Outsourcing Morality by Lying Through Others: Does Delegation Promote Deception?

Glynis Gawn Robert Innes *

February 2016

Abstract

How do agency relationships affect an individual's willingness to lie for monetary advantage? Does lying aversion decline if a lie (or truth) is sent through an agent, rather than sent directly by the individual? In experiments that control for effects of delegation on preferences over payoffs and probabilities of actions, we find that delegation reduces – but does not eliminate – lying aversion. We conclude that delegation reduces the extent to which individuals hold themselves responsible for a lie, even when subjects retain control over the decision.

Keywords: Delegation, Deception, Lying Aversion, Responsibility, Control

JEL Classifications: D03

*Economics, University of California, Merced. School of Social Sciences, Humanities & Arts, U.C. Merced, 5200 N. Lake Rd., Merced, CA 95343. Emails: ggawn@ucmerced.edu, rinnes@ucmerced.edu. Phone: (209) 228-4872.

1 Introduction

Agency relationships are ubiquitous in economic interactions. Outsourced suppliers and subcontractors produce for contracting companies. Employees act for firms. Intermediaries act on behalf of clients in employment and professional services. In many of these settings, agents make decisions with ethical overtones that affect economic costs and benefits to their principals. Outsourced suppliers may misrepresent questionable labor practices (e.g., Nike in the 1990's) or noxious product contents (e.g., Lumber Liquidators in 2015). External auditors may misrepresent financial conditions (e.g., Enron, Worldcom). Employment agencies may hire illegal immigrants as house help for clients (e.g., Mitt Romney, as discussed in Erat, 2013). The complicity of principals in these decisions is often presumed.[1]

One type of moral decision – relevant to many of these examples – involves deception or lying for monetary advantage. The phenomenon of lying is widespread in practice. For example, DePaulo et al. (1996) find that college students and members of a U.S. community lied in 20 to 31 percent of social interactions recorded in daily diaries. In a recent survey, 76 percent of U.S. high school students self-report lying about something significant in the past year, while 38 percent indicate that they sometimes lie to save money (Josephson Institute of Ethics, 2012). More surprising for the economist, individual aversion to lies is pervasive. Contrary to the standard economic model of payoff maximizing behavior, individuals often tell the truth despite monetary costs of doing so (e.g. Gneezy, 2005; Sánchez-Pagés and Vorsatz, 2007, 2009; Fischbacher and Heusi, 2013; Gibson, Tanner and Wagner, 2013; Erat and Gneezy, 2012). Indeed, Abeler, Becker and Falk (2014) find a particularly strong aversion to lies in the field vs. the lab.

Our purpose in this paper is to study how principal-agent relationships can affect lying aversion. If you are choosing to lie or not via an agent, such as an outsourced supplier, is your aversion to the lie less than if you were directly deciding whether or not to lie? In other words, who do you hire as your agent, an agent that is more likely to lie or more likely to tell the truth? In three classroom experiments, we find that delegation increases the propensity

[1] See, for example, O'Rourke (1997) on Nike, McLean and Elkind (2003) on Enron, 60 Minutes (2015) on Lumber Liquidators, and Friedman (2011) on Romney.

for deception.

Answering these questions can reveal not only an economic motive for delegation – suppressing honesty that is costly to a principal – but also how delegation affects moral behavior more broadly, including trust and corruption in economic interchange. For example, Cullen (2008) describes the case of the Canadian engineering group Acres International accused of bribing the chief executive of a Lesotho water project. Charged with violating Lesotho's anti-bribery statutes, "Acres' defence was that they were not responsible for the (bribe) payments... as these were made via an intermediary through a 'representation agreement'." Cullen (2008) writes that "there is an unwritten expectation that the 'facilitator' will do whatever is required, short of overtly criminal acts, to secure the contract or project. For a company like Acres, it is a way of distancing their bid or proposal from the unsavoury aspects of doing business in corrupt countries." Our experiments relate precisely to a situation like this in which the principal can choose an agent knowing the choices that will be made on his/her behalf.

