Gneezy, Uri; Saccardo, Silvia; van Veldhuizen, Roel

Working Paper

Bribery: Greed versus reciprocity

WZB Discussion Paper, No. SP II 2016-203

Provided in Cooperation with:
WZB Berlin Social Science Center

Suggested Citation: Gneezy, Uri; Saccardo, Silvia; van Veldhuizen, Roel (2016) : Bribery: Greed versus reciprocity, WZB Discussion Paper, No. SP II 2016-203, Wissenschaftszentrum Berlin für Sozialforschung (WZB), Berlin

This Version is available at:
http://hdl.handle.net/10419/130838

Wissenschaftszentrum Berlin
für Sozialforschung

Uri Gneezy
Silvia Saccardo
Roel van Veldhuizen

Bribery: Greed versus Reciprocity

Discussion Paper

SP II 2016–203

April 2016

WZB Berlin Social Science Center

Research Area
Markets and Choice

Research Unit
Market Behavior

Wissenschaftszentrum Berlin für Sozialforschung gGmbH
Reichpietschufer 50
10785 Berlin
Germany
www.wzb.eu

Uri Gneezy, Silvia Saccardo, Roel van Veldhuizen
Bribery: Greed versus Reciprocity

Affiliation of the authors:

Uri Gneezy
University of California San Diego

Silvia Saccardo
Carnegie Mellon University

Roel van Veldhuizen
WZB

Wissenschaftszentrum Berlin für Sozialforschung gGmbH
Reichpietschufer 50
10785 Berlin
Germany
www.wzb.eu

Abstract

Bribery: Greed versus Reciprocity

by Uri Gneezy, Silvia Saccardo, Roel van Veldhuizen*

It is estimated that a trillion dollars are annually exchanged in bribes, distorting justice and economic efficiency. In a novel experiment, we investigate the drivers of bribery. Two participants compete for a prize; a referee picks the winner. Participants can bribe the referee. When the referee can keep only the winner's bribe, bribes distort her judgment. When the referee keeps the bribes regardless of the winner, bribes no longer influence her judgment. An extra-laboratory experiment in an Indian market confirms these results. Hence, our participants are influenced by bribes out of greed, and not because of a desire to reciprocate.

Keywords: Bribery, Reciprocity, Laboratory Experiment, Extra–Laboratory experiment

JEL classification: D73, C91, K42

* E-mail: ugneezy@ucsd.edu, ssaccard@andrew.cmu.edu, roel.vanveldhuizen@wzb.eu.

I. Introduction

Bribery affects economic activities around the world. Because it is illegal in most places, getting good empirical data about these activities is difficult. However, the existing data show bribery is likely widespread. The World Bank estimates that $1 trillion exchanges hands in bribes annually (Kaufmann, 2005), and many companies report having to pay bribes to win business—from 15% to 20% in industrialized countries, to 40% in China, Russia, and Mexico (Transparency International, 2011). In some places, these kinds of activities are a major source of income. For example, bribes are estimated to amount to 20% of Russia's GDP in 2005 (INDEM, 2005).

Importantly, bribes are often given in small scale interactions, and not necessarily using money. Examples include paying police person to avoid a ticket for minor traffic violations, sending gifts to teachers, offering free tickets or trips to physicians, etc.

But why do bribes "work?" In particular, if one of the sides in a bribery case does not fulfill his part, the other side cannot take him to court or use traditional enforcement mechanisms. What prevents one, for example, from accepting a payment but then not providing the good? If receiving the bribe is credibly contingent on the briber's success (e.g., winning a contest, or in the case of repeated interactions), traditional economic models with selfish agents can explain behavior. In other one-shot cases in which receiving a bribe is not contingent on delivering the desired outcome, traditional economic assumptions may not be sufficient. In these cases, social preferences may be able to explain the success of bribery. People might be engaged in reciprocal behavior in which one side gives a "gift" and the other reciprocates (Akerlof, 1982; Rabin, 1993; Fehr and Gächter, 2000).

To reduce bribery, it is important to understand what drives it. In the process of understanding the motivation for bribery, experiments are an important tool because they can help us isolate key aspects of the relevant behavior. Our paper distinguishes between the two rival motivations for bribery discussed above: reciprocity and greed (i.e., payoff maximization). This distinction is important from a public policy perspective. If reciprocity drives bribery, policy interventions should focus on reducing social ties and making reciprocity more difficult, for example, by decreasing personal contact through anonymity and staff rotation. If greed drives behavior and individuals only care about maximizing their profit, such policy interventions will not prevent people from engaging in bribery, and traditional anti-corruption methods based on auditing and sanctions may be more effective (Becker and Stigler, 1974; Olken, 2007).

Our ability to distinguish between different motives for bribery comes from the novel game we study. The game captures an important feature that distinguishes bribery from other transactions: a distortionary effect. This kind of distortion is a key element in bribery and occurs when a decision maker uses bribes rather than other objective criteria such as merit, performance, or quality to determine who receives a particular outcome. As a result, public resources may go to the more corrupt people, not necessarily the most talented ones (Pareto, 1896; Goldsmith, 1999; Del Monte & Papagni, 2001). A large empirical literature has shown that such outcomes have detrimental effects on efficiency (see, e.g., Mauro, 1995; Reinikka and Svennson, 2004; Bertrand, Djankov, Hanna, and Mullainathan, 2007; Sequeira and Djankov, 2014; or see Olken and Pande, 2012, for a review).

We are not the first to study bribery using experiments: the existing experimental literature examines different elements of bribing behavior, from the effect of staff rotation (Abbink, 2004) and asymmetric liability (Abbink, Dasgupta, Gangadharan, and Jain, 2014) to culture (Barr and Serra, 2010; Cameron, Chaudhuri, Erkal, and Gangadharan, 2009) and the influence of wages

(Abbink, 2005; Armantier and Boly, 2013; Van Veldhuizen, 2013). See Abbink and Serra (2012) for a comprehensive survey of these experiments. However, in this literature, participants are asked to choose between different monetary allocations. These decisions may include negative externalities on a third party, but they do not include a distortion of facts or judgment. The ability to study the effect of distortion of judgment is, as we show, critical to understanding some bribery behaviors.[1]

To capture this key element, we introduce a new bribery game in which two participants ("the workers") compete on a task. A third participant, the referee, then chooses the winner, who gets a prize. Importantly, as we explain below, the judgment of the quality of the task is subjective. Apart from working on the task, the two workers can also choose to send bribes to the referee. We use this basic design to test whether workers actually send money in an attempt to influence the referee. When workers choose to send money, we investigate whether these bribes distort the referee's judgment. We also vary whether the referee can keep both bribes or only the winner's bribe. This allows us to test whether the distortion is driven by reciprocity or greed, because whenever the referee is able to keep both bribes regardless of her decision, greed cannot influence her choice and only social preferences may drive behavior.

In addition to the laboratory experiments conducted in the United States (San Diego), we also report the results of an extra-laboratory experiment from a market in the city of Shillong in India. The data from a market in a country where corruption is more spread allow us to investigate whether our results generalize beyond the scope of the original laboratory experiments.

[1] Previous literature has shown that choices and judgment can be distorted by social pressure (Asch, 1954; Bond and Smith, 1996), self-serving biases (Babcock et al., 1995; Lord et al., 1979; Kunda, 1990; Haisley and Weber, 2010, see also Bazerman et al., 2002), or through a conflict of interest (Cain et al., 2005; Moore and Loewenstein, 2004).

II. The Bribery Game and Research Questions

II.A. The Bribery Game

Our bribery game involves three players: two workers and a referee. The workers compete against each other on a task and the referee is asked to determine a winner. The worker who wins gets a prize of *p*, and the other worker receives nothing. Additionally, workers can send a bribe ($b_i \in [0, \frac{1}{2}p]$) to the referee, with only integer amounts allowed.

Our main identification relies on two versions of the basic game. In treatment KeepWinner, referees keep the bribe of the winning worker; the other worker's bribe is returned. The referee's monetary payoff maximizing strategy is then to choose the worker who submits the higher bribe. Assuming the referee chooses this strategy, and given the restriction that $b_i \leq 0.5p$, the workers' monetary payoff maximizing strategy is to bribe $1 more than the other worker. This strategy results in a unique Nash equilibrium in which both workers bribe the maximum $b_i = 0.5p$. The referee's equilibrium payoff under these assumptions is $\Pi_R = b_{i*}$, where $i*$ is the winner of the round. The monetary payoff of each worker i is given by

$$\Pi_i = \begin{Bmatrix} -b_i + p & \text{if i wins} \\ 0 & \text{if i loses} \end{Bmatrix}$$

In the second treatment we study ("KeepBoth"), the referee keeps both bribes, and the payoff for the referee in each given round is therefore given by $\Pi_R = b_i + b_j$. The monetary payoff of each worker i is given by

$$\Pi_i = \begin{Bmatrix} -b_i + p & \text{if i wins} \\ -b_i & \text{if i loses} \end{Bmatrix}$$

A monetary payoff maximizing referee will then be financially indifferent between both workers, irrespective of the bribes. The workers' payoff-maximizing strategy depends on their beliefs regarding how the referee will reward bribes. In particular, whenever a worker's belief that referees will select the worker with the higher bribe as the winner (*pRef*) is low enough, the best response will be not to bribe. For the parameters used in the experiment, when $pRef<.6$, workers' optimal strategy is not to bribe. For $.6\leq pRef<1$, a mixed equilibrium exists in which workers bribe with some probability. For $pRef=1$, a pure strategy equilibrium exists in which both workers bribe the maximum. For a more detailed analysis, see Appendix A8.

This game allows us to study whether bribes induce referees to distort the true ranking between workers, resulting in an allocation of the prize based on bribes rather than performance, and to investigate which motives drive distortion. A feature of this game is that the referee's payoff depends only on the bribes and not on worker performance. This reflects real-world situations where decision-makers receive a fixed wage and do not have a direct individual incentive to maintain quality standards. This feature of our design also allows us to cleanly disentangle greed, reciprocity, and moral costs of distortion, as discussed below.

Finally, in order to answer our main research question and isolate the effect of distortion, we deliberately did not introduce other elements often associated with bribery, such as monitoring, punishment, and third-party externalities. Future research could use our bribery game to incorporate these additional features.

II.B. Research Questions

The two treatments described above help us in answering an important question regarding bribery: Does bribing distorts the referee's judgment, and if so, how this distortion interacts with the treatments. In our game, the referee is asked to choose the winner based on the workers' performance on the task. Basing the decision instead on the size of the bribe leads to a distortion of the true ranking between workers. If individuals have some moral costs (e.g., lying costs) associated with distorting their judgment, they may choose to reward the better performer, and bribery would not influence their behavior.[2]

Two important forces could explain why bribery affects judgment: reciprocity and greed. According to the reciprocity, or gift exchange, hypothesis (e.g., Abbink et al., 2002; Malmendier and Schmidt, 2012), if a worker sends money to the referee, the referee might want to reciprocate the favor by choosing to reward the worker who sent her (more) money. In this case, referees will choose the worker who sent the higher bribe, because they want to reciprocate the worker who was nicer to them, and not just because that bribe provides them with more money. In contrast to the gift-exchange explanation, greed implies that referees choose the worker who bribes more only when doing so benefits them financially.

It is important to note that when we refer to reciprocity and gift exchange in the paper, we refer exclusively to the non-strategic one, and not to repeated game consideration that include profit maximization motives (i.e., greed).

Comparing behavior in treatment KeepWinner with behavior in treatment KeepBoth allows us to test whether moral costs, gift exchange, or greed drive behavior. In treatment KeepWinner, a selfish payoff-maximizing referee would base her decision solely on the size of the bribes. Similarly, a referee who cares only about reciprocity will also choose the worker who sent the higher bribe. In treatment KeepBoth, the referee's choice of a winner does not affect her payment. Hence, a selfish payoff-maximizing referee will be indifferent between workers.

[2] See for example Gneezy, (2005), Dreber and Johannesson (2008), Sutter (2009); see also Cappelen, Sørensen, and Tungodden (2013), Erat (2013), and Lightle (2013) for investigations of the factors that affect moral costs in deception behavior, and Belot and Schröder (2013), who examine the relationship between payment schemes and lying and theft in a principal agent setting.

A reciprocal referee will still reward the higher bribe even when doing so does not affect her payoff.

If reciprocity drives the distortion of judgment, the distortionary effects of bribery will be similar in both treatments. If greed drives behavior, referees will distort their judgment in treatment KeepWinner, but in general not in treatment KeepBoth.

If referees also have moral costs, and—all else equal—prefer an allocation that does not require them to distort their judgment, these moral costs may outweigh greed and/or reciprocity concerns, preventing distortion of judgment in both treatments. However, if the power of greed or reciprocity in our experiment is large enough to outweigh moral costs, we expect to see some distortion. By comparing the two treatments, assuming moral costs do not change, we can rank the importance of greed and reciprocity.

A feature of our design is that although treatment is randomized and workers are randomly paired within sessions, bribes are not determined at random. To analyze referees' behavior and make treatment comparisons, referees in the two treatments must face similar combinations of bribes. That is, the distribution of the difference between the bribes the referees receive must be similar across both treatments. We will explore whether this is the case in the results section, where we also discuss how other possible differences in bribing behavior across the two treatments may affect referees' decisions.

II.C. Additional Treatments

Other than the two main versions of the basic game, we ran five additional treatments to provide additional support for our findings and rule out alternative explanations.

Rejecting Bribes.—First, we take into account that, in many cases, the person being bribed (the referee in our game) may choose to reject the bribe. We consider the effect of such an option in treatments KeepWinnerReject and KeepBothReject. These treatments are identical to treatment KeepWinner and treatment KeepBoth, respectively, except that referees also have the option to reject both bribes. Honest behavior may imply choosing a worker but rejecting his bribe. Adding the ability to reject even the winning bribe allows us to investigate how this option affects behavior. Note that equilibrium predictions are unaffected, since no payoff-maximizing referee will ever reject a bribe.

Inequity Aversion.—The third additional treatment involves a higher wage for the referee (treatment HighWage). This treatment is similar to treatment KeepWinnerReject, except that referees receive a higher show-up fee ($20 instead of $5). Changing the referee's wage does not affect equilibrium predictions. This treatment allows us to test whether some sort of inequality preferences (see, e.g., Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000) can explain our results. In treatment KeepWinner (and in all other treatments), the referee starts with less money than the workers ($5 vs. $10). Therefore, accepting the higher bribe will *decrease* inequity by making the referee's income more similar to both the winner's income and the loser's income. By contrast, in treatment HighWage, the referee starts with more money than the workers ($20 vs. $10). Therefore, accepting the higher bribe will now *increase* inequity. Thus, inequity aversion would predict referees would be less likely to let the higher bribe win in treatment HighWage than in treatment KeepWinner. In this treatment, we also included the possibility of rejecting bribes in order to give referees the possibility of decreasing inequity by choosing a winner without keeping either of the two bribes.

The Importance of Moral Costs of Distortion.—In two additional treatments, we investigate the importance of moral costs that may arise when referees need to distort their judgment to award the prize to the worker who send the higher bribe. The fourth additional treatment, Treatment NoTask, is identical to treatment KeepBoth, except that workers no longer compete on a task. In this case, choosing the higher bribe does not require the referee to distort judgment, as there is no better performer. Hence, choosing the higher bribe in this context does not generate a moral cost. By comparing this treatment to treatment KeepBoth, we can therefore test whether moral costs of distortion are important, and whether in the absence of such distortion, gift exchange can account for the results. This treatment is more closely related to the existing bribery games in the literature. As discussed above, these studies do not capture the distortionary effect of bribes that our design introduces.

The fifth additional treatment involves a variation of treatment KeepWinner in which workers compete on a different task (treatment Objective). As we discuss in detail below, the main treatments of this paper use a subjective task. In treatment Objective, we replace this task with a more objective one. When evaluating a subjective task, distorting judgment could be easier because referees may convince themselves that the worker with the higher bribe is also the better performer. If the task is more objective, convincing oneself that the worker with the higher bribe is the better performer could be more difficult, and as a result the moral costs associated with distortion of judgment may be higher. Thus, similar to treatment NoTask, treatment Objective tests the importance of moral costs, but it does so by increasing rather than decreasing them. Removing or changing the nature of the task does not affect equilibrium predictions.

III. Experimental design

III.A. Task

In all treatments except NoTask and Objective, we chose a task that involves creativity and for which the evaluation is not fully objective but depends partly on the referee's subjective taste. In particular, we asked workers to write a joke either about economists (round 1) or psychologists (round 2). All instructions can be found in Appendix B.

We chose to use a subjective task because in many real-life situations in which bribery is relevant, decision makers cannot exclusively rely on objective criteria when deciding how to allocate resources. Procurement auctions are an example of these situations, because the decision to award a procurement contract to a certain supplier is based on both objective parameters (e.g., price, completion time), which can easily be observed, and partially subjective ones (e.g., esthetics), which are left to the auctioneer's discretion (Burguet and Che, 2004).

The task of judging jokes incorporates both the subjective and the objective component. In terms of the subjective component, humor is at least partially a matter of taste, so that for relatively similar jokes, different referees may have different opinions about which joke is the better one. However, when jokes differ enough in quality, one of them can also be objectively regarded as the better joke. We will see below that referees were capable of selecting the (objectively) better joke in such cases. For an overview of some of the jokes written by participants, see Appendix C.

By contrast, in treatment Objective, workers were asked to work on a variation of the Stroop Task (Stroop, 1935). Each participant was shown a sequence of color words (e.g., blue, red, yellow) one after the other and was asked to identify the ink color of each word. We chose to use a congruent version of the task, meaning the color word and its ink color were compatible (e.g., blue was always written in blue letters). Participants were informed that their final score was equal to the number of words successfully identified, and were asked to complete as many words as possible.

Upon completion of the task, their final score was graphically represented on a score sheet in which every successfully identified word was represented by a dot (see Appendix B4 for the instructions and a sample score sheet). This procedure meant referees could count the dots to objectively determine which worker performed better, but had some moral wiggle room in that they needed to expend some effort to objectively evaluate workers' performance. We introduced the score sheet both to add some moral wiggle room and to give referees a non-trivial task to perform.

III.B. Procedure

We conducted the experiment in the laboratory of the Rady School of Management at the University of California San Diego with a total of 333 participants. Participants were recruited using standard recruitment procedures at the laboratory via an online experimental registration system. All UCSD students are able to register for this system to participate in laboratory experiments.

For each session, we invited 10 participants to the laboratory. Every participant from the subject pool was notified about the sessions and was eligible to participate. Each session consisted of exactly six participants, and therefore any time more than six people showed up, we randomly selected six and dismissed the extra participants after paying them a $5 show-up fee.

