteris paribus*.

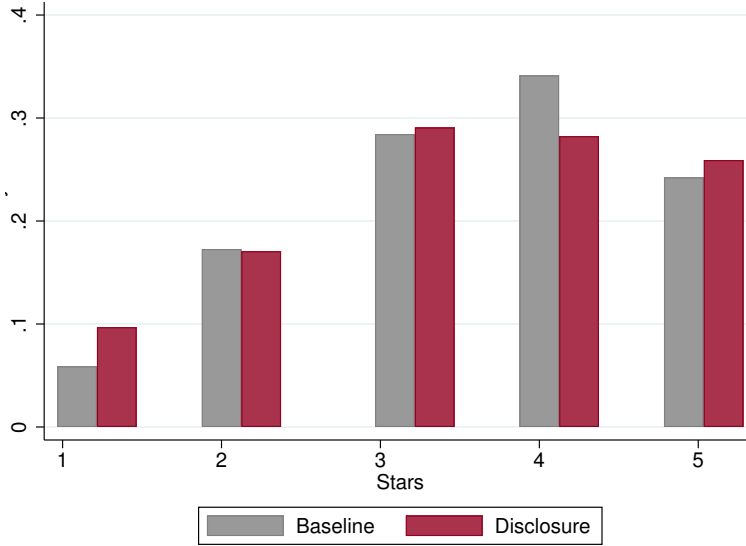
Table 6: Estimated Marginal Effects of Likelihood of Choosing Jar B

VARIABLES	(1)	(2)	(3)
Estimate Jar A - Estimate Jar B	-0.078*** (0.006)	-0.078*** (0.006)	-0.085*** (0.010)
Disclosure		-0.106** (0.041)	-0.101** (0.040)
Disclosure * (Est. Jar A - Est. Jar B)			0.016 (0.020)
Observations	1,044	1,044	1,044
Base Rate	0.384	0.463	0.462
Log Likelihood	-496.7	-493.6	-492.1
Standard errors in parentheses	*** p<0.01, ** p<0.05, * p<0.1		

Rating the Agent

Figure 3 plots how often each star rating was assigned by treatment. The mean number of stars in the Disclosure Treatment (3.397) was slightly lower than the mean in the Baseline (3.487), but the difference is not statistically significant (p -value = 0.226).

Figure 3: Histogram of Stars (Ratings) of Agents



To understand how agent performance impacts the rating a principal gives an agent after observing the true value of the selected jar and the independent noisy signal of the value of the other jar, we rely on the two following specifications. The first specification classifies jars based on whether the agent had a conflict of interest with the jar or not (i.e., Jar A versus Jar B).

$$\begin{aligned}
 Stars_{it} = & \beta_0 + \beta_1 AccurateDirection_{it} + \beta_2 OverEstJarA_{it} + \beta_3 UnderEstJarA_{it} \\
 & + \beta_4 OverEstJarB_{it} + \beta_5 UnderEstJarB_{it} + \varepsilon_{it}
 \end{aligned} \tag{1}$$

where $Stars_{it}$ is the number of stars principal i assigned the selected agent in round t . $AccurateDirection$ is a binary variable indicating that the jar the agent reported as having the highest value also had the highest value based upon the information subsequently revealed to the principal. That is, $AccurateDirection$ takes the value 1 if the agent provided guidance to the principal that was in the “Accurate Direction” and is 0 otherwise. The remaining variables pertain to the difference between the agent’s estimate of a jar’s value and the subsequent information received by the principal. $OverEstJarX$ is the amount the principal believes the agent overestimated the value of Jar X while $UnderEstJarX$ is the amount the principal believes the agent underestimated the value of Jar X.

The second specification classifies jars based on whether the principal selected the jar or not.

$$Stars_{it} = \beta_0 + \beta_1 AccurateDirection_{it} + \beta_2 OverEstPick_{it} + \beta_3 UnderEstPick_{it} + \beta_4 OverEstOther_{it} + \beta_5 UnderEstOther_{it} + \varepsilon_{it} \quad (2)$$

where, similarly to Equation 1, $OverEstPick$ ($UnderEstPick$, $OverEstOther$, $UnderEstOther$) denotes the amount by which the principal believes the agent over- (under-, over-, under-) estimated the value of the selected (selected, not selected, not selected) jar, respectively. This approach allows for the possibility that principals view errors asymmetrically. For both equations 1 and 2 we report 3 estimations. The first is just the basic specification given above. The second specification includes $Disclosure$, a binary variable for the Disclosure Treatment, as well as the interaction of $Disclosure$ and $AccurateDirection$. The third specification adds interaction terms for the type of error an agent made and the treatment to allow for the possibility that overestimation and underestimation are viewed differently between treatments.