A prominent narrative for corrupt transactions such as in the Acres' case pits reluctant transnational corporations (TNCs) against corrupt local bureaucrats, with each company facing a "prisoner's dilemma" with competing TNCs in which the dominant strategy is to acquiesce in a corrupt exchange (see Collier, 2007, p.137, for example).[2] However, if delegation to an intermediary erodes the moral preferences of a principal, as we find in our experiments, then causation also runs in the opposite direction: the process of intermediation promotes dishonesty and related corruption by the TNC, via an agent who is chosen precisely to be deceitful.

Our experiments build on prior literature studying the economic effects of delegation in dictator games, where a principal/dictator allocates money between him- or herself and another player. Hamman, Lowenstein and Weber (HLW, 2010) show that delegation leads

[2] To some extent, this perspective plays out in the recent debate over Basu's (2011) proposal to exempt bribers from any sanction in corrupt transactions (see, e.g. Abbink et al., 2014; Dreze, 2011; Dufwenberg and Spagnola, 2015). Although this literature concerns "harassment bribes" for services to which a briber is entitled, it stresses the central role of public officials, vs. private parties benefiting from public services or contracts, as the driving force in corrupt transactions.

to allocations that are more favorable to the principal and less favorable to the matched player; that is, delegation makes principals more selfish. A similar result is obtained by Drugov, Hamman and Serra (2014) in the context of a corruption game where a "briber" and "bribee" can increase their own payoffs at the expense of a third party. Bartling and Fischbacher (2012), Oexl and Grosmman (2013) and Coffman (2011) show that delegation reduces the extent to which dictators are punished for selfish allocations.

Our study is distinguished from these key results by its focus on lie averse preferences versus preferences for fairness. There are differences between the moral and economic content of the two phenomena that motivate a distinct and growing economics literature on deception (Rosenbaum, Billinger and Stieglitz, 2014). For example, the "selfishness" exhibited in dictator games is sometimes heralded (by economists in particular) for promoting effort and innovation that are central to successful market economies. In contrast, dishonesty and corruption are consistently condemned by churches, community leaders, philosophers and even economists for eroding civility, trust and the moral fabric of society (e.g. Bok, 1978; Sachs, 2011). As a result, lie averse preferences may be more rigid than a preference for "fairness." Rode's (2010) results are illustrative; he finds that subjects' honesty is insensitive to competitive versus cooperative priming; in contrast, a number of studies find that markets promote more "selfish" preferences (see, for example, Heinrich et al., 2010). Certainly, results from dictator games cannot simply be assumed to apply to the phenomenon of deception.[3]

To isolate effects of delegation on lying aversion – our objective – requires attention to two correlated phenomena. The first concerns preferences over allocations, as highlighted in Hamman, Lowenstein and Weber (HLW, 2010). If a lie benefits the liar at the expense of the recipient – a "black lie" as defined in Gneezy (2005) and Erat and Gneezy (2012) – then lying can be motivated by a preference for the self-benefiting payoffs that a lie produces (Hurkens and Kartik, 2009). Effects of delegation on lying behavior might therefore be attributable to the HLW effect, namely, a stronger preference for the selfish (vs. generous) payoff when obtained via an agent. Using parallel dictator games, our experiments control

[3]List (2007) and Bardsley (2008) show that even small framing differences in dictator games can have significant effects on behavior, let alone more profound changes to a game (from dictator to deception).

for this correlated effect and reveal that delegation increases the propensity to lie by more than it increases the preference for the self-benefiting (lying) allocation.

