

Is Dishonesty Contagious?

Robert Innes

and

Arnab Mitra*

RRH: INNES & MITRA: IS DISHONESTY CONTAGIOUS?

ABSTRACT

When an individual believes that peers are predominantly untruthful in a given situation, is he/she more likely to be untruthful in that situation? We study this question in deception experiments patterned after Gneezy (2005) and conducted in Arizona, California, and India. We find evidence that dishonesty is indeed contagious.

Robert Innes
Professor of Economics
Tony Coelho Chair of Public Policy
School of Social Sciences, Humanities and Arts
University of California
Merced, CA 95344
Email: rinnes@ucmerced.edu
Phone: (209) 228-4872

and

Arnab Mitra
Ross School of Business and
School of Natural Resources and Environment
University of Michigan
Ann Arbor, MI
Email: arnabm@umich.com

I. INTRODUCTION

The importance of individual honesty and trustworthiness in economic interactions is well known. These attributes facilitate cooperative relationships, enable contracts, strengthen legal and regulatory institutions, and as a result, promote economic growth (Zak and Knack, 2001; Guiso, Sapienza and Zingales, 2004). Also well known are vast differences in these attributes across cultures and countries. Figure 1 illustrates these differences, showing proportions of world population and world economic activity (respectively) that derive from countries with high, medium, and low levels of corruption, as measured by Transparency International's 2005 corruption perception index (CPI). Without reading too much into these coarse numbers (which, of course, raise complex questions of cause and effect), we note a stylized fact: The distribution of corruption is largely bi-modal, with the vast majority of both population and economic activity in either the low CPI (advanced developed) or high CPI (Third World and transition) countries.¹

<FIGURE 1 ABOUT HERE>

In this paper, we explore a possible contributing explanation for this phenomenon that is rooted in individual preferences. Specifically, we conjecture that honesty is *contagious* in the following sense: If a majority of one's peers are perceived to be honest, an individual is likely to suffer a larger aversion penalty / disutility when behaving dishonestly. If so, honesty breeds honesty and dishonesty breeds dishonesty. Such responses can push countries and cultures toward either predominantly honest or predominantly dishonest behavior, once tipped in one direction or the other. They can thus help explain (or reinforce other explanations for) bimodal outcomes akin to those in Figure 1.² However, they do not tell us why one country or culture goes in one direction – why one culture is “more honest”

and another “less honest.” Moreover, they suggest that behavior is malleable in that, if perceptions of predominating behavior change (from dishonest to honest, for example), individual choices conform. As a result, the bimodal-type outcomes predicted by contagion are potentially fragile; if perceptions of behavior change, then behavior changes with them.

We study the contagion conjecture in the context of a simple deception experiment, wherein we stimulate different subject perceptions of the propensity for honesty in the overall group of experimental subjects. We then examine the resulting impact on an individual’s choice between truthful and untruthful behavior. Our experiments mimic the original deception game designed by Uri Gneezy (2005), who studied the effects of different payoffs on individuals’ aversion to untruthful behavior. Unlike Gneezy (2005), we consider a single set of payoffs in each experiment and focus on the possibility of contagion.³

Because a central motive for our inquiry is to study whether perceived norms of honesty or dishonesty among peers spur more truthful or less truthful conduct in societies that are alternately considered corrupt or not corrupt, we conduct our experiments in both a low CPI country (the U.S.) and a high CPI country (India). Broadly speaking, we find evidence for contagion in both countries.

To our knowledge, the only study that (indirectly) addresses the question of contagion in honesty is Fisman and Miguel’s (2007) famous paper on the tendency for diplomats to garner parking tickets in New York; they find that the immunity-protected foreigners take their home country propensities for lawlessness with them. While these results might be interpreted as evidence against contagion (because diplomats seem to ignore U.S. values in their behavior), we believe that such inferences are misplaced for two reasons. First, there is no *ceteris paribus* in this comparison; diplomats may well temper their lawless behavior,

relative to what they would do if protected by immunity in their home countries. Second, the empirical observation may be a reflection of different relevant peer groups for diplomats from different countries, consistent with the contagion hypothesis. We therefore offer a direct test of contagion in this paper.

Our study is related to recent experimental economics literature on the effects of social information on behavior and a large psychology literature on conformity (see Cialdini and Goldstein, 2004, for a review), but is distinguished from this work by its focus on subjects' truthfulness – our research question. Unlike a significant subset of the literature (but not all), our design also voids prevalent theoretical explanations for conformity and obeying social norms, including social sanctions for the violation of norms (see Fehr and Fischbacher, 2004), incentives to obtain social esteem (Bernheim, 1994), and benefits from others' information (Banerjee, 1992). In our experiments, individual actions are unobservable to anyone other than the individual; there is no possibility of social sanction or building social esteem; and what others do has no bearing on the payoff consequences of individual decisions.⁴

Perhaps most closely related to the present paper is work on social information in dictator games.⁵ In Bicchieri and Xiao (2009), dictators are given different information about the proportion of subjects in a prior session who *were* “fair” vs. “selfish,” and who believe dictators *should be* fair vs. selfish. Their results generally suggest that fairness in actions is contagious. Krupka and Weber (2009) expose dictators to a sample of four (fair vs. selfish) allocations of prior dictators and find a significant increase in the fraction of fair allocations when the sample share is (3/4) or 1, vs. (1/2) or less. Cason and Mui (1998) find that exposing dictators to one prior dictator allocation decision (vs. irrelevant information)

reduces their propensity for selfish allocations. Duffy and Kornienko (2009) show that introducing a tournament that alternately ranks subjects' givings or earnings, significantly promotes generous and selfish allocations respectively; for our purposes, these results could be interpreted as dictators' response to a norm revealed by the choice of tournament. As in our experiments, subjects' actions in these studies are private; there are no social sanctions or rewards; and what others do is irrelevant to payoffs.⁶

The major difference in our experiments is the focus on deception rather than dictator games. This distinction, we believe, is central. Selfish behavior (as in a dictator game) and dishonest behavior (as in a deception game) are very different phenomena. While "fairness" may help promote cooperative relationships (like honesty), "selfish" and acquisitive impulses can promote effort and innovation that are at the core of a thriving market economy. In contrast, the negative consequences of dishonesty and corruption for economic prosperity are well documented. Perhaps for this reason, the culture and psychology of the two phenomena are also different. Selfishness is sometimes heralded as a symptom of the drive to compete and win, as in a game, but in other contexts, scorned as an impediment to cooperative relationships. Honesty, on the other hand, is consistently promoted as a value and virtue by church and community, suggesting that contagion may be less likely. Rode (2010), for example, finds that subjects' honesty is insensitive to competitive vs. cooperative priming.