Each session lasted approximately 50 minutes. Upon being selected to participate, participants were randomly assigned to a computer station and were asked to follow the instructions on the screen. Participants were anonymously matched in groups of three, and each of them was either assigned to the role of workers (called participant A and B in the experiment) or the referee. We then moved the referees to separate rooms (one room for each referee), where they received the remainder of their instructions. Workers continued reading their instructions in the main lab. Neither workers nor referees knew which of the other participants were matched with them.

We then informed participants (except those in the NoTask treatment) about the task and the referee's role in determining the winner. In treatment NoTask, we informed participants about the referee's role in determining the winner but did not ask them to complete any task. In all cases, neither the workers nor the referees were yet informed about the workers' opportunity to send money to the referee.

On their desks, workers had an envelope with their \$10 show-up fee, in \$1 bills. Each referee had an envelope with a \$5 show-up fee in all treatments except treatment HighWage, in which the referee received an envelope with a \$20 show-up fee. The information about the other participants' initial show-up fees was made common knowledge.

After all workers had read their instructions and completed some attention questions, they learned the topic of the jokes for the first round ("economists") and had 10 minutes to type a joke (in the NoTask treatment, workers were told to wait 10 minutes; in the Objective treatment they had 5 minutes to work on the Stroop task). The experimenters then printed and returned each joke (or score sheet in the Objective treatment) to the workers. Workers received only their own joke or score sheet, and were not informed about the jokes or scores of the other workers in the experiment. While the experimenters were printing the jokes, we asked workers to state their expected likelihood of having a better joke than their opponent ("What do you believe is the probability that you will have a better joke than your opponent?").

The workers then received a second set of instructions on the screen, which notified them of the opportunity to send money to the referee. In particular, workers were asked to put the printed copy of their joke (or score sheet in treatment Objective) in a large envelope labeled with their participation ID, and were given the opportunity to add up to $5 of their show-up fee to the envelope. Meanwhile, the referees also received a second set of instructions telling them about the possibility of workers sending them money.

After all workers had prepared their envelopes, an experimenter collected them, privately recorded the monetary content of each, and then gave the envelopes to the referees. Upon receiving the envelopes, each referee had five minutes to rate the quality of the workers' jokes on a scale of 0 to 10 (except in treatment NoTask and treatment Objective), and to place a winner card and a loser card in the winner and loser's envelope, respectively. After five minutes, the workers returned the envelopes to the experimenter, who then recorded the referees' decisions.

In treatments KeepWinner and Objective, the referee could keep only the winner's monetary transfer and had to return the loser's money by putting it back in the envelope. In treatments KeepWinnerReject and HighWage, the referee had to return any money received by the loser, but was asked to decide whether to keep the winner's money or return both bribes. In treatments KeepBoth and NoTask, the referee kept all the money sent by both workers, whereas in treatment KeepBothReject, the referee also had the option to return both the winner and the loser's money. Table 1 summarizes the experimental treatments. Note that we have 60 participants (20 groups) in the two main treatments, KeepWinner and KeepBoth, and in treatment HighWage. For treatment Objective, we have 63 participants (21 groups).[3] For the four remaining treatments, we have 30 participants (10 groups) in each treatment.

TABLE 1—THE EXPERIMENTAL TREATMENTS

	Which bribe does the referee keep?	Task	Participants	Ref. show-up fee
KeepWinner	Only winner's	Jokes	60	$5
KeepBoth	Both	Jokes	60	$5
KeepWinnerReject	Chooses whether to keep winner's	Jokes	30	$5
KeepBothReject	Chooses whether to keep both	Jokes	30	$5
HighWage	Chooses whether to keep winner's	Jokes	60	$20
NoTask	Both	No	30	$5
Objective	Only winner's	Objective	63	$5

The experiment consisted of two rounds with the same matching of participants. To prevent referees from reciprocating the largest bribe in round 1 for strategic reasons, no feedback was provided between rounds. Workers started the second round while the referees were still evaluating their first round. The procedure for round 2 was identical to that of round 1, apart from the topic of the joke.

After the second round, both workers and referees were asked to complete a survey of basic demographic information. The referees were then paid and left the experiment, and workers received back the envelopes for rounds 1 and 2. Each envelope contained either a winner or a loser card indicating the referees' decision. For treatments KeepWinner, KeepWinnerReject, HighWage, and Objective, the envelope with the loser card also contained any money sent to the referee by the worker who lost. For treatments KeepWinnerReject,

[3] In treatment Objective, we had to discard one group because the referee did not follow the instructions and rejected both bribes even though he was instructed not to do so. Therefore, we ran one additional session in order to have at least 20 groups.

KeepBothReject, and HighWage, both envelopes could also contain money returned by the referee if the referee decided to reject both bribes. Workers were then paid $10 for each winner card they had and left the experiment.

III.C. Joke quality

After the experiment was completed, we organized additional sessions in which participants from the same subject pool, who had not previously participated in the experiment, evaluated the quality of several pairs of jokes. The jokes were evaluated by a total of 752 raters who, for each pair of jokes, had to evaluate the quality of each joke (on a scale from 0 to 10) and had to determine which joke was funnier. Raters were shown the same pairs of jokes the referees evaluated during the experiment, without being informed about the bribes sent by the workers. This procedure provides us with a more objective measure of joke quality, which is not biased by the presence of bribery. Each independent rater evaluated up to six pairs of jokes, chosen at random by an electronic randomizer among all the possible pairs of jokes. Each pair of jokes was evaluated by an average of 22 independent raters. The full instructions are in Appendix B3.

IV. Results

Table 2 presents some descriptive statistics on our sample. As the table shows, the treatments are balanced with respect to demographics. Joke quality and confidence levels are also not statistically different between treatments and rounds (Bonferroni or Holm-Bonferroni correction for multiple hypothesis testing).

TABLE 2—DESCRIPTIVE STATISTICS

	Overall	**Keep Winner**	**Keep Both**	**KW Reject**	**KB Rejec**	**High Wage**	**No Task**	**Objective**
Joke Quality	3.62	3.57	3.50	3.47	3.81	3.78		
(Round 1)	(1.18)	(1.10)	(1.17)	(1.28)	(1.43)	(1.11)		
Joke Quality	3.54	3.73	3.27	3.39	3.89	3.51		
(Round 2)	(1.27)	(1.32)	(1.43)	(1.45)	(1.12)	(.98)		
Objective Score	176							176
(Round 1)	(15)							(15)
Objective Score	179							179
(Round 2)	(17)							(17)
Worker Confidence	.51	.48	.53	.49	.55	.41		.62
(Round 1)	(.25)	(.28)	(.27)	(.26)	(.24)	(.24)		(.19)
Worker Confidence	.49	.44	.56	.41	.53	.39		.59
(Round 2)	(.25)	(.28)	(.26)	(.30)	(.28)	(.15)		(.22)
Psychology	.14	.13	.12	.13	.03	.18	.17	.14
	(.34)	(.33)	(.32)	(.35)	(.18)	(.39)	(.38)	(.35)
Economics	.24	.20	.18	.23	.43	.25	.37	.16
	(.43)	(.40)	(.39)	(.43)	(.50)	(.44)	(.49)	(.37)
Other Social Science	.08	.07	.07	.03	.07	.08	.10	.10
	(.26)	(.25)	(.25)	(.18)	(.26)	(.28)	(.31)	(.29)
Biology/Chemistry	.26	.35	.25	.43	.23	.15	.23	.25
	(.44)	(.48)	(.44)	(.50)	(.43)	(.36)	(.43)	(.44)
Engineering/Science	.20	.20	.22	.17	.17	.22	.07	.24
	(.40)	(.40)	(.42)	(.38)	(.38)	(.42)	(.25)	(.43)
Humanities Major	.09	.05	.15	.00	.07	.10	.07	.11
	(.28)	(.22)	(.36)	(.00)	(.25)	(.30)	(.25)	(.32)
Undeclared Major	.01	.00	.02	.00	.00	.02	.00	.00
	(.08)	(.00)	(.13)	(.00)	(.00)	(.13)	(.00	(.00)
Asian Ethnicity	.72	.63	.82	.83	.70	.72	.60	.71
	(.45)	(.49)	(.39)	(.38)	(.47)	(.45)	(.50)	(.46)
Female	.56	.55	.60	.50	.47	.57	.57	.57
	(.50)	(.50)	(.49)	(.51)	(.51)	(.50)	(.50)	(.50)
Nonnative speaker	.16	.15	.18	.13	.23	.15	.20	.13
	(.37)	(.36)	(.39)	(.34)	(.43)	(.36)	(.40)	(.33)
Age	20.7	21.1	20.5	20.6	20.3	20.6	21.0	20.7
	(1.92)	(2.62)	(1.46)	(1.65)	(2.15)	(1.81)	(1.30)	(1.89)
Observations	333	60	60	30	30	60	30	63

Notes: Descriptive statistics. Joke quality is the average rating of the joke by the independent raters. Objective performance is the score on the objective task for treatment Objective. Confidence is the worker's confidence in having a better joke or performance than the other worker. The remaining variables are dummies for the respective majors, Asian participants, females, and nonnative speakers, and a continuous variable for age, respectively. Among the nonnative speakers, 12% are Chinese native speakers, 2% are Spanish native speakers, and the remainder report different languages.

In the remainder of this section, we will use both parametric and non-parametric tests to test for differences between treatments. Whenever we analyze worker behavior, we use one worker as one independent observation; whenever we analyze referee behavior, we use one referee as one independent observation. For non-parametric tests involving data from both rounds, we therefore take the average over both rounds as the unit of observation. We first analyze worker and referee behavior in the two main treatments, KeepWinner and KeepBoth. We then discuss the additional treatments to investigate the robustness of our results and address potential alternative explanations.

IV.A. Worker Behavior

The main goal of the paper is to investigate whether and why bribes distort choices. To do so, we first need to verify that workers were willing to bribe, and that there was sufficient variation in bribe sizes. Figure 1 shows the distribution of bribes in the KeepWinner treatment for both rounds. The first thing to note is that many workers bribed and there was considerable variation in bribe amounts: 41% of bribes were at the maximum \$5 and a further 33% of bribes were positive. In 26% of the cases, workers elected not to send a bribe. Overall, the average bribe was \$2.80.

As Figure 1 also shows, the workers in treatment KeepBoth bribed less than the workers in treatment KeepWinner. Overall the average bribe in this treatment was \$0.90 and no bribe was sent in 66% of cases. The difference in the distribution of bribes between the KeepWinner and the KeepBoth treatments is significant ($p<.001$, Mann-Whitney). In Appendix A1, we further examine whether worker-level characteristics are predictive of bribe size. Over all treatments, we find that non-native speakers, older participants, and men send higher bribes, whereas social science majors (not including economists) send smaller bribes. Importantly, we also show that bribe size is not correlated with joke quality ($p=.76$).

Table 3 reports the distribution of bribes per treatment and the average bribes per round for all treatments. On average, in the KeepWinner treatment, bribes did not change between rounds: the average bribe was $2.83 in the first round and $2.78 in the second round (the difference is not statistically significant). In treatment KeepBoth, the average bribe was $1.13 in round one and $0.68 in round two and this difference is marginally significant (p=.056, Wilcoxon signed-rank test).[4] Additionally, workers who bribed more in round 1 were also likely to bribe more in round 2 (r=.52, p=.005 in KeepWinner; r=.60, p<.001 in KeepBoth).

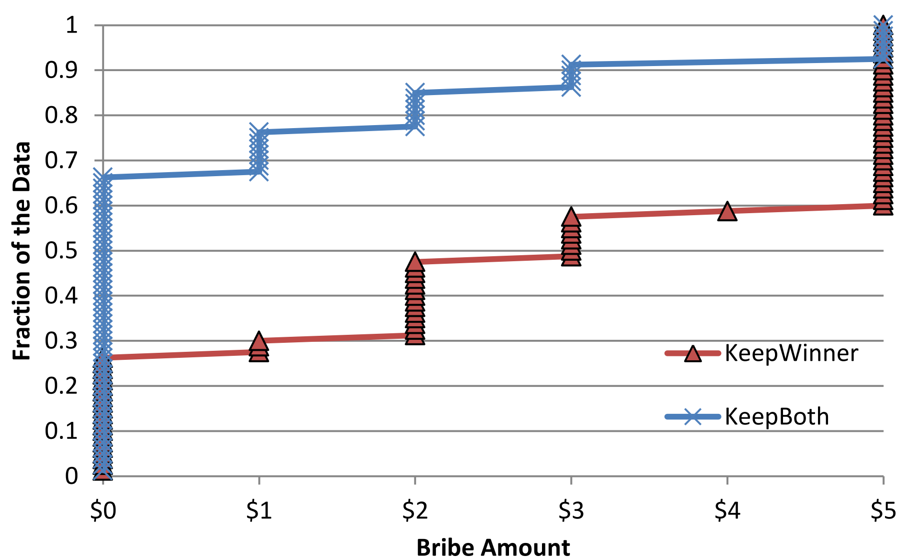

FIGURE 1. CDF OF BRIBES FOR KEEPWINNER AND KEEPBOTH.

Note: Empirical cumulative distribution function of all bribes sent in treatments KeepWinner and KeepBoth.

This analysis shows that there is a treatment difference in worker behavior in terms of bribes. However, in terms of referee behavior, we are interested in

[4] However, this effect is only driven by a few observations. 30 of the 40 workers bribed exactly the same in the second round, two people bribed more, and the remaining eight people bribed less.

analyzing cases where referees receive two bribes of a different size. Specifically, we will analyze whether referees select the better joke or higher bribe as the winner, and whether the likelihood of doing so depends on the absolute *difference* between the two bribes (or difference in quality of the two jokes).

It is therefore important to investigate whether the difference in bribes is similar across treatments. The results are presented in the lower panel of Table 3. We focus on cases where referees received two different bribes, as these are the only cases where the size of the bribe could affect which worker they selected as the winner. We find that for these cases the average difference in bribes is similar in treatment KeepWinner (\$3.03) and treatment KeepBoth (\$2.64 p=.482, Mann-Whitney). Similarly, we find no difference in the distribution of the difference between the workers' bribes (p=1, Kolmogorov-Smirnov). Hence, despite bribes being larger in treatment KeepWinner overall, referees who received bribes of different size still faced similar financial tradeoffs in both treatments.

However, referees in treatment KeepBoth were also more likely to receive two identical bribes, less likely to face two positive bribes, and less likely to receive two large bribes. We control for the first aspect in our regression analysis by examining the cases in which bribes differ separately from the cases where bribes are identical. We discuss the latter two differences in Appendix A2 and show that neither of them affects our main results. In particular, the results suggest referees behaved similarly irrespective of the absolute size or of whether they received one or two positive bribes.

TABLE 3—BRIBES ACROSS TREATMENTS

	Overall	Keep Winner	Keep Both	KW Reject	KB Reject	High Wage	No Task	Objective
				Bribe Frequencies (Both Rounds)				
Bribe=0	41%	26%	66%	13%	58%	53%	25%	36%
Bribe=1	7%	4%	10%	3%	18%	5%	5%	6%
Bribe=2	12%	18%	9%	18%	10%	9%	15%	12%
Bribe=3	8%	10%	6%	13%	5%	4%	18%	10%
Bribe=4	2%	1%	0	3%	3%	1%	15%	1%
Bribe=5	28%	41%	9%	53%	8%	29%	23%	36%
				Bribe Size (Mean)				
Round 1	2.21	2.83	1.13	3.35	1.25	2.08	2.95	2.36
	(2.12)	(2.07)	(1.76)	(1.92)	(1.65)	(2.34)	(1.79)	(2.15)
Round 2	1.97	2.78	.68	3.60	.75	1.58	2.25	2.48
	(2.10)	(2.13)	(1.31)	(1.79)	(1.41)	(2.07)	(1.97)	(2.22)
Both Rounds	2.10	2.80	.90	3.48	1.00	1.83	2.60	2.42
	(2.11)	(2.09)	(1.56)	(1.84)	(1.45)	(2.20)	(1.89)	(2.17)
				Average Difference in Bribes (excluding equal bribes)				
Round 1	3.14	3.07	2.92	2.14	2.83	4.21	2.63	3.40
	(1.47)	(1.59)	(1.56	(.69)	(1.72)	(1.31)	(1.30)	(1.24)
Round 2	2.90	3.00	2.30	3.00	2.60	3.42	3.00	3.29
	(1.51)	(1.36)	(1.64)	(1.07)	(1.82)	(1.73)	(1.41)	(1.54)
Both Rounds	3.03	3.03	2.64	2.60	2.73	3.85	2.82	3.34
	(1.49)	(1.45)	(1.59)	(.99)	(1.68)	(1.54)	(1.33)	(1.37)
N: per round	222	40	40	20	20	40	20	42
N: both rounds	444	80	80	40	40	80	40	84

Notes: The table gives the relative frequency of bribes of different sizes in the upper panel. The middle panel displays average bribe size (over all workers) separately as well as jointly for each treatment and round. The lower panel displays the average difference in bribes (over all workers) separately as well as jointly for each treatment and round. Average bribes are computed using all bribes, including zeros. Average differences in bribes are computed by subtracting the highest bribe from the lowest bribe in a given pair of bribes and are based only on observations in which the two bribes were not identical. The numbers in brackets are standard deviations.

Joke Quality.—In order to investigate referee behavior and make treatment comparisons, it is also important that the quality of the jokes is similar across treatments. To investigate whether this was the case, we use the evaluation provided by the independent raters as an unbiased measure of joke quality.

For our analysis, we are interested in investigating whether the referee selects the better joke as the winner. For this purpose, we will focus on two measures of quality: the difference in average rating between the two jokes in a pair and the fraction of raters that, for a given pair, chose the same joke as the better one (i.e., the degree of agreement across raters). The two measures are highly correlated ($r=.92$, $p<.001$ for all joke treatments combined). Both measures are similarly distributed in the two main treatments ($p=.436$ and $p=.560$ for KeepWinner and KeepBoth respectively, Kolmogorov-Smirnov). As with bribes, in our analysis, we will look at quality *differences* between the jokes written by the two workers in a given pair. The distribution of differences in quality also does not differ between KeepWinner and KeepBoth ($p=.739$, Kolmogorov-Smirnov).

In Appendix A1, we examine predictors of quality and find that non-native speakers, more confident participants and younger participants wrote better jokes. Importantly, we find no evidence that bribes are predicted by joke quality (or the performance on the objective task), suggesting that overall the quality of workers' performance did not affect their bribing behavior.

IV.B. Does Bribery Distort the Referee's Judgment?

In this section, we examine whether and why bribery distorts referees' judgment, using both non-parametric tests and regressions. For the non-parametric tests, we investigate whether the worker with the higher bribe in the pair won the prize. We also investigate whether the worker with the better joke in the pair won the prize.