The second correlated effect concerns the ceding of control. At the heart of delegation are two inter-related changes to the decision environment: delegation can reduce (1) control over a decision and outcomes and (2) attribution of responsibility for a decision. Often the two are considered one and the same. Philosophers argue that individuals are responsible for outcomes only if they can control them (Nelkin, 2004; Gurdal, Miller and Rustichini, 2013). Ceding of control is a central feature of delegation in the recent experimental literature studying its effects.[4] Indeed, Bartling and Fischbacher (2012) explain their results using a new measure of responsibility driven by the extent to which a principal is perceived to affect the outcome.[5]

The ceding of control may be a feature of some delegation situations, but is also a feature of many other situations that produce "moral wiggle room" (Dana, Weber and Kuang, 2007). When negative outcomes for a matched player can be due to either nature or a dictator's decision – with an attendant reduction in the probability that the dictator's decision is implemented – then dictators tend to "hide behind nature" and act more selfishly (Dana et al., 2007; Andreoni and Bernheim, 2009).[6] Decisions in groups, or based on team incentives, can make subjects more selfish (Charness and Sutter, 2012) and more willing to lie (Conrads et al., 2013; Cadsby et al., 2010; Muehlheusser, Roider and Wallmeier, 2015). Studies on responsibility alleviation (Charness, 2000) and hidden costs of control (Falk and Kosfeld,

[4]An exception is Drugov et al. (2014), who find that the presence of a powerless intermediary increases the fraction of subjects who make a bribe that benefits the briber and bribee at the expense of a third player. One interpretation of this result, particularly relevant to our study, is that the HLW effect is present even when control is not ceded.

[5]See also Eisenkopf and Fischbacher (2015) who find that Returners in a trust game differentially reward the person who makes the trust decision, whether a delegate or a Sender.

[6]See also Freshtman and Gneezy (2001) and Bartling and Fischbacher (2012), where punishment is possible. An interesting paper by Haisley and Weber (2010) shows that subjects may "hide behind ambiguity"; adding ambiguity to probability distributions of outcomes – even when ultimate probabilities are the same – leads to more selfish decisions (Haisley and Weber, 2010). Ambiguity does not reduce control objectively, but may be perceived as doing so, as indicated by the authors. Another recent literature shows that, if subjects can be "willfully ignorant" – by opting not to know about the adverse consequences of increasing their own payoff on the payoff of someone else – then they tend to be more selfish (Dana et al., 2007) and less subject to punishment (Bartling, Engl and Weber, 2014). See also Grossman (2015), who finds that few subjects will deliberately choose to be ignorant. While willful ignorance does not cede control per se, it does cede knowledge of consequences.

2006) show how reducing an agent's control can elevate moral hazard (Charness et al., 2012).

While a reduction in control surely implies a reduction in responsibility, delegation might affect responsibility attribution even when it has no effect on control. In many of the examples cited at the outset, delegation does not cede control per se – agents are selected who make decisions that the principals would themselves make if they could immunize themselves from responsibility. Oexl and Grossman (2013) find that delegation reduces the extent to which dictators are blamed and punished for selfish allocations, even when no control is relinquished by the delegation. Delegation might also affect *internal* responsibility attribution – assignment of responsibility to oneself for purposes of evaluating moral trade-offs. This is our central focus in this paper: How does delegation affect lying aversion when there is no loss of control over decisions and no scope for external responsibility attribution or punishment?

In addressing this question, we build upon a key finding of Erat (2013). To our knowledge, Erat (2013) is the first and only prior paper studying effects of delegation opportunities on deception. In his experiments, subjects play a Sender-Receiver game in which each Sender either lies, tells the truth, or delegates; with delegation, the (unknown) decision of a randomly selected player is implemented. Erat (2013) finds that a significant proportion of subjects opt to delegate their decision on whether or not to lie, rather than make the decision themselves. This benchmark result is potentially attributable to a reduction in lying aversion under delegation or, alternately, to the correlated effects of delegation on preferences over payoffs (the HLW effect) and/or on control over outcomes. Our objective is to pinpoint the first explanation – that delegation reduces lying aversion. [7]

After a brief overview of our approach and hypotheses in Section 2, the paper is organized by our three experiments, presented in Sections 3, 4 and 5 below. Section 6 concludes.