Table 7 reports regression estimates using Equation 1 and Table 8 reports regression estimates using Equation 2. In all three specifications in both tables, the coefficient on $AccurateDirection$ is positive and statistically significant meaning that principals rate agents higher when the agent’s rank order Jars in a manner consistent with the independent information the principal receives. From all three specifications in both Tables 7 and 8 it is clear that agents are penalized for errors in their estimation. Specifications (2) and (3) in both tables indicate that ratings are not systematically different across treatments ($AccurateDirection$ & $Disclosure$ is not significant in any specification). The lower panel of both tables provides the p -values for testing for symmetry among the types of errors that an agent could make in her estimations. For each case there are 6 possible tests since there are four coefficients for over- and underestimation.¹²

From Table 7 there is no evidence that overestimation of Jar A is treated differently than overestimation of Jar B, nor is underestimation viewed differently between jars. However, underestimation of a jar is viewed more negatively than overestimation of the same jar, although this difference is not always statistically significant across specifications and treatments. Interestingly, greater concern for underestimation as compared to overestimation holds across jars as well, at least for the Baseline. However, in the Disclosure Treatment when the principal knows the agent has an incentive to make Jar A appear more attractive, this cross-jar asymmetry disappears. That is, in the Disclosure Treatment, the principal views overestimating Jar A the same as underestimating Jar B as these two errors both make Jar A appear more attractive.

Unsurprisingly, Table 8 shows that overestimation of the jar selected is viewed more negatively than underestimation of the jar selected. Additionally, overestimation of the jar selected is viewed more negatively than overestimation of the jar not selected, which is intuitive. Overall, errors regarding the jar selected are viewed more negatively than oppositely signed errors for the non-selected jar. Thus, overestimation of the selected jar leads to a greater reduction in rating than does underestimating the value of the non-selected jar even though both errors have the same impact on assessing the difference in value between the two jars. But this asymmetry is limited to the Baseline as indicated by the tests of specification (3).

The results reported in Tables 7 and 8 support the final finding.

¹²For specification (3) in Tables 7 and 8 there are 6 tests for each treatment since the impact of errors is allowed to vary by treatment.

Table 7: Estimated Impacts on Ratings Given to Agent using Jar Identity

VARIABLES	(1)	(2)	(3)
AccuateDirection	0.555*** (0.134)	0.506*** (0.114)	0.486*** (0.115)
OverEstJarA	-0.094*** (0.022)	-0.094*** (0.022)	-0.127*** (0.024)
UnderEstJarA	-0.175*** (0.024)	-0.176*** (0.024)	-0.208*** (0.022)
OverEstJarB	-0.117*** (0.023)	-0.117*** (0.024)	-0.095*** (0.028)
UnderEstJarB	-0.156*** (0.030)	-0.155*** (0.030)	-0.210*** (0.029)
Disclosure		-0.037 (0.216)	-0.251 (0.305)
AccuateDirection & Disclosure		0.097 (0.266)	0.137 (0.265)
OverEstJarA & Disclosure			0.068 (0.044)
UnderEstJarA & Disclosure			0.048 (0.045)
OverEstJarB & Disclosure			-0.044 (0.046)
UnderEstJarB & Disclosure			0.103* (0.056)
Constant	3.962*** (0.159)	3.982*** (0.168)	4.099*** (0.178)
Observations	1,044	1,044	1,044
Number of Unique IDs	36	36	36
Tests of symmetry			
OverEstJarA = UnderEstJarA	0.000	0.000	0.001
OverEstJarB = UnderEstJarB	0.131	0.153	0.000
OverEstJarA = OverEstJarB	0.496	0.500	0.341
UnderEstJarA = UnderEstJarB	0.633	0.608	0.936
OverEstJarA = UnderEstJarB	0.077	0.083	0.015
UnderEstJarA = OverEstJarB	0.091	0.086	0.000
OverEstJarA + OverEstJarADisc = UnderEstJarA + UnderEstJarADisc			0.005
OverEstJarB + OverEstJarBDisc = UnderEstJarB + UnderEstJarBDisc			0.449
OverEstJarA + OverEstJarADisc = OverEstJarB + OverEstJarBDisc			0.170
UnderEstJarA + UnderEstJarADisc = UnderEstJarB + UnderEstJarBDisc			0.465
OverEstJarA + OverEstJarADisc = UnderEstJarB + UnderEstJarBDisc			0.437
UnderEstJarA + UnderEstJarADisc = OverEstJarB + OverEstJarBDisc			0.715

The results in the first four rows for specification (3) examine symmetry in the baseline and the last four test rows of specification test for symmetry in Disclosure.