A more subtle difference between the two phenomena is also important: Contagion in dictator games (as in Bicchieri and Xiao, 2009) can be explained by other-regarding preferences of the type characterized by Fehr and Schmidt (1999), where subjects are averse to inequities in payoffs between themselves and other players in the experiment. The reason is that a higher propensity for fairness among an experiment's population of dictators (a)

reduces a “fair” dictator’s likelihood of earning less than other dictators and (b) raises a “selfish” dictator’s likelihood of earning more than other dictators. Both effects raise inequity penalties of “selfishness” and thereby promote choice of a fair allocation. In contrast, Fehr-Schmidt preferences imply the *opposite* of contagion in our U.S. deception experiments: Incentives for honesty *fall* with the perceived propensity for honesty in the reference population of subjects.⁷ This suggests that mechanisms underpinning contagion in the two games (dictator vs. deception) are different. This distinction is highlighted when we control for our treatment effects on dictator outcomes in our experiments; we find no effect of our “truthfulness” treatments on dictator choices, but pronounced effects on honesty (Sections II and III).

In sum, honesty (in deception games) and fairness (in dictator games) are different phenomena, and results from dictator experiments cannot be translated to deception experiments; indeed, even small framing differences in dictator games are known to have significant effects on behavior, as shown by List (2007) and Bardsley (2008).⁸

II. EXPERIMENTAL EVIDENCE FROM THE U.S.

The Arizona Classroom Experiment

To elicit honest or dishonest decisions from subjects, we closely follow the deception game designed by Uri Gneezy (2005). In this game, there are two possible payoff distributions for each pair of players, with each pair comprised of a “Sender” and a “Receiver.” The two distributions are represented by Options A and B. Only the Sender is informed about the payments associated with the two options, one of which is advantageous to the Sender and the other of which is advantageous to the Receiver. The Sender sends one of two messages to the Receiver:

Message A: “Option A will earn you (the Receiver) more money than Option B.”

Message B: “Option B will earn you (the Receiver) more money than Option A.”

A message is *truthful* if it truthfully indicates the option that is advantageous to the Receiver. After receiving the message chosen by the Sender, the Receiver chooses an option, which then determines payments. Both players are fully informed about the rules of the game, but Receivers are never informed about the specific monetary consequences of either of the two options.

In our experiment, we focus on a single set of payment options (while randomly varying the A/B labels attached to the two options). In one, the Sender receives \$6 and the Receiver obtains \$3, while in the other, the Sender receives \$4 and the Receiver obtains \$6.⁹

Our objective is to study how different perceptions of the truthfulness of other Senders affects Sender behavior. To do this, we use a between-subjects design where we expose different groups of Senders to different treatments designed to alter perceptions of other Sender behavior. There are different ways to provide this treatment information. In

our California experiments, subjects drew five Sender messages (Truthful or Untruthful) from a box containing actual Sender message choices made in a prior experiment. In Arizona, we exposed Senders to summary statements about the propensity for truthfulness in a (non-random) sample of prior Sender messages. In the control treatment, given to an initial session of subjects, no information on other Sender behavior was given. Using outcomes from the control treatment, Senders in subsequent sessions were told:

“Out of 20 Sender messages from a past session of this experiment, with identical payment options, $X (=Y\%)$ were UNTRUTHFUL and $(20-X) (= (100-Y)\%)$ were TRUTHFUL.”

Four treatments of this form were considered: $Y=15\%$ (heavily truthful), $Y=40\%$, $Y=60\%$, and $Y=85\%$ (heavily untruthful). In all treatments, the higher percentage was reported first (so that, for example, when $Y=40\%$, the number and percent of truthful messages from past sessions was indicated first). In all cases, there was a (non-random selected) set of 20 messages from the control treatment satisfying the above statement. Our approach is similar to that used in other experimental papers in the social influence literature. Frey and Meier (2004), for example, report two different percentages of past students who contribute to a charity based on different outcomes from a recent semester and, alternatively, a ten year interval. Bicchieri and Xiao (2009) report different shares of “fair” choices (40% and 60%) from a past session and argue that the information is truthful because they can select a past session to satisfy either indicated percentage. We designed our statements to highlight the selection of a subset of Sender messages and were careful not to state or imply that the reported messages represent a general pattern.¹⁰

In all treatments, Senders were given general information on the propensity of Receivers to accept their recommendations. Based on results from Gneezy’s (2005)

experiments (where 78 percent of Receivers followed the Sender recommendations), we told all Senders the following:

“In past experiments like this one, roughly 8 out of 10 Receivers chose the Option recommended by their Senders.”

Receivers were not given this information, and Senders were so informed. To verify that Senders generally believed that Receivers would accept their recommendations, we followed Gneezy’s (2005) approach, asking them to predict their Receiver’s choice and paying them for a correct prediction. Overall, 73.4% of Senders predicted that their Receiver would accept their recommendation.¹¹ These results indicate that Senders generally expect their recommendations to be followed; hence, their choices reflect a concern for the “fairness” / morality of lying, and not strategic motives. As it turned out, 73 percent of our Receivers followed their Sender recommendations.

The experiment was conducted in undergraduate economics classes at the University of Arizona in Spring, 2008 and Spring, 2009. In total, there were 233 Sender/Receiver pairs. Receivers were in different classes than any of the Senders. Anonymity of all participants was ensured by identifying subjects with a randomly assigned identification number that was also used to match Senders to Receivers. Class Rosters were used to ensure that no student participated more than once.¹² The experiment took approximately 10 minutes to run. Subject participation was purely voluntary. Subjects were instructed to communicate only with the experimenter and were carefully monitored to this end. Control treatments were run in each Sender class to control for any potential individual course effects.

Table 1 reports the number of Senders exposed to each of the different treatments, and summarizes our results. Table 2 reports results of a probit regression of truthful ($y=1$) vs. untruthful ($y=0$) choices on the treatments and individual class fixed effects. Relative to

the control, the “heavily untruthful” treatment (Y=85%) has a significant impact on subjects’ propensity to be truthful: the proportion of untruthful messages rises from 41 percent (in the control) to approximately 81 percent (under the treatment), almost doubling. Other treatments have no significant effect. *We thus find evidence of contagion in the sense that a strong propensity for untruthfulness is contagious.*

<TABLES 1 AND 2 ABOUT HERE>

Sutter (2009) documents the importance of “sophisticated deception,” when a Sender tells the truth anticipating that his recommendation will not be followed. He proposes a different measure of deception that includes sophisticated deceivers and excludes “benevolent liars” who lie anticipating that their Receiver will reject their recommendation. In our experiment, the treatments concern raw actions of other subjects (truthful vs. untruthful); we are therefore also principally interested in effects on raw actions (as reported above). The conclusions of Table 1 nonetheless extend to Sutter’s measure of sophisticated deception, although the impact of our treatment is attenuated; the proportion of Sutter-truth-tellers is 55.7 percent under the control and 30.8 percent under the heavily untruthful (Y=85%) treatment, a difference that is statistically significant ($z=3.06$).¹³

Criticisms and the California Classroom Experiment

In principle, other effects discussed in the literature might be at play in our experiments.¹⁴ First is the potential for experimenter demand effects, with subjects trying to do what the experimenter appears to want them to do (see Duffy and Kornienko, 2009, for an excellent discussion). We sought to avoid any such effects by ensuring anonymity and no communication to subjects about the rather oblique intent or purpose of the experiment. In Arizona, there was also no significant impact of three of the four (Y=15%, 40%, 60%)

treatments; were there an experimenter demand effect, these treatments would be expected to influence behavior. We nevertheless seek to allay this concern with an alternative design in the California experiment.