[7] Another potential motive for delegation in Erat's (2013) experiments may be the desire to avoid cognitive costs of making a decision (akin to the exit choice in Dana, Cain and Dawes, 2006). Our experiments void such cognitive opt-outs by forcing subjects to make equivalent decisions with and without delegation.

2 Overview of Experiments and Hypotheses

In this paper, we seek to determine whether and how the process of delegation affects an individual's aversion to lies. Are individuals more likely to delegate to liars than they are to lie themselves? In order to give saliency to the delegation decision, we base our experiments on deception interactions in which one person (the Sender) is lying or telling the truth to another person (the Receiver), as in Gneezy's (2005) initial design, rather than recent variations in which there is no one at the receiving end of the lie/truth (e.g., Gibson, et al., 2013; Fischbacher and Heusi, 2013; Abeler, et al., 2014). While this design confounds preferences over allocations with lying aversion, as stressed in the recent literature (Hurkens and Kartik, 2009; Gibson et al., 2013), delegation is arguably trivialized if there is no interpersonal content to the interaction. Our analysis is therefore dedicated, in large measure, to addressing and decomposing joint impacts of delegation on lying aversion and preferences over payoffs.

In each of our experiments, there are two possible payoff allocations, one that is better for the Sender (call it Option 1) and the other that is worse for the Sender (call it Option 2). If a lie is sent on behalf of the Sender, Option 1 is more likely to be implemented; conversely, if a truthful message is sent, Option 2 is more likely. In our Experiments 1 and 2, the first (Option 1) allocation is worse for the Receiver, so that a lie benefits the Sender at the expense of the Receiver. In Experiment 3, Option 1 is better for both Sender and Receiver. Gneezy (2005) and Erat and Gneezy (2012) describe the former situation as a "black lie" and the latter as a "white lie", nomenclature that we will borrow.

In the black lie experiments (1 and 2), subjects participate in two games (deception and dictator) and two treatments (direct choice and delegated choice). The payoff options are identical across games and treatments. In the direct treatments, Senders choose directly whether to lie or not (deception) and which payoff option to select (dictator). In the delegation treatments, Senders each choose one of two agents (Agent 1 vs. Agent 2) and the decision of that agent is implemented. In the deception game, Agent 1 has lied and is more likely to produce a lie on behalf of the Sender; Agent 2 has told the truth and is more likely to produce a truth on behalf of the Sender. Similarly, in the dictator game, Agent 1 is more likely to produce the Sender-benefiting payoff option and Agent 2 is more likely to produce the generous (Receiver-benefiting) payoff option.

In both treatments, Senders express a preference for a lie vs. a truth (in the deception game) or the selfish vs. generous option (in the dictator game), either by their direct choice (in the direct treatment) or by their choice of agent (in the delegation treatment). In this sense, delegation does not cede control over decisions. However, in our benchmark Experiment 1, delegation involves more uncertainty about the decision than does the direct choice. For example, delegating to a liar produces a lower probability that a lie is delivered than does the direct choice of a lie. Although this uncertainty is likely to characterize delegation situations in practice, it sacrifices a degree of control. In Experiments 2 and 3 – which we consider our main designs – we match probabilities across treatments in order to avoid the differential uncertainty. In these experiments, the actual decisions of Agents 1 and 2 are known to the Sender; for example, if a Sender chooses the lying Agent 1 in the delegation treatment, the lie is sent with exactly the same probability as when a lie is chosen in the direct treatment.

Our over-arching prediction is that delegation – implementing a lie through another person rather than directly – reduces the extent to which individuals hold themselves morally responsible for the lie. That is, delegation reduces lying aversion. This prediction has three symptoms for which we test in our experiments:

> *Hypothesis 1 (the difference test).* A larger fraction of subjects will lie when making a delegated decision (through an agent) than when making the decision directly.