However, because joke quality is subjective, it is not enough for one joke to have a slightly higher quality on either of the two quality measures. Instead, we need to know whether one joke is *significantly* better than the other. For this purpose, we consider all joke pairs for which at least 65.1% of the independent raters agreed on the winner. With this threshold, the fraction of independent raters who selected a given joke over the other is significantly different from chance (i.e., 50%) at the 10% level (z=1.28, p=.1, test of proportions for our minimum of 18 independent raters). By this threshold, 63% of joke pairs have a significantly better joke.

In the remaining pairs, the quality of the two jokes was too similar to be statistically distinguishable. In such cases, picking one joke over the other did not constitute a big distortion. Whenever we refer to better-quality jokes in subsequent non-parametric tests, we will only use jokes that are sufficiently different by this criterion. Note, however, that in our regression analysis we will use all the observations, including cases where bribes and/or quality were similar. Appendix A3 presents more details on this threshold and also shows that for a threshold of 69.4%, which corresponds to jokes being significantly different at the 5% level, the results are similar.

For the Objective treatment, we use participants' actual scores as the performance measure. Similar to the other treatments, we omit the 37% least *distinguishable performance* pairs whenever we refer to better-quality performers. In practice, this approach means we only look at pairs in which the difference in performance was at least 11 words.

The KeepWinner Treatment.—Did bribing result in a distortion of the referees' judgment? In 86% of the cases, referees in the KeepWinner treatment chose the worker who offered the higher bribe as the winner. This number is significantly larger than chance (p=.001, Wilcoxon). By contrast, as Figure 2 shows, the better joke (as judged by the independent raters) won only 57% of the time, which is not significantly different from chance (p=.564, Wilcoxon). Thus, these results suggest that bribery distorted referees' judgment, because they chose the worker who paid them more, not the one who wrote the funnier joke.

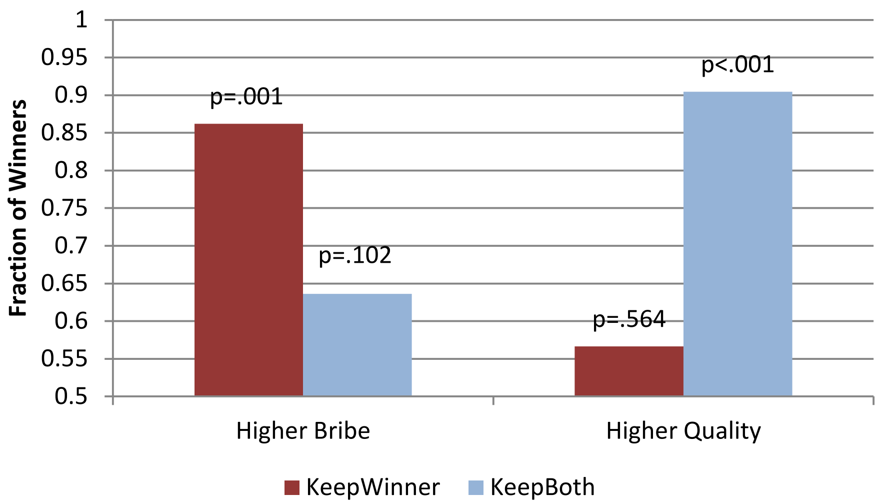

FIGURE 2. WIN CHANCE WHEN HAVING THE HIGHER BRIBE OR BETTER RATING FOR KEEPWINNER AND KEEPBOTH

Notes: The p-values are calculated using a Wilcoxon signed rank test that tests if the fraction is significantly larger than .5. Workers are classified as having a higher quality when at least 65.1% of independent raters agree their joke is better.

We further investigate the effect of bribes and quality using OLS. In the regression, we examine how differences in joke quality (as determined by the independent raters) and bribes between the two workers affect the referee's decision. For a given worker, the regression tells us how an increase in her bribe or joke quality relative to the other worker affects her probability of winning. The more referees care about quality as opposed to bribes, the more beneficial having a better joke should be.

On a more methodological level, because our independent variables are differences between the two workers within a given pair, the observations for the two workers are always the exact inverse of each other. Hence, for the regression, we randomly select one worker per round. We use the same random sample throughout the analysis. In Appendix A6, we show that the reported results do not depend on the particular random combination of workers selected for the regression. Randomly selecting a worker also implies that selected workers on average win approximately 50% of the time; as a result, we do not report the constant in the regression table.

Further, to facilitate comparisons between quality and bribe coefficients, we standardize all independent variables, such that the coefficients represent the effect of a one-standard-deviation increase in the independent variable. For quality, we estimate separate coefficients for cases in which the two bribes are identical and cases in which they differ. The latter coefficient is of particular interest because it allows us to examine the effect of quality when referees could also be influenced by bribes. The former coefficient instead allows us to see whether quality affect the likelihood that a given worker wins the prize when referees have no incentive to distort their judgment.

Finally, we compute the p-values reported in the regression tables using a wild bootstrap procedure. Cameron, Gelbach, and Miller (2008) show that for a small number of clusters and/or small cluster sizes, this approach leads to more accurate (and more conservative) p-values than alternative techniques. In Appendix A5, we also redo all regressions of Table 4, using several alternative techniques, including clustered standard errors and non-parametric bootstraps, with very similar results.

Column (1) of Table 4 presents the results. The coefficient for bribes is large, positive, and statistically significant. Indeed, a one-standard-deviation increase in a given worker's bribe (relative to the bribe of the other worker) increases her likelihood of winning the prize by 31 percentage points. By contrast, the coefficient for quality when bribes differ is small and not statistically significant. Thus, the regression results confirm that bribes, not quality, influenced referees in treatment KeepWinner.

The regression results also show that when bribes are identical and can therefore not distort behavior, an increase in the quality of a given worker's joke (relative to the other worker) does significantly increase her likelihood of winning. This finding shows that despite the subjective nature of the task, referees were indeed capable of identifying the higher-quality joke in the absence of distortionary incentives.

TABLE 4—OLS REGRESSIONS FOR REFEREES IN KEEPWINNER AND KEEPBOTH

Dependent Variable:	Winner (1=Yes)		
	(1)	(2)	(3)
Bribe Difference	.308***	.086	.274***
	(.000)	(.140)	(.000)
Quality Difference (bribes differ)	.014	.262**	.015
	(.762)	(.010)	(.762)
Quality Difference (bribes identical)	.336**	.275**	.255**
	(.020)	(.022)	(.020)
$D_{KeepBoth}$			.008
			(.980)
Bribe Difference X $D_{KeepBoth}$			-.173**
			(.032)
Quality Difference X $D_{KeepBoth}$ (bribes differ)			.222**
			(.014)
Quality Difference X $D_{KeepBoth}$ (bribes identical)			.150
			(.156)
Treatment	KeepWinner	KeepBoth	KeepWinner KeepBoth
Selected Workers	Random	Random	Random
Observations	40	40	80
Clusters	20	20	40

Notes: OLS estimates (p-values). The dependent variable is a dummy that specifies whether the selected worker was selected as the winner. Quality Difference is the difference between the quality of the joke (i.e., the average rating by the independent raters) of the selected worker and the quality of the joke of the other worker in the pair. Bribe Difference is the difference between the bribe sent by the selected worker and the bribe sent by the other worker in the pair. $D_{KeepBoth}$ is a dummy that is equal to one for treatment KeepBoth, and zero otherwise. For column (3), the bribe variable and both quality variables are standardized using the respective variable's combined standard deviation over all included treatments. P-values are calculated using wild bootstraps. For each regression, we randomly select one worker per referee in each round. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

Finally, the results are similar if we perform this regression separately for each round; see Appendix A7. In Appendix A4, we additionally estimate similar regressions that incorporate only those pairs of workers who had different bribes or wrote sufficiently different jokes. Further, we present the results of alternative specifications that include separate coefficients for bribes depending on whether joke quality was similar or significantly different. The results of the additional analyses are in line with the results reported in Table 4.

The KeepBoth Treatment.—Figure 2 and column (2) of Table 4 give an overview of the referees' behavior in the KeepBoth treatment. In 64% of the cases, referees chose the worker who offered the higher bribe as a winner. This number is not significantly larger than chance ($p=.103$, Wilcoxon). By contrast, the better joke, as judged by our independent raters, won 90% of the time. This proportion is significantly larger than chance ($p<.001$, Wilcoxon). In other words, when the referees' payoff did not depend on the choice of winner, bribery did not distort judgment, and referees chose the worker who wrote the funnier joke.

The regression results are similar. Column (2) of Table 4 shows that an increase in the quality of a given worker's joke (relative to the other worker's joke) significantly increased her likelihood of winning, whereas a relative increase in the worker's bribe did not pay off. In addition, the effect of joke quality on the likelihood of winning was similar both when bribes were identical and when bribes differed. This finding confirms that referees chose the worker who wrote the funnier joke in this treatment.

Greed versus Reciprocity.—Figure 2 shows that having a higher bribe was more effective in the KeepWinner treatment (86% vs. 64%, $p=.048$; Mann-Whitney), whereas having a better joke was more effective in the KeepBoth treatment (90% vs. 57%, $p=.004$; Mann-Whitney). The latter effect is driven exclusively by cases for which bribes were different. When bribes were equal, referees *in both treatments* (80% for KeepWinner, 78% for KeepBoth) picked the better joke ($p=.871$; Mann-Whitney). When bribes were *unequal*, referees in KeepBoth selected the better joke 100% of the time, compared with only 45% in KeepWinner (p=.003, Mann-Whitney).

These findings are confirmed by the regression of column (3) in Table 4, where we included data from both treatments and interacted the quality and bribe variables with a dummy for treatment KeepBoth. The interaction terms confirm that when bribes are unequal, the effect of quality is significantly larger in treatment KeepBoth, whereas the effect of bribes is smaller. Further, in cases in which bribes were equal, the importance of quality was approximately the same in both treatments.[5]

In Appendix A2, we report additional analyses in which investigate the effect of absolute bribe size and explore whether receiving one or two bribes affected behavior. We find that referees behaved similarly irrespective of absolute size or the number of bribes received.

Another interesting piece of information from our data comes from referees' evaluations of the quality of the two jokes. In particular, we asked referees in all the joke treatments to rate the quality of both jokes on a scale from 0 to 10. This measure was not incentivized, so participants had no incentive to lie. Whereas in treatment KeepBoth, referees' evaluation of the jokes is highly correlated with the independent raters' quality measure (r=.513, p<.001, OLS), in treatment KeepWinner, this correlation is smaller and not significant (r=.147, p=.194, OLS). This treatment difference in the accuracy of referees' evaluations of quality is significant (p=.034, OLS). This suggests that referees in treatment KeepWinner may have tried to rationalize their choice ex post and that the bribes distorted their quality evaluations.

[5] The reason that the three non-interacted coefficients in column 3 of Table 4 are not exactly identical to the coefficients in column 1 is standardization. For column 1, we standardized all coefficients with respect to the standard deviation of the explanatory variables in KeepWinner, whereas for column 3 we used the standard deviation for the combined data of KeepWinner and KeepBoth.

Overall, the results show that referees awarded the prize to the worker with the higher bribe in treatment KeepWinner, but selected the one with the better joke in treatment KeepBoth. This finding is in line with the greed explanation of bribery. When referees are motivated by greed (treatment KeepWinner), they distort their judgment. However, when only reciprocity could lead referees to select the higher bribe (treatment KeepBoth), they instead select the better joke. This observation suggests greed is more important than moral costs, which is in turn more important than reciprocity.

C. Additional Treatments

In this section, we present the results for the six additional treatments. To facilitate comparisons across treatments, we pool the data from all treatments into a joint regression in Table 5. We use KeepWinner as the reference treatment and interact the bribe and quality difference variables with treatment dummies for all other treatments. This approach allows us to verify whether bribes and quality played a larger or smaller role than in treatment KeepWinner. The corresponding figures are presented in Appendix A9.

Rejecting Bribes.—The results of KeepWinnerReject and KeepBothReject suggest that allowing referees to reject bribes does not change their behavior. In the KeepWinnerReject treatment, referees chose not to keep the winner's bribe in only 10% of the cases. As a consequence, the fraction of winners with the higher bribe (100% vs 86%, p=.176, Mann-Whitney) or with the funnier joke (56.7% vs. 57.1%; p=.978, MW) are similar to those in treatment KeepWinner.

In treatment KeepBothReject, referees chose to reject both bribes in 10% of the cases. As a consequence, the results are similar to the results of treatment

KeepBoth: 86% of the referees awarded the prize to the better joke (vs. 90% in KeepBoth; p=.626, Mann-Whitney). Referees were less likely to award the prize to the worker with the higher bribe than in treatment KeepBoth (36% vs. 64%; p=.043, Mann-Whitney), though this effect disappears when controlling for quality (as in Table 5).

The regression results are presented in Table 5. For treatment KeepWinnerReject, the only difference relative to KeepWinner is that having a larger bribe than the opponent increases the likelihood of winning significantly more, which is due to the fact that referees let the higher bribe win 100% of the time in this treatment. For treatment KeepBothReject, the coefficients for bribes and quality are similar in size to those of treatment KeepBoth, and significantly different from those of treatment KeepWinner.

Overall, the results suggest that allowing referees to reject bribes did not affect their behavior. The treatment effects we find within the Reject treatments is very similar to the treatment effect we found in the main treatments. Regardless of whether they are able to reject the bribes, referees appear motivated by greed, not reciprocity. In Appendix A4, we redo our analysis from the previous section with the merged data from the reject treatments with the data from the main treatments. The analysis reveals that all coefficients estimates are similar, but more precisely estimated. As a result, the main treatment comparison for bribes and quality is significant at the 1% level.

Further, allowing referees to reject bribes did not affect worker behavior. There are no significant differences in average bribes for either of the two treatments (KeepWinnerReject=3.48 vs. KeepWinner=2.80, p=.188, Mann-Whitney; KeepBothReject=1.00 vs. KeepBoth=0.90, p=.343, Mann-Whitney).

TABLE 5—OLS REGRESSIONS FOR REFEREES IN ADDITIONAL TREATMENTS

Dependent Variable: Winner (1=Yes)	(1)
Bribe Difference	.296***
	(.002)
Bribe Difference X $D_{KeepBoth}$	-.187**
	(.032)
Bribe Difference X $D_{KeepWinnerReject}$	.187***
	(.006)
Bribe Difference X $D_{KeepBothReject}$	-.199
	(.106)
Bribe Difference X $D_{HighWage}$	-.084
	(.214)
Bribe Difference X D_{NoTask}	.059
	(.560)
Bribe Difference X $D_{Objective}$	-.226**
	(.042)
Quality Difference (equal bribes only)	.231**
	(.016)
Quality Difference (Bribes Equal) X $D_{KeepBoth}$	.137
	(.162)
Quality Difference (Bribes Equal) X $D_{KeepWinnerReject}$	-.243
	(.288)
Quality Difference (Bribes Equal) X $D_{KeepBothReject}$	-.149
	(.490)
Quality Difference (Bribes Equal) X $D_{HighWage}$	-.206
	(.232)
Quality Difference (Bribes Equal) X $D_{Objective}$	-.081
	(.624)
Quality Difference (Different Bribes)	.015
	(.792)
Quality Difference (Different Bribes) X $D_{KeepBoth}$	.216**
	(.016)
Quality Difference (Different Bribes) X $D_{KeepWinnerReject}$	-.045
	(.462)
Quality Difference (Different Bribes) X $D_{KeepBothReject}$	.463***
	(.006)
Quality Difference (Different Bribes) X $D_{HighWage}$	.023
	(.844)
Quality Difference (Different Bribes) X $D_{Objective}$	.302**
	(.014)
Selected Workers	Random
Treatment Dummies	Yes
Observations	222
Clusters	111

Notes: OLS estimates (p-values). P-values are computed using wild bootstraps. The '$D_{Treatment}$' variables are dummy variables for the respective treatments; KeepWinner serves as the reference treatment. The bribe variable and both quality variables are standardized using the respective variable's combined standard deviation over all treatments. We randomly select one worker per referee in each round. For other variable definitions, see the notes to Table 4.

*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

Inequity Aversion.—In KeepWinner, inequity aversion predicts that referees should accept the higher bribe, whereas in HighWage, it predicts the opposite. However, the higher bribe won 88% of the time, compared with 86% in

KeepWinner, and the better-rated joke won 44% of the time, compared with 57% in KeepWinner. Neither difference is significant in either non-parametric tests or in Table 5. Additionally, the referee chose not to keep the winner's bribe in only 7.5% of the cases. These results suggest referees in this treatment behaved similarly to referees in treatment KeepWinner, contrary to the prediction of inequity aversion. The only difference between the HighWage and the KeepWinner treatment relates to workers' behavior: their average bribe was significantly lower in HighWage than in KeepWinner ($1.80 versus $2.80, p=.020; Mann-Whitney).

The Importance of Moral Costs of Distortion.—Why are higher bribes ineffective in the KeepBoth treatment? One hypothesis is that the moral costs of distorting judgment are stronger than reciprocity, and hence referees will ignore the bribes. In treatment NoTask, we removed all moral costs of distortion by no longer asking referees to judge performance on a task. Hence, we expect reciprocity to become more important.

This is indeed what we find. Whereas the higher bribe won in 64% of the cases in treatment KeepBoth, in the NoTask treatment, this fraction significantly increases to 94% (p=.044, Mann Whitney). Moreover, 94% is also significantly larger than chance (p=.011, Wilcoxon), and not significantly different from KeepWinner (p=.614, Mann Whitney). The regression analysis of Table 5 confirms these results.

Looking at workers' behavior (Table 3), we find the average bribe in the NoTask treatment ($2.60) was higher than in KeepBoth (p<.0001, Mann-Whitney) and similar to KeepWinner (p=.697, Mann-Whitney). Additionally, in this treatment, the fraction of workers who did not send any bribe is lower than in KeepBoth (25% vs. 66%).

The comparison between the KeepBoth and the NoTask treatment provides further evidence that distorting judgment by rewarding the higher bribe presented referees with a moral cost, which previous bribery games did not focus on. Referees were happy to award the prize to the worker who sent them more money when rewarding them did not require a distortion of their judgment, but did not reciprocate the higher bribe when doing so requires them to award the prize to the worst performer. Thus, in the absence of moral costs, reciprocity guides referees' behavior. By contrast, in other treatments, the moral costs of distorting judgment seem stronger than the norm of reciprocity.

Similar to treatment NoTask, treatment Objective tests the importance of moral costs, but it does so by increasing rather than decreasing them. Having a better performance was more effective in the Objective treatment than in the KeepWinner treatment (83% vs. 57%, p=.023; Mann-Whitney), whereas it was equally effective as in the KeepBoth treatment (83% vs. 90%, p=.312; Mann-Whitney). Having a higher bribe was neither less effective than in KeepWinner (76% vs. 86%, p=.443; Mann-Whitney) nor more effective than in KeepBoth (76% vs. 64%, p=.299; Mann-Whitney).