Second, could our results be due to the effect of our treatments on generosity, as opposed to their effect on subjects' aversion to lying? Perhaps when subjects learn that other Senders are more untruthful, their preference for the "selfish" (6-3) option, vs. the "generous" (4-6) option (under the Arizona payoffs) rises; if so, the propensity for untruthful message choices will rise, even absent any effect on subjects' aversion to lying. We would like to distinguish which channel of effect explains the contagion that we observe.

To address these issues, our California experiment modified the Arizona design in three ways: (1) Subjects all played both a deception game and a dictator game, with a coin flip determining which game determined payoffs; (2) we elicited Sender beliefs about the proportion of truthful Senders in the experiment; and (3) we used a different design for the treatments: Rather than reading a statement on outcomes from a non-random sample of prior messages, Senders themselves each drew five Sender messages from a box containing all Sender messages (truthful or untruthful) sent in a prior Arizona experiment.¹⁵

The first change enables us to investigate the effect of information about other Senders' truthfulness on both preferences over allocations ("generosity") and lying aversion. The second enables us to examine explicitly the impact of Sender beliefs about other Senders' truthfulness on their own actions (truthfulness). And the third mitigates the potential for experimenter demand effects.

The payoff options in the California experiment were the same as in Arizona (6-3 and 4-6). All Senders were given the same information about the 80 percent propensity for

Receivers to accept their recommendations, and Receivers were given the same (negligible) information as in the Arizona game. All subjects played both deception and dictator games.¹⁶ In the dictator game, Senders simply chose one of the two payoff options (6-3 or 4-6). Following Gneezy (2005), the Sender-chosen option was realized with 80 percent probability, and the other option with 20% probability, in order to mimic the deception game wherein Receivers accept their Sender recommendations with (approximately) 80 percent probability. The Sender instructions conveyed this probabilistic selection. If the Treatments – the number of Truthful message draws, ranging from zero to five out of five – were to affect Sender generosity, we would expect to observe Dictators choosing the more generous (4-6) option more often when they obtained a higher share of Truthful message draws.

We elicited Sender beliefs about the proportion of Truthful Senders in their experiment by asking the following: “What proportion of Senders in this class do you think will send Truthful messages? CIRCLE ONE OF THE FOLLOWING PERCENTAGES. If your prediction is correct (within five percentage points of the actual choice, plus or minus), you will receive an additional \$1 payment.” Senders were given twenty five-percentage point bands from which to choose (0-5, 5-10, etc.).

The experiment was conducted in undergraduate economics classes at U.C. Merced in the Spring of 2010. As always, subjects were completely anonymous; there was no communication allowed; there were no class overlaps; and treatments were randomly assigned. There were 105 Sender/Receiver pairs, 26 Senders in the Control (no message draws) and 79 in the “message draw” Treatments.

Table 3A reports raw results from the Deception game. Roughly 58 percent of Control subjects were truthful, almost exactly the same proportion as in Arizona. Subjects

who drew two or fewer Untruthful messages (less than half) exhibited no significant difference in their propensity for truthfulness compared with Control subjects. However, subjects who drew three or more Untruthful messages (more than half) exhibited a significantly lower propensity for truthfulness than their Control subject counterparts. Subjects who drew two or fewer Untruthful messages also revealed average beliefs about the fraction of Truthful Senders that were significantly higher than for either the Control subjects or subjects who drew three or more Untruthful messages (see Average Sender Belief in Table 3A). Hence, the experiment provides evidence that the message draws affected Sender beliefs about other Senders' truthfulness in the predicted direction.

<TABLE 3 ABOUT HERE>

Table 3B reports raw results from the Dictator game. Subjects from the “message draw” treatments exhibited no significant differences in their propensity to choose the “Selfish” (6-3) option, when compared to the Control subjects. Contrary to the conjecture that a heavily untruthful treatment would prompt more selfish preferences and thus explain a higher likelihood of an untruthful message choice, subjects who drew *few* Untruthful messages (two or less) were *more* selfish on average (but not significantly more).

These conclusions are reinforced by probit regressions of Sender message and Dictator decisions, as reported in Table 4. Several specifications are reported, using different measures of the treatments. In the last specification, we gauge the impact of Sender beliefs about the proportion of truthful Senders on their message (and Dictator) decisions. Due to the potential for endogeneity between these expressed beliefs and actual decisions, we instrument the belief variable with a dummy for “highly untruthful” (three or more U) message draws; the instrument performs well in the first stage (in the predicted direction)

and, as expected, the instrumented belief variable has a significant positive effect on subjects' propensity for truthfulness and no significant effect on subjects' propensity for selfishness. Direct effects of untruthful treatments (number of U draws, or dummies for a high number of U draws) are also estimated to have a significant negative effect on subjects' propensity for truthfulness. In some cases, these treatments have a significant impact on "selfishness," but this effect is not robust and is negative, contrary to the conjecture that "untruthful" treatments engender more selfish preferences.

<TABLE 4 ABOUT HERE>

Table 5 provides one last examination of whether treatment effects on preferences over allocations (generous vs. selfish) can explain the contagion that we observe in subjects' propensity for truthfulness. We present difference in difference statistics (propensities for untruth minus propensities for selfishness, treatment minus control) for the different treatments. Consistent with our raw results (Table 3A), we find that subjects' *net* excess propensity for untruth vs. selfishness is significantly higher for subjects exposed to the heavily untruthful treatments (with 4 or 5 U draws) than for Control subjects.

In sum, we again find support for the contagion hypothesis in the sense that a strong propensity for untruthfulness is contagious. In addition, we find that this contagion cannot be explained by treatment effects on preferences over allocations.

<TABLE 5 ABOUT HERE>

III. EXPERIMENTAL EVIDENCE FROM INDIA

In our U.S. experiments, we find that information indicating a strong peer propensity for dishonesty promotes untruthful behavior. A central motive for our work is to study the potential for contagion in the other direction: In a country where corruption is high, and a

propensity for dishonesty presumably correspondingly high, can subjects be spurred to more truthful conduct by information suggesting a norm of honesty? India – with a corruption index in the highest tier of countries – is arguably an ideal country in which to examine this question.

In the Spring of 2009 and Summer of 2011, we conducted a deception experiment with a set of 131 Sender/Receiver pairs of undergraduate students at Jadavpur University in Calcutta. Like most university experimental labs, Jadavpur maintains a roster of willing experimental participants and regularly announces opportunities for participation in experiments. Our announcements were made in English, and were only made in Departments where English fluency is required (most of the University, excepting the Bengali major). In the experiment, the following two payoff options were posed:

Option A: 160 Rupees to you (the Sender) and 160 Rupees to the other student (the Receiver).