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 8: Estimated Impacts on Ratings Given to Agent Using Jar Selection

VARIABLES	(1) RateAgnt	(2) RateAgnt	(3) RateAgnt
AccuateDirection	0.517*** (0.113)	0.436*** (0.112)	0.501*** (0.111)
OverEstPick	-0.244*** (0.027)	-0.246*** (0.027)	-0.274*** (0.031)
UnderEstPick	-0.108*** (0.031)	-0.108*** (0.031)	-0.161*** (0.033)
OverEstOther	-0.050** (0.024)	-0.047* (0.026)	-0.087*** (0.024)
UnderEstOther	-0.106*** (0.022)	-0.107*** (0.023)	-0.068*** (0.017)
Disclosure		-0.098 (0.218)	-0.127 (0.272)
AccuateDirection & Disclosure		0.159 (0.263)	0.045 (0.226)
OverEstPick & Disclosure			0.058 (0.053)
UnderEstPick & Disclosure			0.110* (0.062)
OverEstOther & Disclosure			0.056 (0.047)
UnderEstOther & Disclosure			-0.081* (0.043)
Constant	4.011*** (0.139)	4.064*** (0.166)	4.068*** (0.179)
Observations	1,044	1,044	1,044
Number of Unique IDs	36	36	36
OverEstPick = UnderEstPick	0.000	0.000	0.000
OverEstOther = UnderEstOther	0.0630	0.0680	0.476
OverEstPick = OverEstOther	0.000	0.000	0.000
UnderEstPick = UnderEstOther	0.958	0.976	0.006
OverEstPick = UnderEstOther	0.000	0.000	0.000
OverEstOther = UnderEstPick	0.100	0.083	0.080
OverEstPick + OverEstPickDisc = UnderEstPick + UnderEstPickDisc			0.001
OverEstOther + OverEstOtherDisc = UnderEstOther + UnderEstOtherDisc			0.030
OverEstPick + OverEstPickDisc = OverEstOther + OverEstOtherDisc			0.008
UnderEstPick + UnderEstPickDisc = UnderEstOther + UnderEstOtherDisc			0.220
OverEstPick + OverEstPickDisc = UnderEstOther + UnderEstOtherDisc			0.256
OverEstOther + OverEstOtherDisc = UnderEstPick + UnderEstPickDisc			0.730

The results in the first four rows for specification (3) examine symmetry in the baseline and the last four test rows of specification test for symmetry in Disclosure.

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Finding 5: Principals reward agents with higher ratings when the agent appears to correctly identify the best option and punish agents with lower ratings for misestimation. The reduction in rating tends to be greater for underestimation than overestimation of a given jar; however, with Disclosure principals view overestimating one jar and underestimating the other jar as comparable. After selecting a jar, principals are more strongly concerned with errors regarding the selected jar; but, with Disclosure principals view overestimation of the selected jar as being comparable to underestimation of the other jar.

4 Discussion

Conventional wisdom suggests that an agent’s conflict of interest should be disclosed to principals. However, such a policy may prove harmful if disclosure creates a moral license that emboldens agents and encourages them to act in self-serving ways, especially if the magnitude of this effect is not fully appreciated by principals. This negative effect was reported in Cain et al. (2005). Our paper extends the basic structure of Cain et al. (2005) to a repeated interaction situation in which agents have reputation concerns. The goal of our paper is two-fold. First, we seek to understand if the disciplining effect of maintaining a reputation can offset the incentives created by a conflict of interest. Second, we examine the effect of disclosure in a setting where agents have reputation concerns.

In our experimental task, principals rated agents based on the agent’s performance estimating the value of options available to the principal. In turn, principals could select agents based upon those ratings. Agents had a conflict of interest in that they received a direct benefit from increasing the principals belief about the relative value of a particular option. In the absence of disclosure, agents did not provide self-serving biased estimates. This indicates that reputation concerns alone can be enough to offset conflicts of interest. However, when disclosure of the conflict of interest was required self-serving behavior on the part of agents was observed. The observed manipulation involved a mix of inflating estimates of the option that provided a benefit to the agent and deflating the estimate of the other option. Thus, we find that the counterintuitive result of Cain et al. (2005) that disclosure can exacerbate a conflict of interest holds even in the presence of reputation concerns.

While not the main focus of our study, the observed behavior does provide some additional insights. First, we find that mis-estimation of an option, whether it is ultimately the one selected or not, leads to a reduction in the rating a principal assigns to an agent. However, in the Baseline overestimating the benefit from a selected option leads to a greater reduction in the rating than underestimation of the value of an option not selected despite these errors having similar implications for the principal. But, with Disclosure principals view the two errors similarly. Second, we find that disclosure does not change how principals evaluate agents given their performance. In particular, disclosure does not lead to differences in how over- or underestimation of a particular option is treated. Finally, we note that our subjects were unable to accurately guess the value of the options, which were jars of coins, in a separate experiment used to calibrate unbiased expectations of those values. Instead, average estimates were systematically below the actual values contrary to the arguments in Surowiecki (2005).