The fact that both bribes and performance appear to be important may be the result of a positive correlation between differences in bribe size and differences in performance (r=.31; p=.082, OLS, one-sided.). The regression in Table 5 corrects for this correlation. The results show that having a higher bribe than the opponent affected the probability of winning significantly less than in the KeepWinner treatment. The bribe coefficient is similar to the bribe coefficient estimated for the KeepBoth treatment. Similarly, when bribes differed, having a better performance mattered significantly more than in the KeepWinner treatment; the coefficient is similar to the coefficient of the KeepBoth treatment. Overall, the results suggest that moral costs are higher in the Objective treatment, resulting in a larger emphasis on quality.

V. An Experiment in the Market in Shillong, India

The results of our experiment show that the mechanism by which bribes distort the referees' decisions is greed and not reciprocity. Here, we complement the laboratory evidence with evidence from an experiment in a different, more natural setting. Whereas the lab experiments allow us to sharply disentangle between the different mechanisms of bribery in a clean setting, the extra-laboratory experiment (Charness et al., 2013) allows us to investigate whether our results generalize to a population and environment that are more regularly exposed to bribery than UC San Diego students.

V.A. Experimental Design

We conducted the experiment at the market in the city of Shillong, in the state of Meghalaya in northeast India. Bribery and corruption are prevalent in India (Transparency International, 2014), and Meghalaya is thought to be among the most corrupt states in India (Transparency International India, 2008).

The goal of this experiment was to investigate the importance of reciprocity and greed in a different, more natural setting. For this purpose, we approached shoppers at the market and asked them to taste two different pineapples, each purchased from a different vendor, and tell us which of the two they thought tasted better. In addition to tasting the pineapples, participants (i.e., the shoppers) received monetary payments ("bribes") from the vendors. This set up allows us to investigate whether participants chose the higher quality product, or the product from the vendor who made them earn more money. As with the joke task in the lab experiment, we chose this task because selecting the tastier pineapple is partially based on the decision maker's subjective judgment.

The procedure of the experiment was as follows. First, we approached two fruit vendors in the market, A and B, and invited them to participate in the study.

They were both selling pineapples and their stands were not close to each other. We explained to the sellers that, if they participated, we would purchase some of their pineapples. We told the sellers that we would ask shoppers to taste both their pineapple and a pineapple from another seller in the market, and indicate which of the two was tastier. We told the sellers that every time a shopper recommended their pineapple, we would purchase an additional pineapple from their stand at a price of 60 rupees (approximately $1 at the time).

Each seller agreed to pay some money to each shopper who chose his pineapple (other than those in the control treatment, see below). In particular, Seller A agreed to pay 10 Rupees and Seller B agreed to pay 20 Rupees. Both sellers also agreed that in half of the cases, they would pay these amounts even if the shopper did not choose their pineapple.

We then bought several pineapples from each seller, and selected the two best pineapples from seller A (pineapples A1 and A2), and the two worst pineapples from seller B (pineapples B1 and B2). For this purpose, four experimenters tasted each of the pineapples we bought. We chose the four pineapples such that all four experimenters thought pineapple A1 was tastier than pineapple B1, and pineapple A2 was tastier than pineapple B2. We then cut each selected pineapple into small pieces that we placed in separate bowls.

Determining the combination of quality and “bribes” in this way implies that the better tasting pineapple (A1, A2) was always matched to the lower bribe (10 Rupees). In this way, a trade-off between quality and bribes always existed, increasing the power of our study. Further, this procedure ensured that all participants received the same combination of quality and bribes.

We then randomly approached 120 market visitors (shoppers) in the market, one by one. We asked each shopper whether he or she was willing to help us taste a sample from two different pineapples, and tell us which of the two they thought tasted better. The experiment was conducted by two research assistants, a male and a female. The first 60 participants tasted pineapple A1 and pineapple B1; the other 60 participants tasted pineapple A2 and pineapple B2. The procedure for the first and the second group was the same.

We conducted three treatments, with 40 participants in each (20 for each set of pineapples). The first treatment was a control treatment with no bribes, whereas the other two were analogous to treatments KeepWinner and KeepBoth. The shoppers who were approached in the market for our task were never informed that they were taking part in an experiment.

In the Control treatment, we asked the research assistants to follow the following script. After approaching the participant, we asked them to tell the participant the following (translated to the local language--Khasi): "Thank you for agreeing to help us. We will pay you 10 Rupees for your time. We would like to ask you to tell us which of these two pineapples is tastier. It is important for us because we will buy an extra pineapple from the seller who sold us the one you will tell us is tastier. Please taste both and tell us which one is tastier."

The research assistants then asked participants to taste both pineapples and indicate which one was tastier. Participants received their payment of 10 Rupees after making their choice. During the experiment, research assistants were instructed to switch the hands in which they were holding the bowls after each participant, and always start the tasting with the bowl on the left hand. In this way, we counterbalanced any order effect.

Treatment KeepWinner was similar to the control treatment, but instead of paying participants 10 Rupees for tasting the pineapples we told them seller A had offered 10 Rupees to those who chose his pineapple, and seller B had offered 20 Rupees to those who chose his pineapple. Participants were also told they could only keep the money offered by the seller of the pineapple that they indicated as tastier. The following additional wording was added to the script before we asked participants to taste the pineapples: "The seller of this pineapple [the RA holding the bowls raised the bowl containing pineapple A1 or A2] offered you 10 Rupees if you will choose his pineapple, and the seller of this pineapple [now the bowl containing pineapple B1 or B2 was raised] offered 20 Rupees to you if you will choose his pineapple. As a result, you will be paid 10 Rupees if you choose this one and 20 Rupees if you choose this one [again, the respective bowls were raised]." Participants then tasted both pineapples, chose one, and were paid according to their choice.

Treatment KeepBoth was similar to treatment KeepWinner, but participants were told that regardless of their choice, they would be paid both the 10 Rupees offered by seller A and the 20 Rupees offered by seller B. Specifically, the protocol was as follows: "The seller of this pineapple [the RA holding the bowls raised the bowl containing pineapple A1 or A2] offered you 10 Rupees and the seller of this pineapple [the bowl containing pineapple B1 or B2 was raised] offered 20 Rupees. As a result, you will be paid 30 Rupees regardless of your choice." Participants then tasted both pineapples, chose one, and received their payment.

V.B. Results

The results are presented in Figure 3. In the control treatment, 77.5% of the participants indicated pineapple A was tastier. This fraction is significantly larger than predicted by chance (i.e., 50%, $p<.001$, test of proportions), which suggests that pineapple A was indeed tastier than pineapple B. Thus, in this treatment, most participants agreed with the experimenters about which of the two pineapples was tastier.

In treatment KeepWinner, participants chose pineapple A only 35% of the time. The difference between the fraction of participants choosing pineapple A in

KeepWinner and in the control treatment is significant (p<.001, test of proportions).

In treatment KeepBoth, the fraction of participants choosing pineapple A was 67.5%, which is significantly higher than the fraction observed in the KeepWinner treatment (p=.003, test of proportions), and does not differ from the control treatment (p=.317, test of proportions).

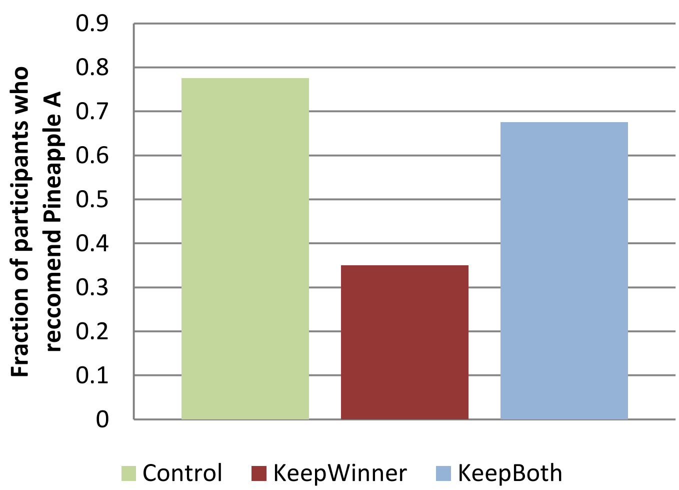

FIGURE 3. THE FRACTION OF PARTICIPANTS WHO CHOOSE THE TASTIER PINEAPPLE (A) BY TREATMENT

We further explore these results using OLS regressions in which we estimate treatment effects on the probability of choosing pineapple A. Since observations are not clustered, we use robust standard errors to compute the p-values. The regression (Table 6) confirms that participants in KeepWinner were significantly less likely to choose the tastier pineapple than participants in the control treatment. Participants in KeepBoth are not significantly less likely to choose the tastier pineapple than participants in the control treatment. The difference between the KeepWinner and the KeepBoth coefficients is also significant (F(1,117)=9.22, p=.003). In column (2), we interact the particular pineapple that was tasted by the subjects with treatment dummies. We find that the treatment effect is similar regardless of the particular combination of pineapples that was tasted.

TABLE 6—OLS REGRESSIONS FOR THE EXTRA-LABORATORY EXPERIMENT IN INDIA

Dependent Variable:	Pineapple A Wins (1=Yes)	
	(1)	(2)
Constant	.775	.750
	(.000)	(.000)
$D_{KeepWinner}$	-.425***	-.400***
	(.000)	(.008)
$D_{KeepBoth}$	-.100	-.050
	(.322)	(.730)
Pineapple A2/B2		.050
		(.712)
Pineapple A2/B2 X $D_{KeepWinner}$		-.050
		(.808)
Pineapple A2/B2 X $D_{KeepBoth}$		-.100
		(.624)
Observations	120	120

Notes: OLS estimates (p-values). The dependent variable is a dummy that specifies whether the referee selected the best-tasting pineapple (A) as the winner. $D_{KeepWinner}$ and $D_{KeepBoth}$ are dummy variables that are equal to one for KeepWinner and KeepBoth respectively, and zero otherwise. Pineapple A2/B2 is a dummy variable that is equal to one if the Pineapples tasted were Pineapple A2 and B2, and equal to zero if the pineapples tasted were A1 and B1. The control treatment serves as the reference treatment. P-values are calculated using robust standard errors.

*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

VI. Concluding Remarks

Bribery is widespread and has an important impact on how decisions are made in politics, business, sports, education, and many other domains, with large economic consequences. Despite some economic arguments that bribes are not necessarily bad for society but are simply used to "grease the wheels" of bureaucracy (e.g., Leff, 1964; Huntington, 1968), even in these cases, when bureaucrats can endogenously choose the level of corruption, bribes clearly have a negative effect on economic efficiency (Banerjee, 1997).

The purpose of the current paper is to investigate the relative importance of greed and social preferences in distorting choices in the presence of bribery. We find that when incentives are contingent on choices, individuals accept and reward bribes: in our experiments, referees systematically rewarded the higher bribe when they could only keep the winner's bribe. However, when bribes were not contingent on delivering a certain outcome, individuals did not distort judgment. This finding supports the greed explanation of bribery. The extra-laboratory experiment we conducted in the market in India confirms the results we observed in San Diego, outside of the lab and with a population that is more accustomed to bribery.

Our experiments show that the norm of reciprocity seems to be weaker than the moral costs of distorting judgment, which are weaker than profit maximization. Our ability to rank these different forces comes from the experimental bribery game that we introduce, which is able to capture the moral costs associated with distortion of judgment that is generated when bribes, rather than performance, are rewarded. We find that distortion plays an important role in explaining whether referees reciprocate the higher bribe. When the decision of which worker wins a prize does not involve a distortion of judgment (as in treatment NoTask), we find that individuals are happy to reciprocate the higher bribe. Further, we find that increasing the moral costs of distortion by providing a more objective task (as in treatment Objective) makes referee less likely to go along with the higher bribe and more likely to choose based on quality. These results thus show that capturing distortion in bribery experiments is important because moral costs of distortion have an important influence on the behavior of participants.

Our investigation of the motives that induce decision makers to accept and reward bribes provides insights into how to better target anti-corruption interventions. One implication is that anti-corruption policies designed to reduce the effectiveness of reciprocity (such as reducing social ties and increasing anonymity) may not be effective. If greed drives behavior, as our results suggest, policy interventions that enforce monetary sanctions may be more effective in preventing corruption.

Policy interventions that focus on increasing the moral costs of distortion may provide an alternative way to reduce bribery. In a related paper (Gneezy, Saccardo, Serra-Garcia, and Van Veldhuizen, 2015), we examine the role of self-deception in distorting judgment, by varying when evaluators are informed about their incentives to recommend one of two options either before or after their initial private judgment. When the information regarding the incentives is provided before (as in treatment KeepWinner here), we find a significant bias in judgment in the direction of the incentive. However, when the information is provided after they privately evaluate the options, the bias in judgment is significantly reduced. In other words, limiting self-deception may increase the moral costs of distortion, which limits the effectiveness of bribery.

Future research could build on our game and findings in at least two other important ways. First, in our experiment, workers who lost because of bribery suffered the negative externality of distorted justice. This negative externality did not reduce the overall wealth of the participants. In many real-world cases of bribery, however, the negative externality could be much larger and reduce the overall earnings. Our game could be extended by incorporating a negative externality (e.g., Falk and Szech, 2013) into the design or by making bribery welfare decreasing.

Second, a future design could study how the chance of being audited and penalized for accepting bribes affects decisions. Our investigation of bribery focused on the case in which the decision of who wins the prize is subjective, in which implementing monitoring and punishment mechanisms is often hard, because of the subjective nature of the choice. Future research could focus on the case in which decision makers have to make decisions that are easily and objectively verifiable, and use this design to study the interplay between the probability of being caught and the size of the penalty, and how this interplay affects the decision to offer or accept a bribe.

REFERENCES

Abbink, Klaus. 2004. “Staff Rotation as an Anti-corruption Policy: An Experimental Study.” *European Journal of Political Economy* 20 (4): 887–906.

Abbink, Klaus. 2005. “Fair Salaries and the Moral Costs of Corruption.” In *Advances in Cognitive Economics*, ed. Boicho N. Kokinov, 2–7. Sofia: NBU Press.

Abbink, Klaus, Utteeyo Dasgupta, Lata Gangadharan, and Tarun Jain. 2014. “Letting the Briber Go Free: An Experiment on Mitigating Harassment Bribes.” *Journal of Public Economics* 111: 17–28.

Abbink, Klaus, Bernd Irlenbusch, and Elke Renner. 2002. “An Experimental Bribery Game.” *Journal of Law, Economics, and Organization* 18 (2): 428–454.

Abbink, Klaus, and Danila Serra. 2012. “Anticorruption Policies: Lessons from the Lab.” In *New Advances in Experimental Research on Corruption (Research in Experimental Economics, Volume 15)*, ed. Danila Serra and Leonard Wantchekon, 77–115. Emerald Group Publishing Ltd.

Akerlof, George A. 1982. “Labor Contracts as Partial Gift Exchange.” *The Quarterly Journal of Economics* 97 (4): 543.

Armantier, Olivier, and Amadou Boly. 2013. “Comparing Corruption in the Laboratory and in the Field in Burkina Faso and in Canada.” *The Economic Journal* 123 (573)): 1168–1187.

Banerjee, Abhijit V. 1997. “A Theory of Misgovernance.” *The Quarterly Journal of Economics* 112 (4): 1289–1332.

Barr, Abigail, and Danila Serra. 2010. “Corruption and Culture: An Experimental Analysis.” *Journal of Public Economics* 94 (11-12): 862–869.

Bazerman, Max H., George Loewenstein, and Don A. Moore. 2002. "Why Good Accountants Do Bad Audits." *Harvard Business Review* 80, no. 11: 96-103.

Becker, Gary S., and George J. Stigler. 1974. "Law Enforcement, Malfeasance, and Compensation of Enforcers." *The Journal of Legal Studies* 3 (1): 1–18.

Belot, Michèle, and Marina Schröder. 2013. "Sloppy Work, Lies and Theft: A Novel Experimental Design to Study Counterproductive Behaviour." *Journal of Economic Behavior & Organization* 93: 233–238.

Bertrand, M., S. Djankov, R. Hanna, and S. Mullainathan. 2007. "Obtaining a Driver's License in India: An Experimental Approach to Studying Corruption." *The Quarterly Journal of Economics* 122 (4): 1639–1676.

Bolton, Gary E., and Axel Ockenfels. 2000. "ERC: A Theory of Equity, Reciprocity, and Competition." *American Economic Review* 90 (1): 166–193.

Burguet, Roberto, and Yeon-Koo Che. 2004. "Competitive Procurement with Corruption." *RAND Journal of Economics* 35 (1): 50–68.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90 (3): 414–427.

Cameron, Lisa, Ananish Chaudhuri, Nisvan Erkal, and Lata Gangadharan. 2009. "Propensities to Engage in and Punish Corrupt Behavior: Experimental Evidence from Australia, India, Indonesia and Singapore." *Journal of Public Economics* 93 (7-8): 843–851.

Cappelen, Alexander W., Erik Ø. Sørensen, and Bertil Tungodden. 2013. "When Do We Lie?" *Journal of Economic Behavior & Organization* 93: 258–265.

Charness, Gary, Uri Gneezy, and Michael A. Kuhn. 2013. "Experimental Methods: Extra-Laboratory Experiments-Extending the Reach of Experimental Economics." *Journal of Economic Behavior & Organization* 91: 93–100.

Del Monte, Alfredo, and Erasmo Papagni. 2001. "Public Expenditure, Corruption, and Economic Growth: The Case of Italy." *European Journal of Political Economy* 17 (1): 1–16.

Dreber, Anna, and Magnus Johannesson. 2008. "Gender Differences in Deception." *Economics Letters* 99 (1): 197–199.

Erat, Sanjiv. 2013. "Avoiding Lying: The Case of Delegated Deception." *Journal of Economic Behavior & Organization* 93: 273–278.

Erat, Sanjiv, and Uri Gneezy. 2012. "White Lies." *Management Science* 58 (4): 723–733.

Falk, Armin, and Nora Szech. 2013. "Morals and Markets." *Science* 340 (6133): 707–11.

Fehr, Ernst, and Simon Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review* 90 (4): 980–994.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *The Quarterly Journal of Economics* 114 (3): 817–868.

Gneezy, Uri. 2005. "Deception: The Role of Consequences." *American Economic Review* 95 (1): 384–394.

Gneezy, Uri, Marta Serra-Garcia, Silvia Saccardo, and Roel van Veldhuizen. 2015. "Motivated Self-Deception, Identity and Unethical Behavior." *Mimeo*.