Option B: 200 Rupees to you (the Sender) and 100 Rupees to the other student (the Receiver).

As in the U.S. experiments, option labels were varied randomly. The payoffs were designed to (a) have the same ratio of Receiver loss to deceit and Sender gain ($3/2$) as in our Arizona experiment, (b) to meet minimum payment requirements, and (c) to give substantial stakes to the choices made. Although 40 Rupees (the Sender gain from dishonesty and Receiver acceptance) are less than one U.S. dollar, average daily per capita consumption expenditures in India are less than 19 Rupees in rural areas and 35 Rupees in urban areas.¹⁷ Put differently (quoting Fehr, et al., 2008), “Fifty Rupees are roughly equal to a day’s skilled wage.” The stakes in our experiment can therefore be considered substantial in context.

Following our Arizona design, we conducted three treatments: a control with no information about Sender behavior in prior sessions of related experiments, a strongly truthful treatment (Y=15%) in which Senders were given the following information:

“Out of 15 Sender messages from a past session of this experiment here in Calcutta, 13 out of 15 (85%) were TRUTHFUL and 2 out of 15 (15%) were UNTRUTHFUL,”

and a corresponding strongly untruthful treatment (Y=85%). Our initial control responses in Calcutta gave us 15 Sender messages satisfying each of the treatment statements.¹⁸

Table 6 reports results from the Calcutta experiment. We find that the honest treatment leads to a significantly higher proportion of truthful messages than in the control. Under the treatment, the proportion of honest messages is more than fifty percent higher than under the control, 69.2% vs. 44.4%. Under the untruthful treatment, the proportion of truthful messages is a third lower than in the control (28.9% vs. 44.4%); however, the difference is statistically significant only in a one-sided test (one-sided p-value=.06). The proportion of Senders predicting Receiver acceptance is high (75.6%) and the proportion of Receivers accepting their Sender recommendations is also high (70.2%), although less than in Gneezy's (2005) experiments and slightly less than in Arizona.¹⁹

<TABLE 6 ABOUT HERE>

In the Summer of 2011, we also conducted an experiment at Jadavpur University that mimics our California design. In this experiment, each treatment subject (Sender) drew five messages from one of two boxes containing all Sender messages (truthful or untruthful) sent in a prior Calcutta experiment. Results from this experiment are presented in Table 7.

Subjects who drew two or fewer Untruthful messages (out of five) exhibited a significantly higher propensity for truthfulness (63 percent) compared with Control subjects (44 percent), consistent with our prior results. In addition, subjects who drew three or more Untruthful

messages (out of five) exhibited a significantly lower propensity for truthfulness (19 percent) than their Control subject counterparts, giving us stronger evidence of contagion in dishonesty. Subjects who drew two or fewer Untruthful messages also revealed average beliefs about the fraction of Truthful Senders that were significantly higher than for either the Control subjects or subjects who drew three or more Untruthful messages (see Average Sender Belief in Table 7A). Hence, the experiment provides evidence that the message draws affected Sender beliefs about other Senders' truthfulness in the predicted direction. Finally, from Table 7B, we see that our treatments had no affect at all on Senders' propensity for "selfishness" in the parallel dictator game: Whether receiving no messages (the Control), two or fewer Untruthful message draws, or three to five Untruthful message draws, approximately seventy percent of subjects chose the option that paid them more and their Receiver less. Hence, as in the U.S., our treatment effects cannot be explained by changes in preferences over allocations.

<TABLE 7 ABOUT HERE>

IV. DISCUSSION AND CONCLUSION

We conduct our experiments in the U.S. and India in order to understand and compare outcomes in a country with a low Corruption Perception Index (CPI) (the U.S.) and one with a high CPI (India). Our related conjecture is that prevailing norms are "more honest" in the U.S., and "more dishonest" in India. International comparisons of experiments is problematic because payoffs cannot be mimicked; a dollar in Rupees is not the same to a student in India as a dollar is to a student in California, and no attempt to "match up" payments can overcome this critique. However, setting this concern aside, our twin experiments yield results consistent with our initial conjecture, *given* our choice of respective

experimental payoffs. On average, control subjects in Arizona and California are honest with relative frequency 58.5%, while control subjects in Calcutta are honest with relative frequency 44.4%; the z-statistic for the difference is 1.742, significant at the 10 percent level (two-sided p-value = .0815).

In both countries, we find evidence of contagion. In our India experiments, treatment effects go in both directions: Exposed to evidence that a large proportion of subjects is dishonest, individuals are themselves dishonest with greater frequency; conversely, exposed to evidence of predominant honesty, individuals tend to be honest more frequently. The first effect supports the conjecture of self-reinforcing dynamics to a social norm that is predominantly dishonest; these are the dynamics that we argue in the introduction can help explain a “two hump” country-level distribution of social norms. The second effect suggests that individual behavior is quite sensitive to social information that is substantially at odds with average behavior. This is noteworthy. Despite relatively large stakes, our Indian subjects become more honest to their economic detriment in response to a social cue that is contrary to the true norm. We interpret this finding as reason for optimism that the promotion of honest norms can be successful, and quickly so, if perceptions of peer behavior can be changed.

In the U.S., we find statistical evidence of one direction of effect: A strong signal of dishonesty leads to more dishonesty. Recall that, in our U.S. experiments, there is a prevailing norm of honesty with roughly 60 percent of our control subjects choosing to tell the truth. Our results thus suggest that this norm is fragile. Even with small stakes, individual propensities for honesty evaporate when peers are thought to be dishonest.

In the other direction, our U.S. subjects, when faced with social information that there is substantial honesty among peers, exhibit a greater tendency to honesty. However, this difference is not statistically significant. A possible interpretation is that individuals who tend to be dishonest, contrary to prevailing social norms, are less subject to social cues than either those who tend to be honest or those whose tendency to dishonesty is in tune with (and potentially attributable to) prevailing norms. However, we are loath to make any strong statements. The insignificance of these estimated effects, in the U.S., could be an artifact of our experimental payoffs.

In summary, we find evidence that both honesty and dishonesty are contagious in the sense that subjects are more likely to be honest (dishonest) when exposed to information suggesting that other subjects have a higher propensity for honesty (dishonesty). We find these responses in both a high-corruption culture (India) and a low-corruption culture (U.S.). These responses shed some light on population dynamics in truthfulness and corruption that may help to explain societal tendencies to be in one camp or the other, highly honest or highly dishonest. Normatively, they suggest value to a culture of honesty in an organization by indicating the fragility of truthful behavior; even with small stakes, our U.S. subjects flocked to the dishonest course when primed with a social pass-go to do so. Conversely, they suggest promise for countering corrupt impulses in the developing world if perceptions of norms can be reversed.