It is worth noting that the success of reputation concerns in deterring self-serving behavior by agents may be dependent upon two key features of our setting. First, principals in our study were effectively consuming a credence good since they could not verify the quality of their choice. However, the principals did receive a noisy signal of the quality. Second, the size of the conflict of interest was relatively small in that a majority of the agent’s earnings were from being selected by a principal rather than having the principal select a particular choice. Whether a conflict of interest would lead to self-serving behavior when reputation concerns are present when either or both of these features is relaxed remains an important question. Similarly, the negative effects of disclosure that we observe may or may not hold as these two factors are relaxed. It is certainly possible that principals who know that agents have a substantial conflict of interest or who have weaker signals of agent performance may behave differently. We hope that our work will spark interest in these important avenues for future research.

References

- Becker, G. S. (1968, April). Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76(2), 169.
- Ben-shahar, O. and C. E. Schneider (2011). The Failure of Mandated Disclosure. *University of Pennsylvania Law Review* 159(3), 647. Publisher: University of Pennsylvania Law School.
- Boatright, J. R. (2000). Conflicts of Interest in Financial Services. *Business and Society Review* 105(2), 201–219. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/0045-3609.00078>.
- Cain, D. M., G. Loewenstein, and D. A. Moore (2005). The dirt on coming clean: Perverse effects of disclosing conflicts of interest. *The Journal of Legal Studies* 34(1), 1–25.
- Cain, D. M., G. Loewenstein, and D. A. Moore (2011, February). When Sunlight Fails to Disinfect: Understanding the Perverse Effects of Disclosing Conflicts of Interest. *Journal of Consumer Research* 37(5), 836–857.
- Campbell, E. G., R. L. Gruen, J. Mountford, L. G. Miller, P. D. Cleary, and D. Blumenthal (2007, April). A National Survey of Physician–Industry Relationships. *New England Journal of Medicine* 356(17), 1742–1750. Publisher: Massachusetts Medical Society _eprint: <https://doi.org/10.1056/NEJMsa064508>.
- Carl, W. (2008, July). The role of disclosure in organized word-of-mouth marketing programs. *Journal of Marketing Communications* 14(3), 225–241. Publisher: Routledge.
- Chen, D. L., V. Levonyan, S. E. Reinhart, and G. Taksler (2019, June). Mandatory Disclosure: Theory and Evidence from Industry-Physician Relationships. *The Journal of Legal Studies* 48(2), 409–440.
- Chen, D. L., M. Schonger, and C. Wickens (2016, March). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance* 9, 88–97.
- Church, B. and X. Kuang (2009, June). Conflicts of Interest, Disclosure, and (Costly) Sanctions: Experimental Evidence. *The Journal of Legal Studies* 38(2), 505–532.
- Darby, M. R. and E. Karni (1973). Free competition and the optimal amount of fraud. *The Journal of law and economics* 16(1), 67–88.
- Dranove, D., D. Kessler, M. McClellan, and M. Satterthwaite (2003). Is More Information Better? The Effects of “Report Cards” on Health Care Providers. *Journal of Political Economy* 111(3), 555. Publisher: The University of Chicago Press.
- Gibbons, R. V., F. J. Landry, D. L. Blouch, D. L. Jones, F. K. Williams, C. R. Lucey, and K. Kroenke (1998, March). A comparison of physicians’ and patients’ attitudes toward pharmaceutical industry gifts. *Journal of General Internal Medicine* 13(3), 151–154.
- Grande, D., J. A. Shea, and K. Armstrong (2012, March). Pharmaceutical Industry Gifts to Physicians: Patient Beliefs and Trust in Physicians and the Health Care System. *Journal of General Internal Medicine* 27(3), 274–279.
- Green, M. J., R. Masters, B. James, B. Simmons, and E. Lehman (2012, May). Do Gifts From the Pharmaceutical Industry Affect Trust in Physicians? *FAMILY MEDICINE* 44(5), 7.
- Guo, T. and S. Sriram (2017). ‘Let the Sun Shine in’: The Impact of Industry Payment Disclosure on Physician Prescription Behavior. *SSRN Electronic Journal*.
- Guo, T., S. Sriram, and P. Manchanda (2017). The Effect of Information Disclosure on Industry Payments to Physicians. *SSRN Electronic Journal*.