Goldsmith, Arthur A. 1999. "Slapping the Grasping Hand: Correlates of Political Corruption in Emerging Markets." *American Journal of Economics and Sociology* 58 (4): 865–883.

Huntington, Samuel P. 1968. *Political Order in Changing Societies*. New Haven: Yale University Press.

INDEM. 2005. "Diagnosing Corruption in Russia: A Sociological Analysis." *Russian Social Science Review* 46 (1): 19–36.

Kaufmann, Daniel. 2005. "Myths and Realities of Governance and Corruption." In *Global Competitiveness Report 2005-06*, 81–98. World Economic Forum.

Leff, Nathaniel H. 1964. "Economic Development Through Bureaucratic Corruption." *American Behavioral Scientist* 8 (3): 8–14.

Lightle, John P. 2013. “Harmful Lie Aversion and Lie Discovery in Noisy Expert Advice Games.” *Journal of Economic Behavior & Organization* 93: 347–362.

Malmendier, Ulrike, and Klaus M. Schmidt. 2012. “You Owe Me.” *NBER Working Paper No. 18543.*

Mauro, Paulo. 1995. “Corruption and Growth.” *Quarterly Journal of Economics* 110 (3): 681–712.

Moore, Don A., and George Loewenstein. 2004. “Self-Interest, Automaticity, and the Psychology of Conflict of Interest.” *Social Justice Research* 17 (2): 189–202.

Olken, Benjamin A. 2007. “Monitoring Corruption: Evidence from a Field Experiment in Indonesia.” *Journal of Political Economy* 115 (2): 200–249.

Olken, Benjamin A., and Rohini Pande. 2012. “Corruption in Developing Countries.” *Annual Review of Economics* 4: 479–509.

Pareto, Vilfredo. 1896. “Course of Political Economy.” In *Sociological Writings*. New York: Praeger, 1966.

Rabin, Matthew. 1993. “Incorporating Fairness into Game Theory and Economics.” *American Economic Review* 83 (5): 1281–1302.

Reinikka, Ritva, and Jakob Svensson. 2004. “Local Capture: Evidence from a Central Government Transfer Program in Uganda.” *The Quarterly Journal of Economics* 119: 679–705.

Sequeira, Sandra, and Simeon Djankov. 2014. “Corruption and Firm Behavior: Evidence from African Ports.” *Journal of International Economics* 94 (2): 277–94.

Stroop, John R. 1935. “Studies of Interference in Serial Verbal Reactions.” *Journal of Experimental Psychology* 18 (6): 643–662.

Sutter, Matthias. 2009. “Deception Through Telling the Truth?! Experimental Evidence From Individuals and Teams.” *The Economic Journal* 119 (534): 47–60.

Transparency International. 2011. “Bribe Payers Index 2011.” http://bpi.transparency.org/bpi2011/results/.

Transparency International. 2014. “Corruption Perceptions Index 2014.” https://www.transparency.org/cpi2014/results.

Transparency International India. 2008. “India Corruption Study - 2008.” http://www.transparencyindia.org/resource/survey_study/India%20Corruptino%20Study%202008.pdf.

Van Veldhuizen, Roel. 2013. “The Influence of Wages on Public Officials’ Corruptibility: a Laboratory Investigation.” *Journal of Economic Psychology* 39: 341–356.

Appendix A FOR ONLINE PUBLICATION

A1. The Determinants of Bribing Behavior

TABLE A1—DETERMINANTS OF BRIBE AMOUNT AND JOKE QUALITY

Dependent Variable:	Bribe Amount	Joke Quality	Obj. Score
	(1)	(2)	(3)
Joke Quality	.066		
	(.554)		
Objective Score	.034		
	(.130)		
Confidence	-.034	.794**	31.88**
	(.940)	(.016)	(.028)
Female	-.417*	-.051	9.73**
	(.088)	(.778)	(.034)
Nonnative speaker	.802**	-.539**	2.99
	(.024)	(.038)	(.818)
Not of Asian Ethnicity	.370	.347**	-6.18
	(.114)	(.050)	(.270)
Age	.138**	-.081*	-1.75
	(.036)	(.078)	(.138)
Economics Major	-.168	.137	7.08
	(.660)	(.556)	(.246)
Psychology Major	.169	.199	-8.33
	(.688)	(.502)	(.218)
Engineering/Science Major	.183	-.192	6.35
	(.610)	(.436)	(.288)
Other Social Science Major	-.890**	.054	6.91
	(.038)	(.894)	(.312)
Humanities Major	-.252	.119	-8.07
	(.582)	(.568)	(.192)
Treatment Dummies	Yes	Yes	Yes
Treatment	All Treatments	All Joke Treatments	Objective
Observations	484	320	84
Clusters	242	160	42

Notes: OLS estimates (p-values). Bribe amount is the bribe amount sent by the worker. Joke quality is the average rating of the joke by the independent judges. Objective score is the worker's score on the objective task. Confidence is the worker's confidence in having a better joke than the other worker. The remaining variables are dummies for females, nonnative speakers, and non-Asians, a continuous variable for age, and dummies for different majors, respectively (the omitted groups are biology/chemistry majors and undeclared majors). P-values (in brackets) are calculated using wild bootstraps.

*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

In this section, we investigate whether worker-level characteristics are predictive of bribe size. For example, workers who wrote inferior jokes (or performed worse on the objective task) might have sent higher bribes. Similarly, some of the demographic variables reported in Table 2 might be predictive of how much a worker decides to bribe.

To check which variables are predictive of bribe size, we regressed bribe size on performance on the task, the confidence question, and all of the demographic variables reported in Table 2, where we use biology/chemistry majors as the reference group. We pooled the data from all treatments to have the largest possible sample size. The regression results are reported in Table A1 (column 1).

Overall, we find that the coefficient for joke quality (or the performance on the objective task) is not significantly different from zero, suggesting that overall the quality of workers' performance did not affect their bribing behavior. Further, the coefficient for workers' beliefs about having a better joke than the opponent is also not significant. These coefficients remain insignificant even if we only include either actual quality or the beliefs in the regression.

Our analysis further reveals that non-native speakers, older participants, and men send higher bribes, whereas social science majors (not including economists and psychologists) send smaller bribes.

Table A1 also presents the results of the analysis on the determinants of joke quality. Columns (2) and (3) show that a correlation exists between joke quality (or performance in the objective task) and workers' belief that they will win the prize. The regressions further show that native speakers and subjects that did not have an Asian ethnicity wrote better-quality jokes, and that women performed better in the objective task.

A2. Differences in Bribes

A feature of our design is that although treatment is randomized and workers are randomly paired within sessions, bribes are not determined at random. To analyze referees' behavior and make treatment comparisons, the referees must face similar combinations of bribes across treatments. That is, whenever a referee receives two bribes of different size, the distribution of the difference between the two bribes must be similar across both treatments. In the result sections, we showed that, despite the average bribe being significantly higher in KeepWinner, the difference between bribes was indeed similar in treatment KeepWinner and treatment KeepBoth.

However, referees in treatment KeepWinner are more likely to face two bribes rather than one. Further, referees in treatment KeepWinner are more likely to have to choose between two large bribes (e.g., \$5 and \$3) than referees in treatment KeepBoth. We will discuss these two differences below and provide evidence suggesting that the respective differences cannot explain our results.

One Bribe versus Two Bribes. Referees were more likely to receive two bribes (e.g., \$5 versus \$2) in treatment KeepWinner than in treatment KeepBoth. Indeed, of all the times they received bribes of different sizes, referees in KeepBoth received only one bribe (e.g., 3\$ vs. \$0) 82% of the time, versus only 41% for the referees in the KeepWinner treatment. To correct for this difference, we re-estimate regressions of Table 4 by including separate coefficients for the two-bribes and one-bribe case for both quality and bribes. This approach allows us to compare the importance of bribes and quality between cases in which referees received one or two bribes, respectively.

TABLE A2—OLS REGRESSIONS FOR REFEREES ONE OR TWO BRIBES POSITIVE

Dependent Variable:	Winner (1=Yes)	
	(1)	(2)
Bribe difference	.330**	.131*
[One bribe positive]	(.022)	(.082)
Quality difference (bribes differ)	.066	.236**
[One bribe positive]	(.416)	(.036)
Bribe difference	.280***	.128
[Both bribes positive]	(.002)	(.140)
Quality difference (bribes differ)	-.056	.646
[Both bribes positive]	(.630)	(.162)
Quality difference (bribes identical)	.342**	.273**
	(.022)	(.022)
Treatment	KeepWinner	KeepBoth
Selected Workers	Random	Random
Observations	40	40
Clusters	20	20

Notes: OLS estimates (p-values). The dependent variable is a dummy that specifies whether the selected worker was selected as the winner. For other variable definitions, see the notes to Table 4. Separate coefficients are included for cases in which one bribe is positive and in which both bribes are positive. P-values (in brackets) are calculated using wild bootstraps. We randomly select one worker per referee in each round.

*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

Table A2 presents the results. For the KeepWinner treatment, the coefficients for bribes are significant and nearly identical in both cases. Similarly, the coefficient for quality is small and not significant in either case. This finding suggests that within treatment, KeepWinner referees do not react differently to bribery in the one-bribe versus two-bribes case.

Treatment KeepBoth does not have enough observations with two different bribes for us to estimate the coefficients for the two-bribes case with any precision. For the one-bribe case, the coefficient for bribes is slightly larger than the coefficient in Table 4 and is significant at the 10% level. At the same time, the (standardized) coefficient for quality is still nearly twice as large as the (standardized) coefficient for bribes, suggesting that quality played a larger role in the referee's decision making. Further, the comparison with the KeepWinner treatment shows the coefficient for bribe difference in KeepBoth is 60% smaller than the coefficient for KeepWinner, whereas the coefficient for quality is substantially larger. These results are similar to the results of Table 4. Thus, the treatment difference between KeepWinner and KeepBoth does not appear to have been the result of referees being more likely to receive only one bribe in the KeepBoth treatment.

Absolute Bribe Size.—A related difference between KeepWinner and KeepBoth that arises from the fact that bribes in our experiment are not determined at random is that referees in KeepWinner are more likely to face two relatively large bribes. Referees might respond differently to receiving two large bribes than to receiving two smaller ones, even when the difference in bribes is the same. Explaining our treatment effect would require that referees care *more* about a difference of \$5 vs. \$4 than \$2 vs. \$1, as could be the case if referees are more willing to reward a bribe when the absolute value of the bribe is high (note, however, that the converse argument could also be made).

To check whether this was indeed the case, we redid the regressions of Table 4 while controlling for the sum of the two bribes. We also interact the sum variable with the bribe-difference variable. The estimate for the interaction effect tells us whether bribes had a larger (or smaller) role when the sum of the two bribes was larger.

Table A3 shows the results. The interaction effect is not significant for either treatment. If anything, the coefficient estimate for the KeepBoth treatment ($p=.234$) suggests referees care *less* about bribes as the sum of the two bribes increases, which would increase the size of the treatment difference. Thus, overall, we find no evidence that differences in absolute bribe size can explain our treatment effect.

TABLE A3—OLS REGRESSIONS FOR REFEREES WITH ABSOLUTE BRIBE SIZE

Dependent Variable:	Winner (1=Yes)	
	(1)	(2)
Bribe difference	.327***	.110
	(.000)	(.278)
Quality difference (bribes differ)	-.012	.291**
	(.904)	(.012)
Quality difference (bribes identical)	.365**	.274**
	(.016)	(.022)
Sum of the two bribes	.066	.083
	(.376)	(.362)
Bribe Difference X Sum of the two bribes	.031	-.130
	(.802)	(.234)
Treatment	KeepWinner	KeepBoth
Selected Workers	Random	Random
Observations	40	40
Clusters	20	20

Notes: OLS estimates (p-values). The dependent variable is a dummy that specifies whether the selected worker was selected as the winner. The sum of the two bribes is the sum of the bribes of both workers in the pair. For other variable definitions, see the notes to Table 4. P-values (in brackets) are calculated using wild bootstraps. We randomly select one worker per referee in each round.

*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

A3. The Quality Threshold

All non-parametric tests in the main text that relate to quality use only those pairs of jokes in which at least 65.1% of independent raters agreed on the winner. With this threshold, the fraction of independent raters who selected a given joke over the other is significantly different from chance (i.e., 50%) at the 10% level by a test of proportions.

In this section, we redo the main analysis using a more stringent threshold of 69.4%, which entails that the two jokes are significantly different at the 5% level (test of proportions, Z=1.65, n=18, p=.050). The table below summarizes the non-parametric results. The results are very similar for both thresholds. A slight difference is that the comparison between the KeepWinner and KeepBoth treatment is significant at the 1% level for the 65.1% threshold and at the 5% level for the more stringent threshold, which is largely due to a loss of observations. Similarly, the difference between KeepWinner and KeepBothReject is no longer significant, because of a lack of observations, although the result is still very similar and not significantly different from KeepBoth (p=.611, Mann-Whitney).

TABLE A4—NON-PARAMETRIC TESTS FOR ALTERNATIVE THRESHOLD

		Better Quality Wins (%)		Difference vs. KeepWinner	
Treatment	Threshold:	65.1%	69.4%	65.1%	69.4%
KeepWinner		56.7%	60.0%		
		(.564)	(.405)		
KeepBoth		90.5%	88.9%	33.80	28.90
		(.001)	(.002)	(.004)	(.030)
KeepWinnerReject		57.1%	58.3%	0.40	-1.70
		(.655)	(.655)	(.978)	(.852)
KeepBothReject		90.5%	81.8%	33.80	21.80
		(.005)	(.014)	(.026)	(.119)
HighWage		44.0%	40.9%	-12.70	-19.10
		(.593)	(.564)	(.439)	(.331)
Objective		83.3%	86.4%	26.60	26.40
		(.001)	(.001)	(.023)	(.023)

Notes: Percentages (P-values). The first two columns display the percentage of times the best joke won for each treatment, for a given threshold. The last two columns display the difference between the percentage of winners between the respective treatment and treatment KeepWinner. In both cases, two different thresholds are used. For the 65.1% threshold, the two jokes in a pair differ significantly at the 10% level by a test of proportions; for the 69.4% threshold, the two jokes differ significantly from each other at the 5% level. P-values are computed using Wilcoxon tests and a Mann-Whitney tests for columns 1 and 2 and columns 3 and 4 respectively.

A4. Alternative Regression Specifications and Demographic Controls

For our main regressions (Table 4 and Table 5), we selected one worker per pair at random for each given round and then investigated how differences in quality and bribes between the selected worker and the opponent affect winning. Table A5 provides the results of two alternative specifications, which we first estimate separately for the two main treatments.

In regression (1) and (2), we focus on the workers who submitted a larger bribe than their opponent in a given round. We then investigate whether for those workers, increasing the quality of the joke relative to the opponent further increased their likelihood of winning. In treatment KeepWinner, the higher bribe already wins 86% of the time. Hence, increasing the quality of the worker's joke (relative to the opponent) does not further increase the likelihood of winning. By contrast, in treatment KeepBoth (column 2), the higher bribe only wins 64% of the time, and increasing the relative quality of the worker's joke is highly beneficial. These results are in line with the results in the main text: quality matters in treatment KeepBoth, but not in treatment KeepWinner.

TABLE A5—ALTERNATIVE OLS REGRESSIONS FOR KEEPWINNER AND KEEPBOTH

Dependent Variable:	Winner (1=Yes)			
	(1)	(2)	(3)	(4)
Bribe Difference			.270** (.032)	.014 (.518)
Quality Difference (bribes differ)	.060 (.254)	.275** (.018)		
Treatment	KeepWinner	KeepBoth	KeepWinner	KeepBoth
Selected Workers	Higher Bribe	Higher Bribe	Higher Quality	Higher Quality
Observations	29	22	30	21
Clusters	16	12	18	17

Notes: OLS estimates (p-values). The dependent variable is a dummy that specifies whether the selected worker was selected as the winner. For other variable definitions, see the notes to Table 4. P-values are calculated using wild bootstraps. For specification (1) and (2), we select only workers with a higher bribe than their opponent in the given round. For specification (3) and (4), we select only workers with a better-quality joke than their opponent in the given round. We consider a joke to be of better quality when the agreement of the independent raters is at least 65.1%.

*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

In regression (3) and (4), we instead focus on the workers who wrote the better joke in a given round, and investigate whether larger differences in bribes between those worker and their opponents make the workers even more likely to win the prize. The results show that for workers who already have the best joke in the pair, bribes have a strong positive effect on the likelihood of winning in treatment KeepWinner, but not in treatment KeepBoth. Thus, similar to the regressions presented in the main text, bribes matter in treatment KeepWinner but not in treatment KeepBoth.

In Table A6 we estimate the regressions reported in Table A5 jointly for all treatments. We report four regressions. In columns (1) and (2), we use KeepWinner as the reference treatment and interact the bribe and quality-difference variables with treatment dummies for all other treatments. This approach allows us to verify whether bribes and quality played a larger or smaller role than in treatment KeepWinner. Columns (3) and (4) are similar to columns (1) and (2) but also include a series of demographic control variables. In particular, we add controls for all the demographic variables reported in Table 2 and for the round.[6]

The results largely replicate the main results from Table 5. Relative to KeepWinner, bribes play a less important role in KeepBoth and KeepBothReject, whereas quality plays a larger role in these treatments and treatment Objective. In addition, including demographic controls does not substantially affect the coefficient estimates. In particular, p-values are very similar across both specifications.

[6] We could not control for demographics in the regressions presented in the main text. When workers are randomly selected, selected workers have a 50% chance of winning, irrespective of the referee's demographics, and hence all demographic variable coefficients are zero in expectation.

TABLE A6—ALTERNATIVE OLS REGRESSIONS FOR ALL TREATMENTS

Dependent Variable: Winner (1=Yes)	(1)	(2)	(3)	(4)
Bribe Difference		.265**		.255**
		(.032)		(.036)
Bribe Difference X $D_{KeepBoth}$		-.249***		-.203**
		(.006)		(.042)
Bribe Difference X $D_{KeepWinnerReject}$		.218***		.232**
		(.006)		(.020)
Bribe Difference X $D_{KeepBothReject}$		-.317**		-332**
		(.012)		(.018)
Bribe Difference X $D_{HighWage}$		.013		.035
		(.892)		(.766)
Bribe Difference X $D_{Objective}$		-.168		-.191
		(.142)		(.128)
Quality Difference (Different Bribes)	.060		.067	
	(.254)		(.258)	
Quality Difference (Different Bribes) X $D_{KeepBoth}$	.181**		.211**	
	(.032)		(.008)	
Quality Difference (Different Bribes) X $D_{KeepWinnerReject}$	-.060		-.041	
	(.274)		(.494)	
Quality Difference (Different Bribes) X $D_{KeepBothReject}$	.381**		.390**	
	(.012)		(.014)	
Quality Difference (Different Bribes) X $D_{HighWage}$	-.020		-.052	
	(.832)		(.670)	
Quality Difference (Different Bribes) X $D_{Objective}$	.228**		.208*	
	(.040)		(.052)	
Selected Workers	Higher Bribe	Higher Quality	Higher Bribe	Higher Quality
Treatment Dummies	Yes	Yes	Yes	Yes
Demographic Controls	No	No	Yes	Yes
Observations	132	134	132	134
Clusters	79	90	79	90

Notes: OLS estimates (p-values). P-values are computed using wild bootstraps. The '$D_{Treatment}$' variables are dummy variables for the respective treatments; KeepWinner serves as the reference treatment. Control variables include the round and all demographic variables of Table 2, where we use bio/chemistry majors and undeclared majors as the reference group. Bribe and quality variables are standardized using the combined standard deviation over all included treatments. For other variable definitions and explanation of how we selected the workers in each regression, see the notes to Table 4.