In principle, evidence for Hypothesis 1 from our black lie experiments (1 and 2) could be attributable to impacts of delegation on preferences over allocations (the HLW effect). More subjects may lie under delegation because more subjects prefer the "selfish allocation". To net out the latter effects and thereby enable inferences about lying aversion, we have:

> *Hypothesis 2 (the difference-in-difference test).* Delegation will increase the frac-

tion of subjects who lie (by selecting the lying agent) by more than it increases the fraction of subjects who choose the selfish payoff allocation (by selecting the "selfish" agent).

The on-line Appendix presents a simple conceptual model of responsibility that frames our experiments. Allowing for a general distribution of preferences over payoffs (to Sender and Receiver) and lies (vs. truths), the model produces Hypothesis 2 as a test of treatment effects on lying aversion when the HLW effect is present.

Hurkens and Kartik (2009) propose a different approach to disentangling treatment effects on payoff preferences. Because only those who prefer the "selfish" payoff would lie – and these subjects would only tell the truth due to an aversion to lying – Hurkens and Kartik (2009) argue that lying aversion can be measured by conditioning on a selfish choice in the dictator game. Conditional on preferring the self-benefiting payoff allocation, to what extent do subjects lie versus tell the truth? If the conditional proportion of lies falls in one setting versus another, then one can conclude that lying aversion has risen on average.

Hypothesis 3 (the Hurkens and Kartik test). Delegation increases the fraction of "selfish" subjects who lie.

3 Experiment 1: The Dot Delegate Experiment

The Deception Experiment. Our baseline experiment follows Gneezy (2005): Players are paired across sessions, with each Sender matched with a corresponding Receiver. The paired players obtain one of two possible payoff allocations, Option A or Option B. Only Senders are 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. One of two messages is sent to the Receiver on behalf of the Sender:

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. Based only on the message, the Receiver chooses an option (A or B) and payments are then made according to that option. 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. Under one payoff option, the Sender receives $7 and the Receiver obtains $3; under the other, the Sender receives $5 and the Receiver obtains $6.[8]

In all cases, we seek to avoid strategic motives by giving Senders general information on the propensity of Receivers to accept their recommendations. Based on results from Gneezy (2005) (where 78 percent of Receivers followed the Sender recommendations), all Senders are told:

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

Receivers are not given this information, and Senders are so informed. To verify that Senders generally believe that Receivers will accept their recommendations, we ask them to predict their Receiver's choice and pay them $1 for a correct prediction. 72.7 percent of Senders predict that their Receiver will choose the recommended option. As it turns out, 72.9 percent of our Receivers follow their Sender recommendations.

In order to control for preferences over allocations, subjects play both a Gneezy deception game and a parallel dictator game, with one of the games randomly selected for payment (each with 50 percent probability). This approach follows Hurkens and Kartik (2009) with a twist: When the dictator game is selected for payment, the chosen dictator allocation is implemented with 80 percent probability in order to mimic deception game payoffs (as in Gneezy, 2005). Senders/dictators are informed of this procedure.

Delegation Using the Dot Experiment. We distinguish between deception decisions made directly by a subject (our direct treatments) vs. indirectly by choice of an agent (our delegation treatments). To implement delegation, we give subjects a choice between agents who are likely to be more truthful vs. less truthful and more generous vs. less generous.

[8]There is no obvious choice of payment options. Our payments correspond roughly with those in Innes and Mitra (2013), with a gain to Sender dishonesty of $2 and a corresponding Receiver loss of $3.

To define potential agents/delegates, we exploit a Lopez-Perez and Spiegelman (2012) "dot experiment" conducted at an earlier time with a different set of subjects. In the dot experiment, each subject observes the color of a dot, either blue or green. For the sake of exposition, let us say that the true color is blue. The subject reports on the experimental survey that the dot is blue (a truth) or that the dot is green (a lie). By reporting the true color (blue), a participant earns \$1; by instead reporting the false color (green), he/she earns \$2.