- Hannan, E. L., H. Kilburn, M. Racz, E. Shields, and M. R. Chassin (1994, March). Improving the Outcomes of Coronary Artery Bypass Surgery in New York State. *JAMA* 271(10), 761–766. Publisher: American Medical Association.
- Kerschbamer, R. and M. Sutter (2017). The economics of credence goods—a survey of recent lab and field experiments. *CEifo Economic Studies* 63(1), 1–23.
- Levitt, S. D. and J. A. List (2007). What do laboratory experiments measuring social preferences reveal about the real world? *The Journal of Economic Perspectives* 21(2), 153–174.
- Mainous, A. G. (1995, April). Patient perceptions of physician acceptance of gifts from the pharmaceutical industry. *Archives of Family Medicine* 4(4), 335–339.
- Mukamel, D. B. and A. I. Mushlin (1998). Quality of Care Information Makes a Difference: An Analysis of Market Share and Price Changes after Publication of the New York State Cardiac Surgery Mortality Reports. *Medical Care* 36(7), 945–954. Publisher: Lippincott Williams & Wilkins.
- Pope, D. G. (2009). Reacting to rankings: evidence from “America’s Best Hospitals”. *Journal of health economics* 28(6), 1154–1165.
- Rose, S. L., S. Sah, R. Dweik, C. Schmidt, M. Mercer, A. Mitchum, M. Kattan, M. Karafa, and C. Robertson (2019, March). Patient responses to physician disclosures of industry conflicts of interest: A randomized field experiment. *Organizational Behavior and Human Decision Processes*.
- Scholl, B. and A. A. Hung (2018). The Retail Market for Investment Advice. Technical report, Security and Exchange Commission, Washington, D.C.
- Surowiecki, J. (2005). *The wisdom of crowds* (1st Anchor books ed ed.). New York: Anchor Books. OCLC: 61254310.
- Wolinsky, A. (1993). Competition in a Market for Informed Experts’ Services. *The RAND Journal of Economics* 24(3), 380–398.

Appendices

A.1 Instructions

The experiment instructions were computerized, but subjects were informed that they could raise their hand if they had any questions. Screenshots of instructions are below.

Guess Treatment

The Guess Treatment had two instruction screens.

Pay





Welcome and thank you for participating in this experiment! Your task is to view photos of jars containing coins and then correctly guess the value of the coins in the jar.

There will be 30 rounds. Each round your lab earnings equals $\$25 - \$2 * (\text{Actual value} - \text{Your guess})^2$. If your earnings would be less than zero for a round, then they will be zero. This equation may look complicated but what it means is that you should provide the most accurate guess you can in order to earn the most money that you can. The more accurate your prediction the more you will earn. Your final payment in \$US will be the AVERAGE of your lab earnings across all rounds plus the show-up payment.

Next

The Task

All the pictures in the study are of jars containing between \$10.00 and \$25.00. To help everyone understand the task, here are some example pictures. Notice, that the height of the coins is not the only factor to consider when estimating value. The composition of the coins must also be considered. For the examples below, the two jars on the left are only about half full, while one contains \$16.00 and the \$22.00. However, the two on the right are nearly filled to the top, and also \$16.00 and \$22.00.

\$16.00	\$22.00	\$16.00	\$22.00
			

Next

Main Treatments

The main treatments had four instruction screens with conditional text. Annotation shows what text was conditional on treatment.

Instructions

Welcome and thank you for participating in this experiment! There are two roles in the experiment, 'Investors' and 'Experts'. You will be assigned one of these roles and remain in that role for the entire experiment.

There will be **four blocks**, each which will last a random number of rounds. Each block lasts at least 5 rounds. Every round after the 5th round there is 25% chance the block will end.

[Next](#)

Steps

Each round you will be assigned to a group, consisting of two *Investors* and two *Experts*. Your group is randomly determined each round so you will be interacting with different people over the course of the block and across blocks although you may encounter the same people multiple times.

1. Each *Investor* picks an *Expert* for the round (except the first round when they are randomly assigned).
2. Each round, *Experts* see photos of two jars, A and B, and estimate the value of the coins in each jar.
3. *Investors* then pick an *Expert* and see that *Expert's* estimates for both jars, but will not see the photos or the estimates of the *Expert* not picked.
4. The *Investor* then picks a jar and learns the exact amount in the chosen jar. The *Investor* will also get a computer generated signal about the amount in the jar that was not selected. The signal is the actual amount in the jar that was not chosen plus a random number (between -\$2 and +\$2).
5. The *Investor* then rates the chosen *Expert*.
6. At the start of every round after the first round, *Investors* will be able to see the ratings of the two *Experts* in the group.

[Next](#)

Pay

The *Investors's* earning for a round equals the actual amount in the Jar. At the end of the experiment, one round will be randomly selected and the amount in Jar selected by the *Investors* that period will be the payoff the *Investors* receives.





An *Expert's* earnings depend in part on the number of *Investors* who pick that *Expert*. Every round an *Expert* earns \$0.75 for each *Investor* who picks that *Expert*. In addition, an *Expert* earns an extra \$0.50 for each of those *Investors* who goes on to pick Jar A. Only in Disclosure Treatment

At the end of the experiment, two blocks will be randomly selected for each *Expert*. *Experts* will receive all money earned in all periods in those blocks.

[Next](#)

The Task

All the pictures in the study are of jars containing between \$10.00 and \$25.00. To help everyone understand the task, here are some example pictures. Notice, that the height of the coins is not the only factor to consider when estimating value. The composition of the coins must also be considered. For the examples below, the two jars on the left are only about half full, while one contains \$16.00 and the \$22.00. However, the two on the right are nearly filled to the top, and also \$16.00 and \$22.00.