*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

Identical Quality—In the regressions reported in the main text, we estimate separate coefficients for quality when bribes are equal and when bribes differ. Here, we present a similar analysis in which we also estimate separate coefficients for bribes when quality is equal (using the 65.1% threshold) and when quality differs. The latter coefficient allows us to focus on cases in which quality differs, which is where a larger treatment difference should be expected. We do not use this approach in the main text because the threshold between 'similar' jokes and different jokes is not as clear-cut as the threshold between identical and non-identical bribes.

Table A7 reports the results. For treatment KeepWinner, bribes have a larger effect when quality is equal and therefore cannot influence judgment. Similarly, in treatment KeepBoth, the coefficient for bribes is larger when quality is equal, though it is still not significant. The interaction terms in column (3) shows that the treatment difference in the importance of bribes overall seems to be largely driven by cases in which quality differs, as expected. These results indeed suggest that bribes play a larger role when quality is equal in both treatments.

TABLE A7—OLS REGRESSIONS FOR REFEREES WITH SEPARATE BRIBE COEFFICIENTS

Dependent Variable:	Winner (1=Yes)		
	(1)	(2)	(3)
Bribe Difference (quality identical)	.453**	.171	.348**
	(.016)	(.180)	(.016)
Bribe Difference (quality differs)	.267**	.033	.254**
	(.032)	(.458)	(.032)
Quality Difference (bribes differ)	.009	.270**	.010
	(.894)	(.020)	(.894)
Quality Difference (bribes identical)	.339**	.276**	.257**
	(.012)	(.034)	(.012)
$D_{KeepBoth}$			.024
			(.862)
Bribe Difference (quality identical) X $D_{KeepBoth}$			-.145
			(.274)
Bribe Difference (quality differs) X $D_{KeepBoth}$			-.216**
			(.024)
Quality Difference (bribes differ) X $D_{KeepBoth}$			.234***
			(.008)
Quality Difference (bribes identical) X $D_{KeepBoth}$			.151
			(.164)
Treatment	KeepWinner	KeepBoth	KeepWinner KeepBoth
Selected Workers	Random	Random	Random
Observations	40	40	80
Clusters	20	20	40

Notes: OLS estimates (p-values). The dependent variable is a dummy that specifies whether the selected worker was selected as the winner. For other variable definitions, see the notes to Table 4. P-values (in brackets) are calculated using wild bootstraps. We consider jokes to be of identical quality when fewer than 65.1% of independent raters agree on the winner. We randomly select one worker per referee in each round.

*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

Main and Reject treatments. The analysis of treatment KeepWinnerReject and KeepBothReject reported in the main text reveals that giving referees the opportunity to reject bribes did not change their behavior, as compared to treatment KeepWinner and KeepBoth respectively. As an additional robustness check for the results presented in Table 4, we can therefore merge the data of Reject treatments and the main treatments and redo the analysis with the merged data. The results (Table A8) are very similar to Table 4. Referees in treatment KeepWinner select the highest bribe as the winner, whereas referees in treatment KeepBoth select the best joke. The main difference with Table 4 is that the main treatment comparisons are now significant at the 1% level, which is due to the larger sample size used.

TABLE A8—OLS REGRESSIONS FOR REFEREES IN KEEPWINNER AND KEEPBOTH INCLUDING REJECT TREATMENTS

Dependent Variable:		Winner (1=Yes)	
	(1)	(2)	(3)
Bribe Difference	.337***	.052	.311***
	(.000)	(.274)	(.000)
Quality Difference (bribes differ)	.006	.318***	.006
	(.866)	(.002)	(.866)
Quality Difference (bribes identical)	.257*	.219**	.197*
	(.054)	(.012)	(.054)
$D_{KeepBoth}$			.050
			(.564)
Bribe Difference X $D_{KeepBoth}$			-.252***
			(.002)
Quality Difference X $D_{KeepBoth}$			.301***
(bribes differ)			(.002)
Quality Difference X $D_{KeepBoth}$			.089
(bribes identical)			(.448)
Treatment	KeepWinner	KeepBoth	KeepWinner KeepBoth
Selected Workers	Random	Random	Random
Observations	60	60	120
Clusters	30	30	60

Notes: OLS estimates (p-values). The dependent variable is a dummy that specifies whether the selected worker was selected as the winner. Quality Difference is the difference between the quality of the joke (i.e., the average rating by the independent raters) of the selected worker and the quality of the joke of the other worker in the pair. Bribe Difference is the difference between the bribe sent by the selected worker and the bribe sent by the other worker in the pair. $D_{KeepBoth}$ is a dummy that is equal to one for treatment KeepBoth, and zero otherwise. For column (3), the bribe variable and both quality variables are standardized using the respective variable's combined standard deviation over all included treatments. P-values are calculated using wild bootstraps. For each regression, we randomly select one worker per referee in each round.

*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

A5. Robustness Check: Alternatives to the Wild Bootstrap

In the main regression analysis, we computed p-values using wild bootstraps, as suggested by Cameron et al. (2008). In this section, as a robustness check, we provide the results of several alternative techniques, which we use to compute the p-values of the KeepWinner and KeepBoth treatment regressions reported in Table 4.

We use four different techniques. The first is the wild bootstrap procedure, which we use in the main text. Cameron et al. (2008) show that this approach (which they refer to as the "wild bootstrap-t" procedure) leads to more consistently accurate (and more conservative) rejection rates than alternative approaches, and therefore recommend its use in case of a small number of clusters. For more details on how the technique works, see Cameron et al. (2008).

For column (2), we recalculate our p-values using clustered standard errors, a standard approach in experimental economics. For column (3), we use a non-parametric bootstrap, which is, to our knowledge, the most widely used bootstrap method in experimental economics. Finally, we also redo the main regressions using probit (with clustered standard errors).

The top half of Table A9 reports the results for treatment KeepWinner; the lower half presents the results for treatment KeepBoth. Each column presents the results of a different estimation technique. Column (1) reports the results of Table 4. The other columns show that alternative estimation techniques result in very similar p-values. In particular, bribes are significant at the 1% level across specifications, and quality (for different bribes) is never significant in any specification. The only difference is that quality (for equal bribes) is significant at the 5% level with the wild bootstrap, but significant at the 1% level in all other specifications.[7]

For KeepBoth, the results are also very similar across specifications. Bribes are not significant, and quality (for different bribes) is significant at the 1% level in all specifications. As with KeepWinner, the main difference is that the quality variable for equal bribes is significant at the 5% level in the wild bootstrap and probit, but at the 1% level in the other specifications.

TABLE A9—SEPARATE OLS REGRESSIONS WITH ALTERNATIVE P-VALUE ESTIMATES FOR MAIN TREATMENTS

Dependent Variable: Winner (1=Yes)	(1)	(2)	(3)	(4)
(KeepWinner)				
Bribe Difference	.308***	.308***	.308***	.488***
	(.000)	(.000)	(.000)	(.000)
Quality Difference (bribes differ)	.014	.014	.014	.054
	(.762)	(.794)	(.830)	(.619)
Quality Difference (bribes identical)	.336**	.336***	.336***	.526***
	(.020)	(.000)	(.002)	(.002)
(KeepBoth)				
Bribe Difference	.086	.086	.086	.154
	(.140)	(.161)	(.222)	(.205)
Quality Difference (bribes differ)	.262**	.262***	.262***	.613***
	(.010)	(.001)	(.004)	(.005)
Quality Difference (bribes identical)	.275**	.275***	.275***	.420**
	(.022)	(.000)	(.001)	(.012)
Selected Workers	Random	Random	Random	Random
Technique	OLS	OLS	OLS	Probit (Marg. Eff.)
P-Values	Wild BS	Clustered SE	Non-Par BS	Clustered SE
Observations	40	40	40	40
Clusters	20	20	20	20

Notes: Regression estimates (p-values). Each column presents the results of two regressions, which for the top and bottom panel are analogous to regressions 1 and 2 in Table 4, respectively. For variable definitions and other details, see Table 4 and its notes. Columns (1)-(3) use OLS, column (4) uses probit. P-values are computed either by wild bootstrap (column (1)), clustered standard errors (column (2) and (4)) or a non-parametric bootstrap (column (3)).

*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

Table A10 presents the regression with interaction effects (column (3) in Table 4). We do not estimate the probit model here, because interaction terms are

[7] The marginal-effect estimates for probit are typically larger than the coefficient estimates of OLS, which is the result of the marginal effect being estimated at the sample average, where the probability of winning is approximately .5.

difficult to interpret with marginal effects. Similar to Table A8, the p-values are quite robust with respect to the estimation technique we use.

Overall, the wild bootstrap technique and alternatives yield similar p-values in both Table A9 and A10. If anything, the wild bootstrap tends to be the most conservative technique, which is in line with Cameron et al. (2008).

TABLE A10—JOINT OLS REGRESSIONS WITH ALTERNATIVE P-VALUE ESTIMATES FOR MAIN TREATMENTS

Dependent Variable: Winner (1=Yes)	(1)	(2)	(3)
Bribe Difference	.274***	.274***	.274***
	(.000)	(.000)	(.000)
Quality Difference (bribes differ)	.015	.015	.015
	(.762)	(.791)	(.839)
Quality Difference (bribes identical)	.255**	.255***	.255***
	(.020)	(.000)	(.005)
$D_{KeepBoth}$	.008	.008	.008
	(.980)	(.942)	(.947)
Bribe Difference X $D_{KeepBoth}$	-.173**	-.173**	-.173*
	(.032)	(.035)	(.058)
Quality Difference X $D_{KeepBoth}$	.222**	.222***	.222**
(bribes differ)	(.014)	(.009)	(.049)
Quality Difference X $D_{KeepBoth}$	.150	.150	.150
(bribes identical)	(.156)	(.142)	(.320)
Treatment	KeepWinner	KeepWinner	KeepWinner
	KeepBoth	KeepBoth	KeepBoth
Selected Workers	Random	Random	Random
Technique	OLS	OLS	OLS
P-Values	Wild BS	Clustered SE	Non-Par BS
Observations	40	40	40
Clusters	20	20	20

Notes: OLS estimates (p-values). Each column presents the results of a single regression, which is analogous to regression 3 in Table 4. For variable definitions and other details, see Table 4 and its notes. P-values are computed either by wild bootstrap (column (1)), clustered standard errors (column (2)), or a non-parametric bootstrap (column (3)).

*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

A6. Robustness Check: Alternative Random Samples of Workers

In our main regression analyses, we randomly selected one worker per referee in a given round. We used the same random sample for all regressions. Here, we present evidence that shows the results we present do not depend on the particular sample of workers we randomly selected. For this purpose, we re-estimate of the regressions of Table 4 for 1,000 alternative random samples of workers. In every random sample, each referee and round combination is represented exactly once; the only thing that differs across samples is whether worker A or worker B is included in the analysis (for a given pair and round).

Table A11 below presents the average resulting coefficient estimates as well as the standard deviation of the coefficient estimates (in square brackets). Overall, coefficient estimates do not differ much across the different samples. The only exception is the dummy for treatment KeepBoth in column (3); this exception is due to the random sample sometimes containing more winners for treatment KeepBoth than for KeepWinner and vice versa. Overall, the results are very robust with respect to the particular random sample selected for these regressions.

TABLE A11—OLS REGRESSIONS WITH ALTERNATIVE RANDOM SAMPLES

Probability (winning)	(1)	(2)	(3)
Bribe Difference	0.304	0.092	0.275
	[0.011]	[0.011]	[0.010]
Quality Difference (bribes differ)	0.014	0.253	0.014
	[0.011]	[0.019]	[0.012]
Quality Difference (bribes identical)	0.341	0.274	0.251
	[0.028]	[0.020]	[0.015]
$D_{KeepBoth}$			-0.004
			[0.095]
Bribe Difference X $D_{KeepBoth}$			-0.170
			[0.017]
Quality Difference X $D_{KeepBoth}$			0.222
(bribes differ)			[0.021]
Quality Difference X $D_{KeepBoth}$			0.144
(bribes identical)			[0.028]
Treatment	KeepWinner	KeepBoth	KeepWinner KeepBoth
Observations	40	40	80
Clusters	20	20	40

Notes: OLS estimates [standard deviations]. For each column, we re-estimate the regression of, respectively, columns (1), (2), and (3) of Table 4 1,000 times. For each replication, we randomly select a new sample of one worker per referee per round. The presented estimates are the average coefficient estimates (over all replications). The standard deviations are the standard deviations of the coefficient estimates. For variable definitions, see Table 4 and its notes.

*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

A7. Comparison between rounds

The analysis of referees' behavior aggregates observations for rounds 1 and 2. However, it seems possible that some referees changed their behavior between rounds. To allow for this, Table A12 re-estimates regressions (1) and (2) from Table 4 separately for each round. This allows us to investigate whether the impact of quality and bribes was different in round 1 and round 2. This analysis uses only one observation per cluster, and hence we no longer compute standard errors using wild bootstraps, but use robust standard errors to compute p-values instead.

TABLE A12 – OLS REGRESSIONS FOR KEEPWINNER AND KEEPBOTH SEPARATELY FOR EACH ROUND

Probability (winning)	(1)	(2)	(3)	(4)
Bribe Difference	.292***	.295***	.184**	-.049
	(.000)	(.001)	(.025)	(.556)
Quality Difference	-.112	.113	.270***	.289**
(bribes differ)	(.118)	(.246)	(.000)	(.014)
Quality Difference	.342***	.372***	.280**	.293***
(bribes identical)	(.003)	(.002)	(.018)	(.001)
Treatment	KeepWinner	KeepWinner	KeepBoth	KeepBoth
Selected Workers	Random	Random	Random	Random
Standard Errors	Robust	Robust	Robust	Robust
Round	1	2	1	2
Observations	20	20	20	20
Clusters	20	20	20	20

Notes: OLS estimates (p-values). The dependent variable is a dummy that specifies whether the selected worker was selected as the winner. For other variable definitions, see Table 4 and its notes. P-values (in brackets) are calculated using robust standard errors. For all specifications, we randomly select one worker per referee in each round.

*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

Columns (1) and (2) show that in the KeepWinner treatment, bribes matter equally in both rounds. Further, we find that quality has a similar impact in both rounds when bribes are identical, and is not significant in either round when bribes differ. For treatment KeepBoth, columns (3) and (4) show the effect of quality is more important than the effect of bribes in both rounds. Both quality coefficients are very similar across rounds and are significant in both cases. The only difference from the analysis presented in the main text is that the coefficient for bribes is zero in round 2 and positive and significant in round 1. This finding suggests referees may have been more receptive toward bribes in round 1 than in round 2, though the effect of quality stays the same through both rounds.[8]

Another potential explanation for why behavior might differ across rounds is that referees may have alternated between workers across rounds. This explanation may in particular have been the case when bribes and/or joke quality were equal. In the remainder of this section, we check whether alternating played an important role in either of the two main treatments.

First, we show that alternating cannot by itself explain our results. Indeed, if all referees were alternating, neither quality nor bribes would be significant in any treatment. To illustrate this, we ran simulations in which we used workers' actual behavior from the experiment but replaced referees' choices with a random winner in round 1 and then had them choose the other worker in round 2.

The results of this simulation are displayed in Table A13 and show that, as expected, the average estimated coefficient over all simulations is essentially zero for all variables and both treatments. Further, the quality and bribe variables are only significant in approximately 5% of all simulations in both treatments, again as expected by chance. Third, in only .1% of simulation samples (i.e., 2 out of 2,000 across both treatments) were the same coefficients significant (at the 5% level) and had the same sign as in the actual regression estimates. Thus, the simulations strongly suggest our results cannot have been generated by alternating alone.

[8] One potential reason for this result is that, by chance, a greater number of pairs of jokes in round 1 were very similar in quality. If referees found that deciding between two very similar jokes was difficult, they may have selected the better bribe instead. In line with this reasoning, we showed in appendix A4 that referees appear to have been more likely to care about bribes when jokes were of similar quality.

TABLE A13—SIMULATIONS FOR ALTERNATING

	Actual Estimate	Simulation Estimate	Sim. Est. Significant
KeepWinner			
Bribe Difference	.308***	-.001	5.7%
Quality Difference (bribes differ)	.014	.003	5.0%
Quality Difference (bribes identical)	.336**	-.001	5.6%
KeepBoth			
Bribe Difference	.086	.001	4.9%
Quality Difference (bribes differ)	.262**	.000	5.1%
Quality Difference (bribes identical)	.275**	.001	5.2%

Notes: The first column (Actual Estimates) gives the actual coefficient estimates from column (1) (upper panel) and column (2) (lower panel) of Table 4. The second column (Simulation Estimate) presents the corresponding average coefficient estimates from the simulations with alternating. Column 3 (Sim. Est. Significant) presents the percentage of times in which the respective estimated coefficient was significant at the 5% level (in either direction) in a simulation. Here, we used clustered standard errors to calculate significance, because using bootstraps would have been too computationally intensive.

*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level (Column 1 only)

Though referees alternating across rounds cannot fully explain our results, alternating behavior might have still played a role in some cases. For example, referees might have alternated in cases in which bribes and/or quality were very similar. To check whether alternating played a role, we control for it directly by adding a dummy for round 1 winners to our regression and interacting it with a round 2 dummy. If alternating plays a role, round 1 winners should be less likely to win in round 2, and hence the coefficient for this interaction effect should be negative and significant.

Table A14 below presents the results in columns (2) and (4). Columns (1) and (3) report the relevant regressions from Table 4 (columns (1) and (2) respectively) for ease of comparison. In KeepWinner, the coefficient for round 1 winners is *positive* and significant at the 10% level. A positive coefficient is inconsistent with alternating, but note the effect is small and could be spurious, for example, if the round 1 winner had the better bribe in both round 1 and round 2. Importantly, the comparison between column (1) and (2) shows that the coefficients for bribes and quality remain mostly unaffected by allowing for alternating.