Of course this begs the question: By what mechanism can norms be changed? Recent findings suggest that this may be tough. Rode (2010), for example, finds that dishonesty is insensitive to cooperative priming, and Fisman and Miguel (2007) find that foreign diplomats do not respond to American values of lawful behavior. However, empirical evidence

indicates that aid and trade can reduce corruption (Tavares, 2003; Gokcekus and Knorich, 2006). Our results suggest a coarse mechanism for this effect, but leave much unanswered. For example, what determines whether “honest norm” partners bend to the norms of “dishonest norm” partners or vice versa? If an “honest norm” agent seeks to trade in a “dishonest norm” country that is relatively closed, then is the honest trader likely to bend to local (dishonest) norms? And how does a subject’s exposure to another country’s norms affect his behavior in his own country? These questions lend themselves to further experimental work that can illuminate not only the nature of the contagion we identify, but also how it can be exploited for positive social ends and what implications it has for one of the key pillars of the globalization debate: benefits of trade in reducing corruption.

REFERENCES

- Acemoglu, D., S. Johnson, and J. Robinson. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*, 91, 2001, 1369-1401.
- Andvig, J. and K. Moene. "How Corruption May Corrupt." *Journal of Economic Behavior & Organization*, 13, 1990, 63-76.
- Ball, S., C. Eckel, P. Grossman, and W. Zame. "Status in Markets." *Quarterly Journal of Economics*, 116, 2001, 797-817.
- Banerjee, A. "A Simple Model of Herd Behavior." *Quarterly Journal of Economics*, 107, 1992, 797-817.
- Banerjee, A., and T. Besley. "Peer Group Externalities and Learning Incentives: A Theory of Nerd Behavior." John M. Olin Discussion Paper no. 68. Princeton University, 1990.
- Bardhan, P. "Corruption and Development: A Review of Issues." *Journal of Economic Literature*, 35, 1997, 1320-46.
- Bardsley, N. "Dictator Game Giving: Altruism or Artefact?" *Experimental Economics* 11, 2008, 122-33.
- Battigalli, P. and M. Dufwenberg. "Guilt in Games." *American Economic Review* 97, 2007, 170-6.
- Berg, J., J. Dickhaut, and K. McCabe. "Trust, Reciprocity and Social History." *Games and Economic Behavior*, 10, 1995, 122-42.
- Bergstrom, T. "Evolution of Social Behavior: Individual and Group Selection." *Journal of Economic Perspectives*, 16, 2002, 67-88.
- Bernheim, B.D.. "A Theory of Conformity." *Journal of Political Economy*, 102, 1994, 841-77.

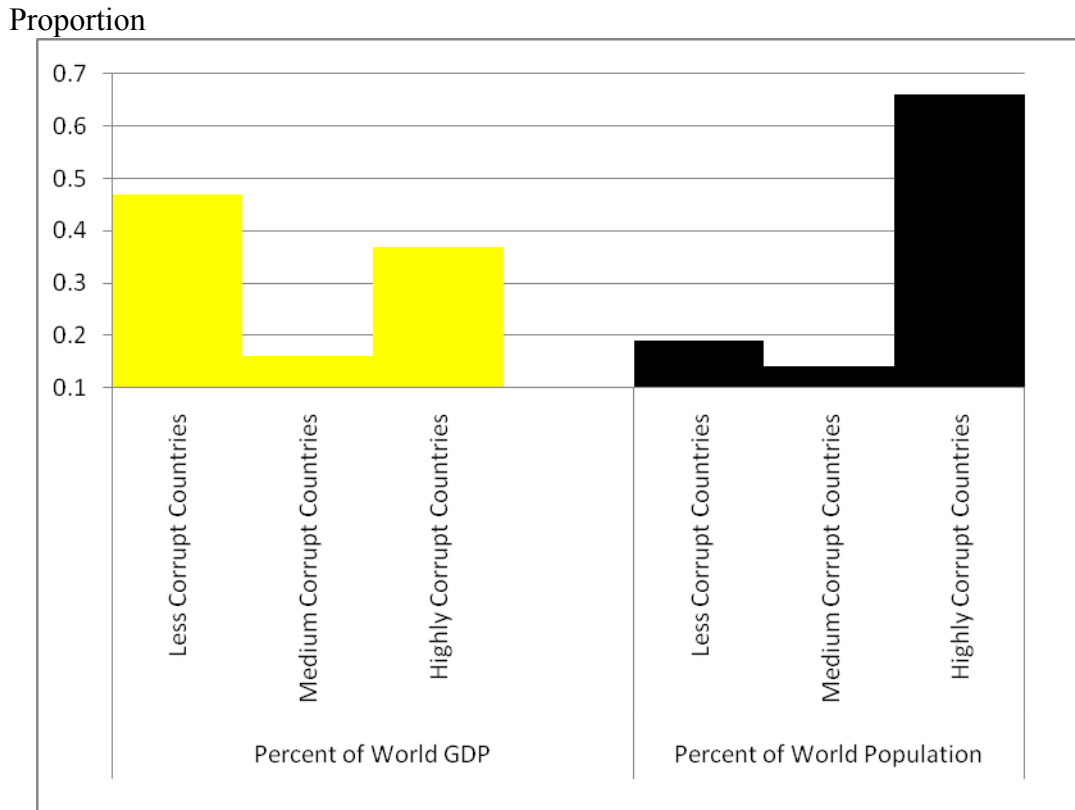
- Bertrand, M., and S. Mullainathan. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review*, 94, 2004, 991-1013.
- Bicchieri, C., and E. Xiao. "Do the Right Thing: But Only if Others Do So." *Journal of Behavioral Decision Making*, 22, 2009, 191-208.
- Binmore, K., A. Shaked, and J. Sutton. "Testing Noncooperative Bargaining Theory: A Preliminary Study." *American Economic Review*, 75, 1985, 1178-80.
- Bohnet, I. and R. Zeckhauser. "Social Comparisons in Ultimatum Bargaining." *Scandinavian Journal of Economics*, 106, 2004, 495-510.
- Bolton, G. and A. Ockenfels. "ERC: A Theory of Equity, Reciprocity, and Competition." *American Economic Review*, 90, 2000, 166-93.
- Bonetti, S. "Experimental Economics and Deception." *Journal of Economic Psychology*, 19, 1998, 377-95.
- Cason, T. and V. Mui. "Social Influence in the Sequential Dictator Game." *Journal of Mathematical Psychology*, 42, 1998, 248-65.
- Charness, G. and M. Dufwenberg. "Promises and Partnership." *Econometrica*, 74, 2006, 1579-1601.
- Charness, G. and M. Dufwenberg. "Deception: The Role of Guilt." Working Paper, UC Santa Barbara and University of Arizona, 2005.
- Chaudhuri, A., S. Graziano and P. Maitra. "Social Learning and Norms in Public Goods Experiment with Inter-Generational Advice." *Review of Economic Studies*, 73, 2006, 357-80.