\$16.00	\$22.00	\$16.00	\$22.00
			

[Next](#)

A.2 Programming Error

In Round 1, agents saw Jars 1 and 2; for Agent 1 Jar 1 was Jar A and Jar 2 was Jar B, vice versa for Agent 2. In Round 2, agents saw Jars 3 and 4, assigned to A and B as above. Rounds continued like this until Round 15. In Round 15, agents saw Jars 1 and 2 a second time, with the assignment to A and B reversed. For Agent 2 Jar 1 was Jar A and Jar 2 was Jar B. This reversal occurred in Rounds 16 - 28 as well with each round being a swap of the round 14 rounds earlier. However, in round 29, the last given the random termination, Agents saw Jar 29 and Jar 30, which they had not seen before. Because of the odd number of rounds not every jar was observed twice.

Table A.1: KEEP ALL Regression Estimates for Agents' Estimates of Jars

VARIABLES	(1)	(2)	(3)	(4)
Constant	9.254*** (0.655)	3.544*** (0.529)	-131.457*** (12.882)	3.011** (1.279)
Value	0.399*** (0.029)		8.889*** (0.770)	
MeanGuess		0.880*** (0.024)		0.926*** (0.080)
Baseline & JarA	-0.205 (0.200)	-0.205 (0.163)	-0.205 (0.175)	-0.205 (0.175)
Disclosure & Jar A	-0.075 (0.560)	-0.075 (0.554)	-0.075 (0.573)	-0.075 (0.573)
Disclosure & Jar B	-0.480 (0.560)	-0.480 (0.554)	-0.480 (0.515)	-0.480 (0.515)
Observations	2,088	2,088	2,088	2,088
Number of Unique IDs	36	36	36	36
Overall R2	0.0701	0.318	0.356	0.356
Jar # Controls	N	N	Y	Y
Round # Controls	N	N	Y	Y
Clustering	N	N	Y	Y
JarA - JarB for Disclosure	0.405	0.405	0.405	0.405
SE	0.200	0.163	0.223	0.223
P> t	0.043	0.013	0.069	0.069
Standard errors in parentheses		*** p<0.01, ** p<0.05, * p<0.1		

Had every subject always been Agent 1 or always been Agent 2, every agent would have two estimates for each jar (1 to 28); one when the jar was Jar A and one when it was Jar B. However, there was an error in the logic of our program. A subject who was Agent 1 in the first half, e.g., Round 1, of the experiment could be Agent 2 14 rounds later, e.g., Round 15. As a result, while we have two estimates for each jar from 1-28 from each agent, for some agents some jars were estimated twice as Jar A and for others the jars were estimated twice as Jar B. For the analysis in the body of the paper, we only included estimates for which there were Jar A and Jar B estimates for the same jar by the same agent. Table A.1 reports estimates from repeating the regressions in Table 4 including all the estimates. The results are qualitatively similar. In the Baseline, guesses for the value of Jar B are larger than those for Jar A, but not by a statistically significant margin. In contrast, in the Disclosure Treatment, the guesses for the value of Jar A were 41¢ higher than those for Jar B.

A.3 Ordered Probit Regressions

Tables A.2 through A.4 present results from analysis complementary to that reported in Table 7. However, rather than treating the number of stars awarded to an agent as a continuous variable, this supplementary analysis treats the ratings as discrete ranked options and thus uses ordered probit regressions. Each table reports the marginal effects of the probit regression in parallel to one of the specifications in Table 8. For these tables, each column is the effect from a single independent variable and each row is the change in the likelihood of an agent receiving that number of stars for a one unit change in the variable. The results are consistent with those from Table 8.

Table A.2: Order Probit of Impacts on Rating given to Agent, Jar A vs. B, Model 1

VARIABLES	(1) DIRECTACC	(2) OverEstJarA	(3) UnderEstJarA	(4) OverEstJarB	(5) UnderEstJarB
1 Star	-0.0640*** (0.0189)	0.0106*** (0.00312)	0.0192*** (0.00391)	0.0131*** (0.00302)	0.0176*** (0.00387)
2 Stars	-0.0785*** (0.0179)	0.0130*** (0.00342)	0.0236*** (0.00377)	0.0160*** (0.00344)	0.0216*** (0.00467)
3 Stars	-0.0512*** (0.0125)	0.00845*** (0.00215)	0.0154*** (0.00346)	0.0105*** (0.00319)	0.0141*** (0.00402)
4 Stars	0.0440*** (0.0109)	-0.00727*** (0.00231)	-0.0132*** (0.00267)	-0.00900*** (0.00230)	-0.0121*** (0.00324)
5 Stars	0.150*** (0.0342)	-0.0247*** (0.00560)	-0.0449*** (0.00660)	-0.0306*** (0.00645)	-0.0412*** (0.00793)
Observations	1,044	1,044	1,044	1,044	1,044
Standard errors in parentheses			*** p<0.01, ** p<0.05, * p<0.1		

Table A.3: Order Probit of Impacts on Rating given to Agent, Jar A vs. B, Model 2