In KeepBoth (column (4)), however, alternating does seem to have played a role: round 1 winners are 53.1 percentage points less likely to win in round 2 than round 1 losers (who have a 75% chance of winning in round 2). The coefficient for bribes also increases slightly and is now significant, though it is still substantially smaller than the quality coefficients and the bribe coefficient for KeepWinner. Allowing for alternating does not affect the coefficients for quality.

Why does alternating play a role in KeepBoth but not in KeepWinner? One possible explanation is that in the absence of profit-maximization motives, referees typically select the better joke. However, in some cases, the two jokes were of very similar quality, and referees might alternate in such cases. Thus, alternating may have emerged when neither greed nor joke quality could determine referee behavior. At the same time, it is important to emphasize that controlling for alternating does not significantly change the coefficient estimates for quality, which still plays a more important role in treatment KeepBoth. Similarly, controlling for alternating does not affect the coefficient estimates for KeepWinner, where bribes are more important than quality.

TABLE A14 – OLS REGRESSIONS ALLOWING FOR ALTERNATING

Probability (winning)	(1)	(2)	(3)	(4)
Bribe Difference	.308***	.272***	.086	.143**
	(.000)	(.000)	(.140)	(.020)
Quality Difference	.014	.006	262**	.264***
(bribes differ)	(.762)	(.894)	(.010)	(.008)
Quality Difference	.336**	.326**	.275**	.342**
(bribes identical)	(.020)	(.014)	(.022)	(.014)
Round 2		-.029		.252*
		(.852)		(.072)
Round1Winner		.257*		-.531**
*Round 2		(.094)		(.010)
Treatment	KeepWinner	KeepWinner	KeepBoth	KeepBoth
Selected Workers	Random	Random	Random	Random
Observations	40	40	40	40
Clusters	20	20	20	20

Notes: OLS estimates (p-values). The dependent variable is a dummy that specifies whether the selected worker was selected as the winner. Round 2 is a dummy for round 2 observations, and round 1 winner is a dummy for workers who won in round 1. For other variable definitions, see Table 4 and its notes. P-values (in brackets) are calculated using wild bootstraps. We randomly select one worker per referee in each round.

*** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level

A8. Equilibrium for KeepBoth

This section gives an overview of equilibrium bribing behavior for workers in treatment KeepBoth, under the assumption that the referee allocates the prize based on the bribes. Since referees get to keep the bribes of both workers in treatment KeepBoth, however, they are financially indifferent between the two workers. Therefore, we allow for the possibility that the referee may instead use an allocation rule that rewards the *lower* bribe with some probability. Specifically, the referee will reward the higher bribe with probability δ, where δ can range from 0 (referees always select the lower bribe) to 1 (referees always select the higher bribe), and allocates the prize randomly (with equal probability) if both bribes are equal.

Given this allocation rule, the expected value of bribing a given amount *b* equals:

$$E\Pi(b_i = b) = P(b > b_j)\delta p + P(b = b_j)\frac{1}{2}p + P(b < b_j)(1 - \delta)p - b$$

Here, b_j is the bribe of the competing worker, and $p = 10$ is the prize obtained by the winning worker. In equilibrium, each worker i chooses a strategy $\sigma_i = \{\sigma_b\}^i = (\sigma_0, \dots, \sigma_5)$ that specifies the probability that worker i plays any given bribe, such that the expected payoff is maximized given the strategy of the competing worker, σ_j. Since the best response functions depend on the referee's allocation rule, equilibrium worker behavior depends on δ.

We focus on pure strategy equilibria of the game, if they exist, and otherwise specify the symmetric mixed equilibrium. Deriving the equilibrium for a given δ can be done in three steps. The first step is to eliminate strictly dominated bribes. If more than one bribe amount remains, the second step is to then check for pure strategy equilibria. If necessary, the final step is to compute the symmetric mixed equilibrium (ME). For the ME, no profitable deviations should be possible. This means that all bribe amounts that are assigned strictly positive probabilities in equilibrium need to have equal expected payoffs, and other bribe amounts need to have a lower (or equal) expected payoff. Hence, deriving the ME entails solving the system of equations $E\Pi(b = 0) = E\Pi(b = 1) =..$ subject to the constraint that all σ_b are non-negative and sum to one (for both workers).

For $\delta \leq 0.5$, referees on average let the worse bribe win. As a result, the unique equilibrium is for both workers to bribe zero (i.e., $\sigma_0 = 1$ for both workers), since any larger bribe is both costly and results in a weakly smaller likelihood of winning.

Figure A1 below plots the equilibria for increasing values of δ, on the interval $\delta \in [0.5,1]$, for the parameters used in the experiment. For values close to $\delta = 0.5$, choosing a bribe of 1 or more would mean a sure loss of the bribe (i.e., 1 or more) versus an expected gain of $((\delta - 0.5) * 10 < 1)$. Hence for $\delta \in [0.5,0.6)$, both workers bribing zero is also the unique equilibrium.

For $\delta = 0.6$, there is an additional pure strategy equilibrium where both workers bribe 1, as well as an equilibrium where one worker bribes 1 and the other bribes 0.

For $\delta \in (0.6,1)$ there are only mixed equilibria. For $\delta \in (0.6,0.7]$, the equilibrium is for both workers to randomize between 0, 1 and 2, where 0 and 2 are chosen with equal probability. Intuitively, there can be no pure strategy equilibria, since the best responses to 0, 1 and 2 are 1, 2 and 0 respectively. As δ increases over the interval, 2 is becoming an ever more attractive bribe, since the advantage relative to a bribe of 1 is increasing, and the disadvantage with respect to a bribe of 0 is decreasing. Therefore, in equilibrium, the frequency of bribes of 0 and 2 must increase.

For $\delta \in [0.7,0.8]$ the equilibrium is for workers to randomize between 0, 1, 2, 3 and 4, where 0, 2 and 4 are chosen with equal probability, as are 1 and 3. Bribes of 3 and 4 are added to the equilibrium since they are no longer strictly dominated by a bribe of zero, otherwise the intuition is similar to the above.

For $\delta \in [0.8,0.9]$, equilibrium randomization occurs between 0,1,2,3, and 5, again with 0 and 2 chosen equally often, and similarly for 1 and 3, with 5 chosen more often than any of the other bribes. For $\delta \in [0.9,1)$ workers randomize between 0,1 and 5, with 5 chosen most often. Intuitively, as the highest bribe starts winning with high regularity, bribing 5 becomes more attractive, which is amplified by the fact that bribes of 6 or higher were not permitted in the experiment. Given that many workers bribe 5, it is no longer attractive to bribe 4 or other intermediate amounts, since these bribes will always lose to a bribe of 5 but still incur the certain cost of sending the bribe. However, bribes of zero are still attractive, since they are costless and have a positive probability of wining even against bribes of 5. Finally, bribes of 1 are also not very costly and have the added advantage of beating the 0 bribes with high likelihood.

In addition, for $\delta \in \{0.7,0.8,0.9\}$, any linear combination between both relevant aforementioned equilibria is also a ME. Finally, for $\delta = 1$, the equilibrium is a pure strategy equilibrium where both workers bribe the maximum. Note also that the fraction of referees who picked the highest bribe in treatment KeepBoth in the experiment was .6364. The mixed equilibrium corresponding to this fraction is for workers to bribe 0 or 2 with probability .267 and bribe 1 with probability .467.

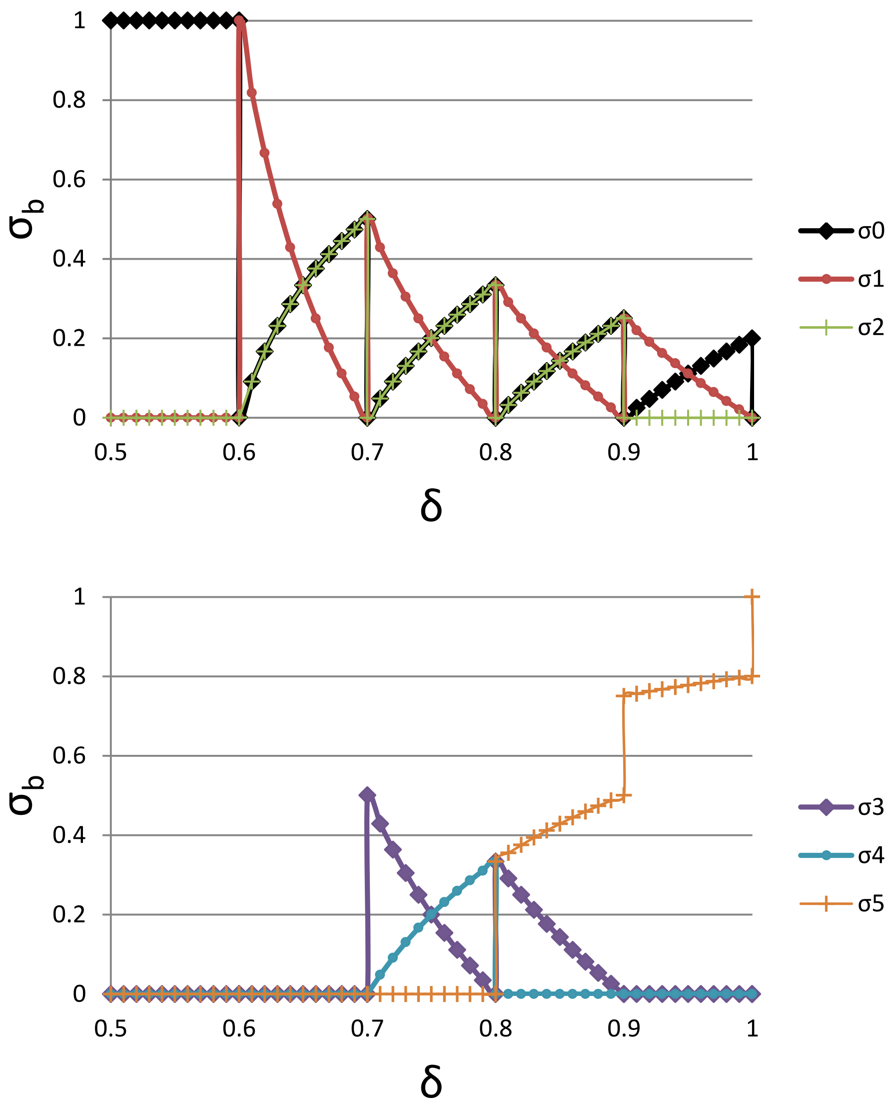

FIGURE A1. EQUILIBRIA FOR TREATMENT KEEPBOTH.

Notes: The figure displays the probability σ_b that a bribe b is chosen in equilibrium as a function of the referee's allocation rule δ. Here, the referee's allocation rule specifies the fraction of times the referee chooses the highest bribe as the winner.

A9. Additional Figures

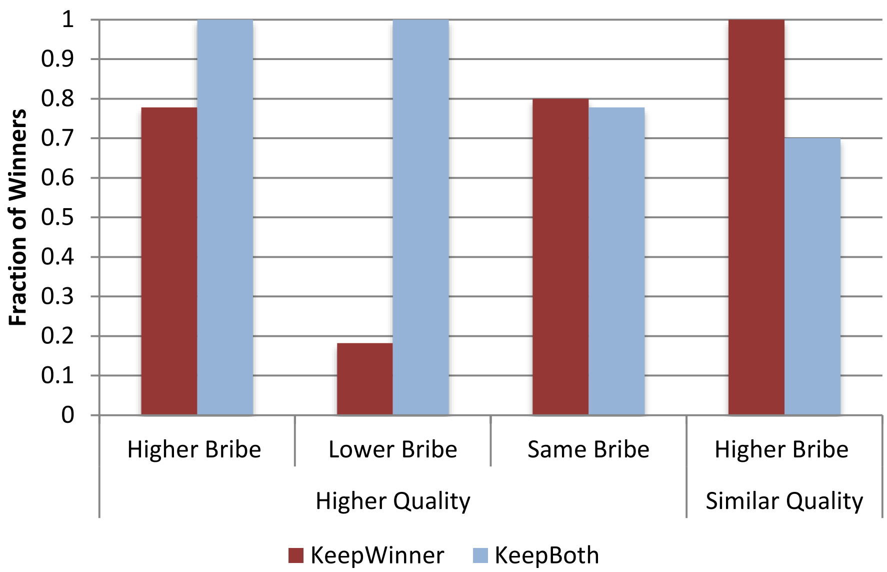

FIGURE A2. WIN CHANCE FOR KEEPWINNER AND KEEPBOTH
Notes: Workers are classified as having a better rating when at least 65.1% of independent raters agree their joke is better.

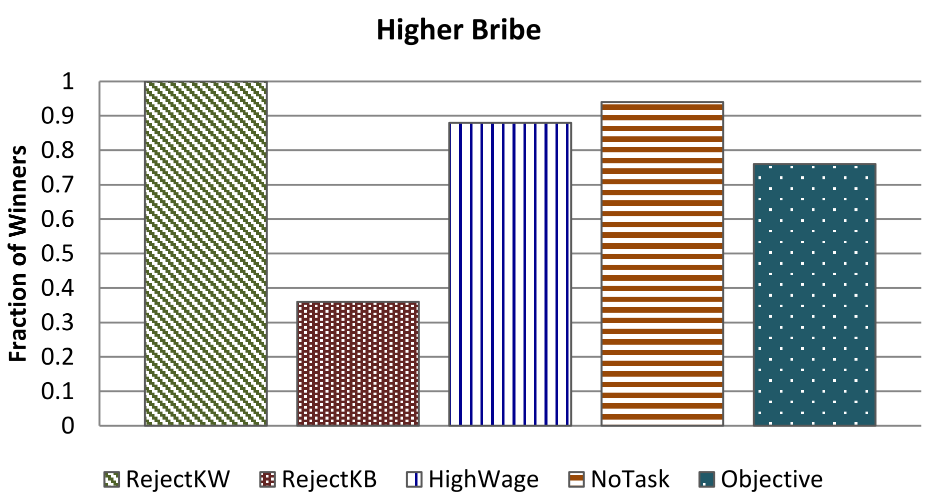

FIGURE A3. WIN CHANCE WHEN HAVING THE HIGHER BRIBE FOR THE ADDITIONAL TREATMENTS

Notes: The figure displays the fraction of cases referees let the higher bribe win in the six additional treatments. RejectKW and RejectKB refer to the KeepWinnerReject and KeepBothReject treatment respectively.

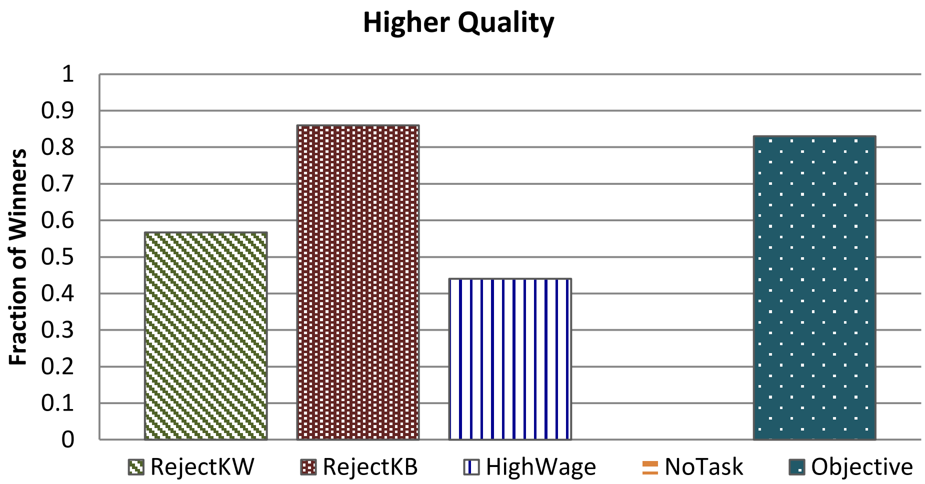

FIGURE A4. WIN CHANCE WHEN HAVING THE BEST JOKE FOR THE ADDITIONAL TREATMENTS

Notes: The figure displays the fraction of cases referees let the higher bribe win in the six additional treatments. RejectKW and RejectKB refer to the KeepWinnerReject and KeepBothReject treatment respectively. Workers are classified as having a higher quality when at least 65.1% of independent raters agree their joke is better.

Appendix B FOR ONLINE PUBLICATION

B1. Worker Instructions

Welcome to today's experiment.[9] Please read the instructions carefully. If you have any questions, please raise your hand and one of the experimenters will come to your desk to answer your question.[10]

You have been assigned to the role of Participant A. For the remainder of the experiment you will be matched with two other participants: Participant B and a Referee. The Referee will now be moved to a different room.

On your desk you can find an envelope with 10 dollars. This is your show-up fee for taking part in this experiment. Both you and Participant B have received a $10 show-up fee whereas the Referee has received a $5 show-up fee. Please do not remove the money from the envelope until you are instructed to.

Both you and Participant B will be asked to work on a task for two rounds. The task will be explained below. After each round the Referee will decide whether you or Participant B performed the task better. The Participant that performed better (as decided by the Referee) will receive an additional $10 prize on top of the show-up fee. The other Participant will receive nothing.

You will be matched to the same Referee and Participant B in both rounds. None of you will ever know the identity of the other two participants.

[9] These are the instructions for the KeepWinner treatment. The instructions for treatment Objective are presented below. The instructions for the other treatments are similar to KeepWinner and available upon request.

[10] A horizontal line indicates that participants went to the next screen.

Do you have any questions before we explain the task to you?

Your task:

Your task is to come up with a joke about a certain topic, which will be announced after the instructions. In total, you will have 10 minutes to come up with a joke. The joke can be short or long, a simple one liner or a full anecdote. The experimenter will let you know when you have 5 minutes as well as 1 minute left for the round.

Check-up questions

How much will you earn (in dollars) in a given round if you are the winner?

Who is going to evaluate your task?

True or false: the Referee and Participant B will be the same participants in both rounds of the experiment.

You are now ready to start the experiment. Please raise your hand when you are ready to start the task. Do not proceed to the next page. The experimenter will instruct you to start when the other participants have finished reading the instructions.

Please write a joke about economists. You have 10 minutes to complete the task.

What do you believe is the probability that you wrote better jokes than Participant B?

Please wait while we are printing your joke. After you have received your joke, please put it into the large envelope with the number so it can be handed over to your Referee. You also have the option to add money for the Referee and put it in the envelope with the joke. For this purpose, you can take up to $5 out of the smaller envelope with your show-up fee and put it into the larger envelope together with the joke you wrote.

Participant B also has the option to add up to $5 to the envelope he/she sends to the Referee.