- Chen, Y., M. Harper, J. Konstan and S. Xin Li. "Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens." *American Economic Review*, 100, 2010, 1358-98.
- Cialdini, R. and N. Goldstein. "Social Influence: Compliance and Conformity." *Annual Review of Psychology*, 55, 2004. 59-621.
- Duffy, J. and N. Feltovich. "Does Observation of Others Affect Learning in Strategic Environments? An Experimental Study." *International Journal of Game Theory*, 28, 1999, 131-52.
- Duffy, J. and T. Kornienko. "Does Competition Affect Giving?" University of Pittsburgh and University of Edinburgh, 2009.
- Dufwenberg, M. and U. Gneezy. "Measuring Beliefs in an Experimental Lost Wallet Game." *Games & Economic Behavior*, 30, 2000, 163-82.
- Eckel, C. and R. Wilson. "Social Learning in Coordination Games: Does Status Matter?" *Experimental Economics*, 10, 2007, 317-29.
- Ederer, F. and E. Fehr. "Deception and Incentives: How Dishonesty Undermines Effort Provision." Discussion Paper No. 3200, IZA, Bonn, 2007.
- Fehr, E., and U. Fischbacher. "Social Norms and Human Cooperation." *Trends in Cognitive Sciences*, 8, 2004, 185-90.
- Fehr, E., K. Hoff and M. Kshetramade. "Spite and Development." *American Economic Review*, 98, 2008, 494-9.
- Fehr, E. and K. Schmidt. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics*, 114, 1999, 817-68.

- Fischbacher, U., S. Gächter and E. Fehr. "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment." *Economics Letters*, 71, 2001, 97-104.
- Fisman, R., and E. Miguel. "Corruption, Norms and Legal Enforcement: Evidence From Diplomatic Parking Tickets." *Journal of Political Economy*, 115, 2007, 1020-48.
- Frey, B. and S. Meier. "Social Comparisons and Pro-Social Behavior: Testing 'Conditional Cooperation' in a Field Experiment." *American Economic Review*, 94, 2004, 1717-22.
- Gneezy, U. "Deception: The Role of Consequences." *American Economic Review*, 95, 2005, 384-94.
- Gokcekus, O. and J. Knorich. "Does Quality of Openness Affect Corruption?" *Economics Letters*, 91, 2006, 190-6.
- Guiso, L., P. Sapienza, and L. Zingales. "The Role of Social Capital in Financial Development." *American Economic Review*, 94, 2004, 526-56.
- Innes, R. "A Theory of Moral Contagion." Working Paper, U.C. Merced, 2009.
- Katz, M., and C. Shapiro. "Technology Adoption in the Presence of Network Externalities." *Journal of Political Economy*, 94, 1986, 822-41.
- Keizer, K., S. Lindenberg, and L. Steg. "The Spreading of Disorder." *Science*, 12, 2008, 1681-85.
- Knez, M. and C. Camerer. "Outside Options and Social Comparison in Three-Player Ultimatum Game Experiments." *Games and Economic Behavior*, 10, 1995, 65-94.
- Krupka, E. and R. Weber. "The Focusing and Informational Effects of Norms on Pro-Social Behavior." *Journal of Economic Psychology*, 30, 2009, 307-20.
- List, J. "On the Interpretation of Giving in Dictator Games." *Journal of Political Economy*, 115, 2007, 482-93.

- List, J., and D. Reiley. "The Effects of Seed Money and Refunds on Charitable Giving: Experimental Evidence from a University Capital Campaign." *Journal of Political Economy*, 110, 2002, 215-33.
- Mo, P. "Corruption and Economic Growth." *Journal of Comparative Economics*, 29, 2001, 66-79.
- Rode, J. "Truth and Trust in Communication: Experiments on the Effect of a Competitive Context." *Games and Economic Behavior*, 68, 2010, 325-38.
- Sanchez-Pages, S. and M. Vorsatz. "An Experimental Study of Truth-Telling in a Sender-Receiver Game." *Games and Economic Behavior*, 61, 2007, 86-112.
- Schotter, A. and B. Sopher. "Social Learning and Coordination Conventions in Intergenerational Games: An Experimental Study." *Journal of Political Economy*, 111, 2003, 498-529.
- Shang, J. and R. Croson. "Field Experiments in Charitable Contribution: The Impact of Social Influence on the Voluntary Provision of Public Goods." *The Economic Journal*, 119, 2009, 1422-39.
- Sutter, M. "Deception through Telling the Truth? Experimental Evidence from Individuals and Teams." *The Economic Journal*, 119, 2009, 47-60.
- Tavares, J. "Does Foreign Aid Corrupt?" *Economics Letters*, 79, 2003, 99-106.
- Treisman, D. "The Causes of Corruption: A Cross-National Study." *Journal of Public Economics*, 76, 2000, 399-457.
- Zak, P. and S. Knack. "Trust and Growth." *The Economic Journal*, 111, 2001, 295-321.

Figure 1. Distribution of World GDP and World Population by Level of Corruption



*Corruption is measured by Transparency International’s Corruption Perception Index. “Less Corrupt” countries are those with CPI values in the top third of the range; “Medium Corrupt” in the middle third; and “Highly Corrupt” in the bottom third. We exclude India and China from the population distribution, but include them in the GDP distribution.

Table 1
Results of Arizona Classroom Experiment

Treatment (Reported Propensity Untruthful Senders)	Number of Subjects	Percent Truthful	z-statistic (Control - Treatment)	Percent Predicting Receiver Acceptance
Control	97	58.8%		74.2%
Y=15%	25	64.0%	-.480	76.0%
Y=40%	26	53.8%	0.455	80.8%
Y=60%	33	54.5%	0.430	63.6%
Y=85%	52	19.2%	5.349***	73.1%
Overall	233	49.3%		73.4%

*** denotes significant at 1% level (two-sided).

Table 2
Probit Regression of Arizona Sender Message Choices (Truthful vs. Untruthful)
with Course Fixed Effects

Variable	Coefficient	z-statistic	Marginal Effect	z-statistic
Constant	0.7004**	2.06		
Y=85% Treatment	-1.0385***	-3.92	-0.3703***	-4.58
Y=60% Treatment	-0.4358	-1.17	-0.1480	-1.25
Y=40% Treatment	-0.5554	-1.44	-0.1873	-1.58
Y=15% Treatment	0.0206	0.06	0.0073	0.06

Note: N=233. ** denotes significant at 5% (two-sided). *** denotes significant at 1%.
 Dependent variable: Sender message choice (truthful=1, untruthful=0).
 We report average marginal effects. The fixed course effects are jointly insignificant, with χ^2 (df=5) test statistic (p-value) 7.43 (.1904).