VARIABLES	(1) DIRECTACC	(2) OverEstJarA	(3) UnderEstJarA	(4) OverEstJarB	(5) UnderEstJarB	(6) DISC	(7) DirectAccDISC
1 Star	-0.0603*** (0.0143)	0.0106*** (0.00311)	0.0193*** (0.00394)	0.0131*** (0.00300)	0.0176*** (0.00384)	0.00277 (0.0231)	-0.00725 (0.0302)
2 Stars	-0.0741*** (0.0180)	0.0130*** (0.00333)	0.0236*** (0.00372)	0.0161*** (0.00347)	0.0216*** (0.00478)	0.00341 (0.0281)	-0.00890 (0.0365)
3 Stars	-0.0483*** (0.0132)	0.00849*** (0.00208)	0.0154*** (0.00341)	0.0105*** (0.00319)	0.0141*** (0.00406)	0.00222 (0.0184)	-0.00580 (0.0237)
4 Stars	0.0416*** (0.0116)	-0.00730*** (0.00228)	-0.0133*** (0.00267)	-0.00900*** (0.00234)	-0.0121*** (0.00331)	-0.00191 (0.0157)	0.00499 (0.0204)
5 Stars	0.141*** (0.0293)	-0.0248*** (0.00549)	-0.0451*** (0.00659)	-0.0306*** (0.00645)	-0.0411*** (0.00804)	-0.00649 (0.0538)	0.0170 (0.0700)
Observations	1,044	1,044	1,044	1,044	1,044	1,044	1,044
Standard errors in parentheses			*** p<0.01, ** p<0.05, * p<0.1				

Table A.4: Order Probit of Impacts on Rating given to Agent, Jar A vs. B, Model 3

VARIABLES	DIRECTACC	OverEstJarA	UnderEstJarA	OverEstJarB	UnderEstJarB	DISC
1 Star	-0.0582*** (0.0142)	0.0145*** (0.00364)	0.0231*** (0.00439)	0.0109*** (0.00361)	0.0239*** (0.00466)	0.0304 (0.0345)
2 Stars	-0.0715*** (0.0179)	0.0178*** (0.00391)	0.0283*** (0.00387)	0.0134*** (0.00371)	0.0294*** (0.00441)	0.0373 (0.0391)
3 Stars	-0.0468*** (0.0131)	0.0116*** (0.00268)	0.0185*** (0.00409)	0.00876*** (0.00328)	0.0192*** (0.00496)	0.0244 (0.0268)
4 Stars	0.0399*** (0.0113)	-0.00991*** (0.00283)	-0.0158*** (0.00356)	-0.00747*** (0.00247)	-0.0164*** (0.00339)	-0.0208 (0.0219)
5 Stars	0.137*** (0.0294)	-0.0339*** (0.00606)	-0.0541*** (0.00593)	-0.0256*** (0.00759)	-0.0562*** (0.00855)	-0.0712 (0.0779)

VARIABLES	DirectAccDISC	OverEstJarADisc	UnderEstJarADisc	OverEstJarBDisc	UnderEstJarBDisc
1 Star	-0.0111 (0.0298)	-0.00789 (0.00492)	-0.00587 (0.00513)	0.00421 (0.00522)	-0.0120* (0.00655)
2 Stars	-0.0136 (0.0355)	-0.00970 (0.00614)	-0.00722 (0.00619)	0.00517 (0.00671)	-0.0147** (0.00743)
3 Stars	-0.00891 (0.0232)	-0.00634 (0.00419)	-0.00472 (0.00424)	0.00338 (0.00432)	-0.00963* (0.00552)
4 Stars	0.00760 (0.0197)	0.00541 (0.00355)	0.00403 (0.00375)	-0.00288 (0.00370)	0.00822** (0.00410)
5 Stars	0.0260 (0.0688)	0.0185 (0.0115)	0.0138 (0.0117)	-0.00987 (0.0125)	0.0281* (0.0150)
Observations	1,044	1,044	1,044	1,044	1,044

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Tables A.5 through A.7 present results from analysis complementary to that reported in Table 8. However, rather than treating the number of stars awarded to an agent as a continuous variable, this supplementary analysis treats the ratings as discrete ranked options and thus uses ordered probit regressions. Each table reports the marginal effects of the probit regression in parallel to one of the specifications in Table 8. For these tables, each column is the effects from a single independent variable and each row is the change in the likelihood of an agent receiving that number of stars for a one unit change in the variable. The results are consistent with those from Table 8.