The Referee will be given both your envelope with the joke and the money and Participants B's. He/she will then be asked to read the jokes and decide which one wins. If the Referee chooses your joke, then you will get an additional $10 and the Referee will keep the money you sent him/her. Participant B will get the money he/she sent to the Referee back. If the Referee chooses Participants B's joke, then Participant B will get an additional $10 and the Referee will keep the money he/she sent to him. In this case you will get back the money that you sent to the Referee.

Please raise your hand when the envelope for the Referee is ready. The experimenter will bring it to the Referee in the next room. After the Referee has determined the winner, the envelope will be collected by the experimenter. The envelope will be returned to you after the Referee has finished grading the second round of jokes.[11]

---

[11] Instructions for round 2 started from "please write a joke about economists" onwards and were identical to the instructions for round 1, except that workers were instructed to write a joke about psychologists instead.

B2. Referee Instructions

Welcome to today's experiment. Please read the instructions carefully. If you have any questions, please raise your hand and one of the experimenters will come to your desk to answer your question.

You have been assigned to the role of the Referee. For the remainder of the experiment you will be matched with two other participants: Participant A and Participant B.

Please raise your hand. The experimenter will escort you to a different room.

Welcome to the experiment. On your desk you will find a small envelope with $5. This is your show-up fee for taking part in this experiment; Participant A and Participant B have received a $10 show-up fee for the experiment.

Today your task is to rate the quality of a joke written by Participant A and a joke written by Participant B. You will be matched to the same Participant A and Participant B in both rounds. None of you will ever know the identity of the other two participants.

Both Participants have 10 minutes to write a joke. After Participants A and B have finished their jokes, they will print them and put them in an envelope which will be brought to you by an experimenter.

You will then have 5 minutes to read both jokes and determine who of the two did the best job, i.e. determine the winner. The winner will receive a prize of $10, whereas the loser will receive nothing. Please make sure to indicate the winner by placing a winner card in the winner's envelope and a loser card in the loser's envelope.

You will also be asked to rate the quality of both the winner's and the loser's joke on a scale from 0 to 10 (on the evaluation form).

The envelopes will then be collected by the experimenter and you will be asked to grade a second round of jokes, written by Participants A and B while you were grading.

The envelopes for both rounds will be returned to Participants A and B at the end of the second round.

The topic for the first round will be 'economists', the topic of the second round will be announced to you after you finish grading the first round.

Please wait while Participants A and B finish writing their jokes. If you have any questions in the meantime, please ask the experimenter.

In a few moments you will receive two envelopes containing the jokes written by Participants A and B. To grade their jokes, please indicate you rating for both Participants on the evaluation form on a scale from 0 to 10.

Participants A and B also have the opportunity to add money to their envelope. You have the option to keep the money sent to you by either Participant A or Participant B. If you keep the money of a Participant, he or she will automatically be the winner. The loser's money will then be returned.

After determining the winner, please make sure to indicate the winner by placing a winner card in the winner's envelope and a loser card in the loser's envelope. After five minutes, an experimenter will collect the envelopes. The envelopes will be returned to Participants A and B at the end of the second round.

Please remain patient while we are printing the jokes.[12]

[12] Instructions for round 2 contained the topic of the second round. Otherwise, they were identical to the last page of the instructions for round 1 (from "In a few moments" onwards).

B3. Independent Raters Instructions

Welcome!

In this experiment you will be shown six pairs of jokes. Jokes in each pair will either be about economists or about psychologists. Participants in a previous experiment wrote the jokes in 10 minutes. For each pair of jokes, you will be asked to rate the quality of both jokes and to indicate which one is better.

Please rate the quality of the following jokes about economists (psychologists). Make sure to read both jokes before rating.

What is the quality of this joke? (0-10)

Which one is the best joke? (Joke A, Joke B)

B4. Worker Instructions Treatment Objective

Welcome to today's experiment. Please read the instructions carefully. If you have any questions, please raise your hand and one of the experimenters will come to your desk to answer your question.

You have been assigned to the role of Participant A. For the remainder of the experiment you will be matched with two other participants: Participant A and a Referee.

The Referee will now be moved to a different room.

On your desk you can find an envelope with 10 dollars. This is your show-up fee for taking part in this experiment. Both you and Participant B have received a $10 show-up fee whereas the referee has received a $5 show-up fee. Please do not remove the money from the envelope until you are instructed to.

Both you and Participant B will be asked to work on a task for two rounds. The task will be explained below. After each round the Referee will decide whether you or Participant B performed the task better. Your goal is to complete as many words as possible in 5 minutes.

The Participant that performed better (as decided by the Referee) will receive an additional $10 prize on top of the show-up fee. The other Participant will receive nothing. You will be matched to the same Referee and Participant B in both rounds. None of you will ever know the identity of the other two participants.

Do you have any questions before we explain the task to you?

During each round of the experiment you will be shown a sequence of words. These words will be printed in different colors: yellow, blue, purple, orange, or red. Your task is to indicate the color of each word. Only the colors named correctly will counts towards your total. This task will last for a total of 5 minutes.

You can indicate the color of your choice using the keyboard. The relevant keys are y (for yellow), r (red), p (purple), o (orange) and b (blue). The key-color combinations will also be visible at the bottom of the screen throughout the task. Be aware: if you press any key other than the one corresponding to the correct color, this will not be counted as a correct response. This also holds for keys that do not refer to any color. On the next page you will have the opportunity to practice the task with a sequence of 10 words.

After you finish the task, your score will be printed on a score sheet that will be handed over to your referee. Your score sheet will be similar to the example below. Every color you successfully indicated will be represented by a dot on the score sheet.

Please answer the following questions before proceeding to the next page.

Question 1: How much will you earn (in dollars) in a given round if you are the winner?

Question 2: Who is going to evaluate your task? Participant A, Participant B or the Referee?

Question 3: True or false: the Referee and Participant will be the same participants in both rounds of the experiment.

You are now ready to start the experiment.

Please raise your hand when you are ready to start the task.

Do not proceed to the next page. The experimenter will instruct you to start when the other participants have finished reading the instructions.

What do you believe is the probability that you have a better score than Participant ?

Please wait while we are printing your score sheet.

After you have received your score sheet, please put it into the large envelope with the number so it can be handed over to your Referee.

You also have the option to add money for the Referee and put it in the envelope with the score sheet. For this purpose, you can take up to $5 out of the smaller envelope with your show-up fee and put it into the larger envelope together with your score sheet.

Participant B also has the option to add up to $5 to the envelope he/she sends to the Referee.

The Referee will be given both your envelope with the score sheet and the money and Participant B's envelope. He/she will then be asked to determine which Participant wins. If the Referee decides that you win, then you will get an additional $10 and the Referee will keep the money you sent him/her. Participant B will get back the money he/she sent to the Referee.

If the Referee decides that Participant B wins, then Participant B will get an additional $10 and the Referee will keep the money he/she sent to him. In this case you will get back the money that you sent out to the Referee.

Please raise your hand when the envelope for the Referee is ready. The experimenter will bring it to the Referee in the next room.

You are now ready to start round 2. This round will be similar to round 1: you will again have to indicate the color of a sequence of words and the task will again be graded by the Referee. Please remember that you will be matched to the same Referee and the same Participant as before. You will again have the option to send money to the referee after you finish your task.[13]

[13] Instructions for round 2 were the same as round 1, starting from "what do you believe is the probability …".

B4. Referee Instructions Treatment Objective

Welcome to the experiment. On your desk you will find a small envelope with $5. This is your show-up fee for taking part in this experiment; Participant A and Participant B have received a $10 show-up fee for the experiment.

Today your task is to determine the score of Participant A and Participant B on a task. You will be matched to the same Participant A and Participant B in both rounds. None of you will ever know the identity of the other two participants.

Participant A and B's task is to determine the color of a series of words displayed on their computer screen. The participants will be shown a sequence of words one at the time and they will have to indicate the colors of the words. Their goal is to complete as many words as possible in 5 minutes. A screenshot of the task has been provided to you on a separate sheet.

Both Participants have 5 minutes for the task. After 5 minutes, their scores will be printed on a score sheet and each one of them will get his or her own printout. The printout score sheet will be similar to the sample score sheet provided to you as an example. Each color successfully determined by the participants will be represented by a single dot on the score sheet. Each participant will then put his/her score sheet in an envelope that will be brought to you by the experimenter.

You will then have 5 minutes to determine the winner. The winner will receive a prize of $10, whereas the loser will receive nothing. Please make sure to indicate the winner by placing a winner card in the winner's envelope and a loser card in the loser's envelope.

The envelopes will then be returned to Participants A and B and you will be asked to grade a second round of score sheets representing the number of colors successfully indicated by Participants A and B while you were grading.

Please wait while Participants A and B complete the first round. If you have any questions in the meantime, please ask the experimenter.

In a few moments you will receive two envelopes containing the score sheets of Participants A and B.

Participants A and B also have the opportunity to add money to their envelope. You have the option to keep the money sent to you by either Participant A or Participant B. If you keep the money of a Participant, he or she will automatically be the winner. The loser's money will then be returned.

After determining the winner, please make sure to indicate the winner by placing a winner card in the winner's envelope and a loser card in the loser's envelope. After five minutes, the experimenter will collect the envelopes and return them to Participants A and B in the other room.

Please remain patient while we are printing the score sheets.

Please wait while Participants A and B are finishing the second round. After Participants A and B have finished the second round, the procedure will be similar to round 1.

You will again receive two envelopes containing the score sheets of Participants A and B.

Participants A and B also again have the opportunity to add money to their envelope. You have the option to keep the money sent to you by either Participant A or Participant B. If you keep the money of a Participant, he or she will automatically be the winner. The loser's money will then be returned.

After determining the winner, please make sure to indicate the winner by placing a winner card in the winner's envelope and a loser card in the loser's envelope. After five minutes, an experimenter will collect the envelopes and return them to Participants A and B in the other room.

Please remain patient while Participants A and B are finishing the second round.

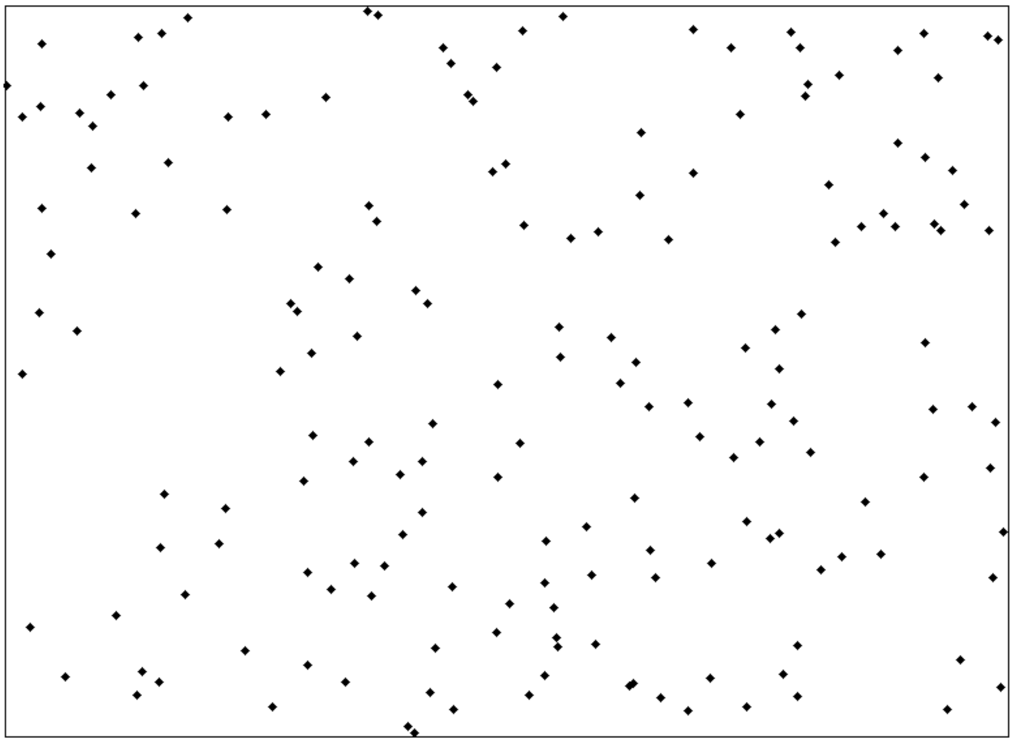

FIGURE B1. SAMPLE SCORE SHEET FOR TREATMENT OBJECTIVE

Notes: the figure was used as an example for participants in the experiment. It was generated using the same procedure as the score sheets used in the experiment, with every dot representing one correct response on the task.

Appendix C Examples of Jokes FOR ONLINE PUBLICATION

In this section, we present 9 examples of jokes written by participants in the experiment. The following jokes are the three best, the three worst and the three median jokes (as determined by the independent raters). All other jokes are available upon request.

C1. Good Jokes

A psychologist was conducting a group therapy session about addictions and obsessions, in which four mothers and their children were participating. Each of the mothers was asked by the psychologist to share their obsession as well as their kid's names. The first mother said, "I am obsessed with eating, and my daughter's name is Candy." The second mother said, "I am obsessed with money, and my daughter's name is Penny." The third mother said, "I am an alcoholic, my daughter's name is Brandy." The fourth mother got up, took her son by the hand, and whispered in his ear, "Come on, Dick, let's go home." (Average Rating 8/10)

A man is at the library and is trying to find an open seat to study. He finds an open spot next to an attractive young woman and asks if he can sit there. She responds rather loudly, "NO, I DON'T WANT TO SPEND THE NIGHT WITH YOU!" Everyone in the library turns to stare at the man. Embarrassed by the attention, the man goes on and finds another spot to study. Later, the same young woman goes up to the man and tells him, "I'm a psychologist and I study social behavior. I know I made you feel embarrassed, right?" The man looks up and responds rather loudly, "$200 DOLLARS JUST FOR ONE NIGHT?! THAT'S TOO MUCH!" Everyone in the library turns to stare at the young woman this time. The man then proceeds to tell her in a subdued voice, "I'm a lawyer. I know how to make people feel guilty." (Average rating 7.63/10)

Why did the psychologist get kicked out school? The professor caught him committing Freud (Average rating 6.6/10)

C1. Median Jokes

A psychologist, an economist, and a physicist were asked for their professional input on ways to improve execution by guillotine. The physicist said "To make the execution less painful, the blade should be heavier because then the blade will travel more quickly and kill the victim sooner". The economist said "The blade shouldn't be cleaned in between executions, because then you can save the cost of cleaning supplies. They're going to die anyway, so sterilizing the blade isn't an ethical concern". The psychologist said "How do we know how much pain the person is in to make it less painful? I think we need more trials, but that's not possible because people only have one life to live. We should use cats! They have nine lives!" Everyone else decided to use the psychologist for a trial because they all owned cats. (Average Rating 3.5/10)

Economists. What my friends think I do: sit back and stare at money. What my parents think I do: earn money. What my colleagues think I do: scam money and help with money laundering. What the academics think I do: create awesome financial theories and win the Nobel Prize. What the public think I do: nothing. What I really do: look at lines and graphs all day long (Average Rating 3.5/10)

Three psychologists are looking up at the stars. The first, a Freudian, sees the Big Dipper in close proximity to Orion's belt and understands instantly the sexual frustration nestled there. The second, a social psychologist, scoffs, and asks the first what Cassiopeia's Little Dipper means, then. He sees the sky in aggregate, a multitude of decision-making stars cohering to a wider social contract. The third is silent. "Hey, Silence of the Lambs," the Freudian psychologist calls out. "Who's right about the stars?" Number Three, an abnormal psychologist, is rather convinced that the stars are, in fact, a 1970s construct remnant from the Star Wars campaign, part of a government conspiracy, and also happen to be transmitting this very conversation to (secretly) Red Russia. Then a goat comes along and speaks. None pay it any heed. (Average Rating 3.5/10)

C1. Bad Jokes

One economist one day went to the shooping centre to buy a keyboard, the price labbled on the hat was $59.99. While the keyboard is using solar as its battery, he start to computer the profit he can get from the keyboard. Since the solar keyboard is much expensive than the normal one, he think that he can use it 3 years, and if he uses the normal keyboard the battery is ... As he thinking, here is a college student came to the store, he bought the keyboard without thinking, and the solar keyboard is out of stock! (Average Rating 1.1/10)

One day the economics was walking beside the beach and began to wonder what the white coloring was up ahead. / Once the economics had reached it then it suddenly had the realization that it was the face of mist. Economics decided to walk into it, and the mist decided to walk through economics as well....and what became of them after that?...That was when economists began their journey! (Average Rating 1.1/10)

A man asked a Psychologist, "Sir, I had a dream that I was swinging on a tree swing like the one in my old house when I younger. The trees were sturdy Oaks that were many years old, and I remember their branches being slightly gnarled. I remember the sun peeking through the leaves and my mother called my name, but as I was running to reach her, the ground opened up and swallowed me. What can this mean?" The Psychologist looked at the man with a furrowed brow, leaned his head back and stared at the ceiling. "Well, the branches mean you are very sexually repressed as they blocked your view of the sun clearly, and that the woman of your dreams is your mom. Clearly you have an Oedipus complex, as the ground breaking up is a sign of your father stopping you from gaining access to your mother. All in all, you love your mom and need to kill your dad." The man blinked a few times, then stood up. "You made that up didn't you?" he asked the Psychologist. With a hearty sigh the Psychologist sat up straight and looked the man dead in the eyes. "Yup." (Average Rating 1.2/10)

Discussion Papers of the Research Area Markets and Choice 2016

Research Unit: **Market Behavior**

David Danz, Steffen Huck, Philippe Jehiel SP II 2016-201
Public statistics and private experience:
Varying feedback information in a take-or-pass game

Jana Friedrichsen SPII 2016-202
Signals sell: Designing a product line when consumers have social image concerns

Uri Gneezy, Silvia Saccardo, Roel van Veldhuizen SPII 2016-203
Bribery: Greed versus reciprocity

Research Unit: **Economics of Change**

Armin Falk, Nora Szech SP II 2016-301
Pleasures of skill and moral conduct

Thomas Deckers, Armin Falk, Fabian Kosse, Nora Szech SP II 2016-302
Homo moralis: Personal characteristics, institutions, and moral decision-making

Jenny Simon, Justian Mattias Valasek SP II 2016-303
The political economy of multilateral aid funds

Vittorio Bassi, Steffen Huck, Imran Rasul SP II 2016-304
A note on charitable giving by corporates and aristocrats:
Evidence from a field experiment

Ludwig Ensthaler, Steffen Huck, Johannes Leutgeb SP II 2016-305
Games played through agents in the laboratory – A test of Prat & Rustichini's model

Research Professorship: **Market Design: Theory and Pragmatics**

Kevin McLaughlin, Daniel Friedman SP II 2016-501
Online ad auctions: An experiment