Table 3
California Experiment: Raw Results

Treatment (Number of Untruthful Draws)

	Control (No draws)	Truthful (0-1-2 U draws out of 5)	Untruthful (3-4-5 U draws out of 5)
Observations	26	27	52
Percent Truthful	57.69%	62.96%	28.85%
z Stat 1 (Cont.-Treat.)	-	-0.3925	2.4980**
z Stat 2 (012-345)	-	-	3.0412***
z Stat 3 (Cont.+012 – 345)	-	-	3.4279***
Average Sender Belief (SD)	28.65% (25.47)	50.28% (25.91)	32.31% (23.22)
z Stat for Belief (012-Ct/345)	3.064***	-	3.027***

A) Deception

B) Dictator

Percent “Selfish”	69.23%	81.48%	65.38%
z Stat 1 (Cont.-Treat.)	-	-1.0436	0.3434
z Stat 2 (012-345)	-	-	1.6144
z Stat 3 (Cont.+012 – 345)	-	-	1.1388

Table 4
Probit Regressions for California Sender Decisions

A) Truthful (y=1) vs. Untruthful (y=0)

	<u>Model</u>			
	1	2	3	4
Constant	.5592* (.3285)	.3309 (.2461)	.1992 (.2048)	-2.1557*** (.6788)
Control (=1 if Control)	-.3652 (.4113)	-.1368 (.3490)	-.0052 (.3212)	.9326*** (.3358)
Treatment:				
Number of U Messages Drawn	-.2611*** (.0967)/[-.1031]			
Dummy for 3-4-5 U Messages		-.8888*** (.3072)/[-.3398]		
Dummy for 4-5 U Messages			-.8930*** (.2961)/[-.3349]	
Sender Belief (about Percent Truthful Senders) Instrumented*				.0495** (.0196)/[.0195]

A) Selfish (y=1) vs. Generous (y=0)

	<u>Model</u>			
	1	2	3	4
Constant	1.3079*** (.3953)	.8958*** (.2799)	.8994*** (.2362)	-.5033 (.7096)
Control (=1 if Control)	-.8055* (.4717)	-.3934 (.3802)	-.3970 (.3494)	.2083 (.3403)
Treatment:				
Number of U Messages Drawn	-.2308** (.1081)/[-.0785]			
Dummy for 3-4-5 U Messages		-.5000 (.3321)/[-.1704]		
Dummy for 4-5 U Messages			-.6207** (.3086)/[-.2169]	
Sender Belief (about Percent Truthful Senders) Instrumented*				.0278 (.0185)/[.0095]

Notes: Number of Observations = 105. Standard errors are in round parentheses. Marginal effects are in square brackets.

* The Sender Belief is instrumented and identified with the treatment dummy for a draw of 3,4, or 5 Untruthful Sender messages. The first-stage F statistic (p-value) for the identifying instrument is 9.57 (.0026).

Table 5
Difference in Difference (California Experiment)

Treatment	N	Difference (%U-%S)	z-stat (Diff-in-Diff, Treatment-Control)
Control	26	-26.92%	
0 U Messages	4	-50.00%	-0.75
1 U Messages	12	-41.67%	-0.81
2 U Messages	11	-45.45%	-0.98
3 U Messages	11	-27.27%	-0.02
4 U Messages	25	12.00%	2.85***
5 U Messages	16	18.75%	2.35**
0-1-2 U Messages	27	-44.44%	-1.22
3-4-5 U Messages	52	5.77%	2.49**

Table 6
Results of Calcutta Lab Experiment

Treatment (Reported Propensity Untruthful Senders)	Number of Subjects	Percent Truthful	z-statistic (Control - Treatment)	Percent Predicting Receiver Acceptance
Control	54	44.4%		75.9%
Y=15%	39	69.2%	-2.475**	76.9%
Y=85%	38	28.9%	1.550	73.7%
Overall	131	47.3%		73.6%

** denotes significant at 5% level (two-sided).

Table 7
Calcutta Experiment: California Design

Treatment (Number of Untruthful Draws)

	Control (No draws)	Truthful (0-1-2 U draws out of 5)	Untruthful (3-4-5 U draws out of 5)
<u>A) Deception</u>			
Observations	54	46	37
Percent Truthful	44.44%	63.04%	18.92%
z Stat 1 (Cont.-Treat.)	-	-1.940*	2.686***
z Stat 2 (012-345)	-	-	4.597***
Average Sender Belief (SD) ⁺	34.90% (19.05)	66.52% (22.60)	35.61% (20.79)
z Stat for Belief (012-Ct/345) ⁺	6.248***	-	6.476***
<u>B) Dictator</u>			
Observations	25	46	37
Percent "Selfish"	72.00%	69.56%	72.97%
z Stat 1 (Cont.-Treat.)	-	0.217	-0.084
z Stat 2 (012-345)	-	-	-0.342

⁺ Based on 25 Control observations

* The authors thank the Editor and two anonymous reviewers for prescient comments on a prior version of this paper. We are indebted to Martin Dufwenberg and John List for their generous advice on this research; to Bruce Beattie, Dennis Cory, George Frisvold, Paul Wilson, Alex Whalley, Todd Neumann, Trevor Kollmann, Jennifer Pullen, and Jon Carlson for their generosity with class time; and to Gautam Gupta, Sanmitra Ghosh, Abhishek Das, Nilay Tikadar, Bijoy Sarkar, Daisy Paniagua, and Shannon Iraniha for their help in running experiments. We are also grateful to seminar participants at the University of Arizona, UC Davis, UC Berkeley, UC Merced, NC State, Purdue, University of Nebraska, Cal Poly, and the University of Washington for a variety of helpful suggestions, particularly Subhasish Dugar and Chuck Knoeber. The usual disclaimer applies.

¹ We stress the illustrative character of Figure 1. While the CPI is widely considered a good measure of corruption (see, for example, Treisman, 2000; Mo, 2001), the twin phenomena of dishonesty (our focus) and corruption (Figure 1) are, while clearly related, not isomorphic. The two words are often used interchangeably, with bribe-taking officials alternately called “corrupt” and “dishonest” (e.g., Bardhan, 1997, p. 1332).

However, we do not study links between these two phenomena in this paper.

² Andvig and Moene (1990), among others, develop models of corruption that give rise to a “Shelling diagram” of behavior in which there are two stable equilibria: all corrupt and none corrupt (see Bardhan, 1997). Drawing on a variant of Frank’s (1987) model, Innes (2009) produces a similar diagram of honest vs. dishonest behavior. Another possible explanation for bimodal corruption outcomes is a vicious cycle in which low incomes

promote corruption which, in turn, deters growth and so on. There is a vast literature on the evolution of institutions and their relationship to corruption and growth (see, for example, Acemoglu, et al, 2001). One interpretation of our paper, in the context of this literature, is that there may be self-reinforcing dynamics to the evolution of bad and good economic institutions.

³ Other recent experimental work on deception games include Sanchez-Pages and Vorsatz (2007), who study links between a subject's willingness to punish lies of others and their aversion to lying; Ederer and Fehr (2007), who study impacts of deception and aversion to lying in a principal agent game; Sutter (2009), who identifies sophisticated deception; Rode (2010), who studies effects of competitive and cooperative priming on subjects' honesty; and Charness and Dufwenberg (2005) who provide an alternative (guilt aversion) interpretation of Gneezy's (2005) findings.