Table A.5: Order Probit of Impacts on Rating given to Agent, Selected vs. Other, Model 1

VARIABLES	(1) DIRECTACC	(2) OverEstPick	(3) UnderEstPick	(4) OverEstOther	(5) UnderEstOther
1 Star	-0.0592*** (0.0156)	0.0274*** (0.00486)	0.0122*** (0.00353)	0.00496* (0.00273)	0.0117*** (0.00334)
2 Star	-0.0728*** (0.0156)	0.0337*** (0.00452)	0.0150*** (0.00461)	0.00610* (0.00350)	0.0144*** (0.00320)
3 Star	-0.0485*** (0.0116)	0.0224*** (0.00463)	0.00999*** (0.00363)	0.00406 (0.00251)	0.00962*** (0.00223)
4 Star	0.0404*** (0.00918)	-0.0187*** (0.00361)	-0.00831*** (0.00284)	-0.00338* (0.00205)	-0.00800*** (0.00227)
5 Star	0.140*** (0.0297)	-0.0649*** (0.00753)	-0.0289*** (0.00835)	-0.0117* (0.00656)	-0.0278*** (0.00569)
Observations	1,044	1,044	1,044	1,044	1,044

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table A.6: Order Probit of Impacts on Rating given to Agent, Selected vs. Other, Model 2

VARIABLES	(1) DIRECTACC	(2) OverEstPick	(3) UnderEstPick	(4) OverEstOther	(5) UnderEstOther	(6) DISC	(7) DirectAccDISC
1 star	-0.0521*** (0.0126)	0.0275*** (0.00491)	0.0122*** (0.00356)	0.00473* (0.00282)	0.0118*** (0.00339)	0.00946 (0.0235)	-0.0139 (0.0300)
2 Stars	-0.0641*** (0.0178)	0.0339*** (0.00442)	0.0150*** (0.00463)	0.00582 (0.00368)	0.0145*** (0.00323)	0.0116 (0.0280)	-0.0171 (0.0357)
3 Stars	-0.0427*** (0.0131)	0.0225*** (0.00453)	0.00996*** (0.00363)	0.00387 (0.00261)	0.00968*** (0.00224)	0.00775 (0.0189)	-0.0114 (0.0237)
4 Stars	0.0355*** (0.0112)	-0.0187*** (0.00360)	-0.00827*** (0.00288)	-0.00322 (0.00216)	-0.00804*** (0.00228)	-0.00644 (0.0153)	0.00947 (0.0195)
5 Stars	0.123*** (0.0293)	-0.0652*** (0.00754)	-0.0288*** (0.00838)	-0.0112 (0.00684)	-0.0280*** (0.00583)	-0.0224 (0.0551)	0.0330 (0.0698)
Observations	1,044	1,044	1,044	1,044	1,044	1,044	1,044

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.7: Order Probit of Impacts on Rating given to Agent, Selected vs. Other, Model 3

VARIABLES	(1) DIRECTACC	(2) OverEstPick	(3) UnderEstPick	(4) OverEstOther	(5) UnderEstOther	(6) DISC
1 Star	-0.0599*** (0.0137)	0.0311*** (0.00549)	0.0185*** (0.00457)	0.00899*** (0.00310)	0.00754*** (0.00242)	0.0161 (0.0301)
2 Stars	-0.0736*** (0.0177)	0.0382*** (0.00487)	0.0227*** (0.00452)	0.0110*** (0.00359)	0.00926*** (0.00264)	0.0197 (0.0357)
3 Stars	-0.0491*** (0.0126)	0.0255*** (0.00586)	0.0151*** (0.00414)	0.00736*** (0.00259)	0.00618*** (0.00189)	0.0132 (0.0247)
4 Stars	0.0404*** (0.0117)	-0.0210*** (0.00441)	-0.0124*** (0.00335)	-0.00606*** (0.00227)	-0.00508*** (0.00177)	-0.0108 (0.0194)
5 Stars	0.142*** (0.0280)	-0.0739*** (0.00902)	-0.0438*** (0.00868)	-0.0213*** (0.00660)	-0.0179*** (0.00479)	-0.0381 (0.0709)
Observations	1,044	1,044	1,044	1,044	1,044	1,044

VARIABLES	(7) DirectAccDISC	(8) UnderEstPickDisc	(9) OverEstPickDisc	(10) UnderEstOtherDisc	(11) OverEstOtherDisc
1 Star	-0.000301 (0.0248)	-0.0131* (0.00738)	-0.00737 (0.00597)	0.00873* (0.00525)	-0.00595 (0.00541)
2 Stars	-0.000370 (0.0305)	-0.0161** (0.00787)	-0.00906 (0.00727)	0.0107* (0.00599)	-0.00730 (0.00632)
3 Stars	-0.000247 (0.0203)	-0.0107* (0.00570)	-0.00604 (0.00538)	0.00715* (0.00400)	-0.00487 (0.00410)
4 Stars	0.000203 (0.0167)	0.00881* (0.00471)	0.00497 (0.00423)	-0.00588* (0.00341)	0.00401 (0.00345)
5 Stars	0.000716 (0.0589)	0.0311* (0.0159)	0.0175 (0.0143)	-0.0207* (0.0116)	0.0141 (0.0123)
Observations	1,044	1,044	1,044	1,044	1,044

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1