⁴ An interesting paper by Keizer, et al. (2008) finds that, in public field experiments, subjects tend to violate one norm (the target) more frequently when they see that another norm (the contextual norm) is violated. For example, if graffiti is present (despite a "no graffiti" sign), subjects are more likely to litter. There are three key differences with the present study: (1) Keizer, et al. (2008) are concerned with cross-norm contagion whereas we focus on within-norm contagion; (2) we focus on deception, rather than norm-violating public actions (littering or swiping cash); and (3) subject decisions are public, whereas they are private in our study. The last difference implies that, in Keizer, et al. (2008), subject behavior may potentially be due to effects of the contextual norm treatments on the perceived risk of either law enforcement or social sanctions against violation of the target norm. Social sanctions are not possible in our experiments.

⁵ In a dictator game, a “dictator” allocates a given pot of money (e.g., \$10) between himself and one other player (his Receiver). “Fair” allocations split the money more evenly (50-50 or 60-40), while “selfish” allocations split the money less evenly (80-20, for example).

⁶ The possible exception is Duffy and Kornienko (2009), where winning a tournament, even if the winner is only identified anonymously (by ID number), may provide some intrinsic reward.

⁷ A higher perceived propensity for honesty among an experiment’s population of Senders increases the probability that Receivers obtain their high payoff, which raises the net benefit to dishonesty by lowering the implied cost of inequity. For our U.S. experiments, this effect tilts incentives toward dishonesty, contrary to the contagion hypothesis. Similar logic (for both dictator and deception games) applies to spiteful (Fehr, et al., 2008) and ERC (Bolton and Ockenfels, 2000) preferences. In our expanded paper, we explore in detail these and other possible mechanisms underpinning the contagion in honesty that we observe. We argue that other-regarding preferences, such as Fehr-Schmidt (1999) inequity aversion or Dufwenberg and Gneezy (2000) guilt aversion (see also Charness and Dufwenberg (2006) and Battigalli and Dufwenberg (2007)), cannot explain our findings. As a result, we conjecture that contagion is the outcome of “hard-wired” preferences, which motivates an examination of evolutionary advantages of this trait (Innes, 2009).

⁸ Effects of social information have been studied in a number of other contexts. In ultimatum games, Knez and Camerer (1995) and Bohnet and Zeckhauser (2004) examine effects of information about other proposer offers on proposer and responder behavior,

finding evidence of a pre-existing norm of equity. Several authors study the role of social information in achieving social learning and conditional cooperation in coordination games (Berg, Dickhaut and McCabe, 1995; Fischbacher, Gächter and Fehr, 2001; Schotter and Sopher, 2003; Chaudhuri, Graziano, and Maitra, 2006; Eckel and Wilson, 2007); in this work, unlike ours and like studies by Chen, et. al (2009) on on-line participation in MovieLens and Duffy and Feltovich (1999) on learning in ultimatum and best shot games, the social information is potentially payoff relevant. Recent field experiments on charitable contributions document that subjects contribute more often when they believe a higher fraction of their peers contribute (Frey and Meier, 2004) and contribute more when told that a prior contributor contributed more (Shang and Croson, 2009); in this work, the social information can also be payoff relevant by signaling the virtue of the charity. Overall, this work provides evidence of contagion in different contexts, but with potential channels of effect that are not at play in our experiments.

⁹ There is no obvious choice of payment options. We conducted preliminary surveys on alternative options that varied (1) the gain to the Sender from lying G_S (assuming Receiver acceptance of recommendations), and (2) the corresponding Receiver loss L_R . Consistent with expectations, incentives to lie rise with G_S and fall with L_R . Our survey evidence implied an approximate Sender propensity for truthfulness equal to 58 percent for $G_S=2$ and $L_R=3$ (our chosen options). Armed with this evidence – and the conjecture (wrong as it turned out) that actual dollar stakes would raise incentives to lie – we settled on the indicated options.

¹⁰ A norm in experimental economics is that the experimenter be honest with his/her subjects. We obey this norm with our approach. However, our treatments are intended to

influence perceptions. We note that experimental designs with such objectives are common in the experimental economics literature. Prominent examples are influential papers that report a subject's "awarded" status "to suggest to the (other) subjects that the high status was deserved" when in fact it was randomly assigned (Ball, et al., 2001), that expose subjects to resume's with fictitious racial profiles (Bertrand and Mullainathan, 2004), that elicit contributions for a public project given fictitious variation in seed money (List and Reiley, 2002), and that use a standard experimental protocol to not inform subjects that they will be playing in subsequent rounds or roles (e.g., Binmore, et al, 1985; Duffy and Kornienko, 2009). See Bonetti (1998) for a lucid discussion of this topic.

¹¹ In principle, risk aversion could motivate an "accept" prediction by truthful Senders and a "reject" prediction by untruthful Senders, with the prediction serving as insurance. (If the Sender obtains the low payoff for the game, the loss is offset by payment for the prediction.) However, the proportion of truthful Senders predicting Receiver accept decisions (74.8%) is essentially identical to the proportion of untruthful Senders predicting accept decisions (72.0%) in our experiment. Hence, risk aversion does not appear to motivate Sender predictions in our experiment.

¹² There was no overlap between the Receiver class and any of the Sender classes. Two students who were enrolled in two of the Sender classes were not present when the second class experiment was performed.

¹³ The proportions of Sutter-truth-tellers are 56%, 57.7% and 60.6% under the other (Y=15%, 40% and 60%) treatments. Note that Sutter (2009) finds almost no "benevolent liars" in his experiment. In contrast, the proportion of "benevolent liars" in our subject

pool (14.2%) is roughly the same as the proportion of “sophisticated liars” (12.4%); we also find no clear pattern in this behavior across treatments (for example, in the proportion of liars who are benevolent or the proportion of truth-tellers who are sophisticated liars). These observations loosely suggest that the predictions of our Reject-predicting subjects may be random, reflecting an anticipation that the Receiver essentially flips a coin when making his choice.

¹⁴ We are indebted to a prior reader for highlighting these issues.

¹⁵ Each Sender drew from one of two boxes, one containing messages from the Arizona Control treatment experiment and one containing messages from the heavily untruthful treatment (Y=85%).

¹⁶ The Deception game was denoted by *K* and the dictator game by *L*. Senders were given the following instructions: “You and your Receiver will participate in two different decision-making situations, which we identify by *K* and *L* below. Both of you will be paid for ONE of the two situations. The situation for which you will be paid will be determined by a flip of a coin after all decisions have been made by all participants. You should therefore make your decision in each situation as if it is the one for which you will be paid.

- You and your Receiver will be paid for situation *K* if the coin toss comes up **Heads**.
- You and your Receiver will be paid for situation *L* if the coin toss comes up **Tails**.”

Receivers were given parallel instructions. The “Heads” vs. “Tails” determinants of outcomes was varied randomly between Senders.

¹⁷ See “Household Consumption Expenditure in India (January-June 2004),” NSSO, Government of India, 23 November, 2005.

¹⁸ As in Arizona, all Senders were told that roughly 80% of Receivers accepted their Sender recommendations in a similar prior experiment; none of this information was provided to Receivers and Senders were so informed.

¹⁹ In the India experiment, a slightly higher fraction of truthful Senders predict Receiver acceptance than do untruthful Senders (80.6% vs. 71.0%), but the difference is not statistically significant ($z=1.298$).