In the prior experiment, participants completed both the dot experiment and the Gneezy (2005) deception experiment with parallel dictator game, as described above. The prior participants are called either "\$1 Senders" or "\$2 Senders", with the former truthfully reporting the color of the dot and the latter reporting the incorrect color in return for a \$2 payment.

In the delegation treatment, student participants also play a Gneezy deception game – the same one played by the prior participants – with the only difference being that they do not themselves decide on messages and allocations; instead they each choose between a randomly selected "\$1 Sender" and a randomly selected "\$2 Sender". For example, if the student delegates to a \$1 Sender, then one of the \$1 Senders is randomly selected from the prior experiment and the decision of THAT SENDER, in the Gneezy deception experiment, is implemented for the delegating student and his/her matched Receiver. The student separately chooses a delegate (\$1 or \$2 Sender) in the deception experiment and the parallel dictator game.

In order to keep the choice setting as natural as possible, we do not tell students the rates at which \$1 Senders and \$2 Senders made truthful or generous choices. However, they are asked to verify their understanding of the situation by indicating which of the two Sender types earned more money by reporting the incorrect color of the dot. 83.3 percent of respondents provide the correct answer to this question. We also elicit the students' beliefs about probabilities of the different Sender types choosing the self-interested Message and dictator allocation, respectively. Each correct prediction – within the correct ten percentage point band – is rewarded with a \$1 payment. Responses indicate broad but not unanimous understanding of differential behavior by $1 and $2 Senders. Thirty percent of delegating subjects predict a lower probability of a truthful choice by a $1 versus $2 Sender, contrary to expectations. For the dictator game, twenty-five percent predict a lower probability of a generous choice by a $1 versus $2 Sender. These percentages are significantly lower than 50 percent, indicating broad success in conveying the nature of the delegation choice. As a robustness check on our results, we consider a sample of delegators that excludes the anomalous predictors whose delegation decisions potentially do not reflect the intended correspondence between more and less truthful agents.[9]

In the prior session, $1 and $2 Senders behaved as expected. 62.2 percent of $1 Senders were truthful in the Gneezy game, compared with 35.9 percent of $2 Senders; similarly, 55.6 percent of $1 Senders were generous in the dictator game, compared with 27.8 percent of $2 Senders.

Logistics. We conduct the experiment in three upper-division undergraduate economics classes at U.C. Merced. In total, there are 142 Sender/Receiver pairs, with 72 Senders in the direct treatment and 70 Senders in the delegated treatment. Receivers are in different classes than paired Senders. To ensure anonymity, each subject is identified by an identification number that is attached to the experimental questionnaire. Participation is purely voluntary and class rosters are used to avoid a student participating more than once. Subjects are instructed to communicate only with the experimenters and are carefully monitored to this end. Direct and delegated Sender treatments are implemented in equal proportions in each of three sessions, with questionnaires randomly distributed to students.[10] Dot colors and option labels are randomly varied.

[9]We do not want to over-emphasize responses to the belief questions. Prior work in psychology and economics documents that subjects often make decisions that reflect a subconscious understanding of the situation even when they cannot explain this understanding. See, for example, the celebrated "red card/blue card" paper by Bechara et al. (1997) and a recent paper by Friedman et al. (2015) indicating that large numbers of repetitions can produce cooperation even though participant answers reveal that they do not understand the economic environment underpinning the interactions.

[10]Our experiments are conducted in classrooms with limited time and subject anonymity. As a result, we have only one observable variable to judge cross-treatment balance, namely, gender (Male). The proportions of male subjects in the direct and delegated samples (full and restricted) are, respectively, 40.0% (direct), 36.23% (full delegated) and 35.42% (restricted delegated); differences are not significant.

Results of Experiment 1. Figure 1 and Table 1 present results of Experiment 1. The Table gives coarse summary statistics, as well as relevant difference and difference-in-difference tests. Our baseline (direct treatment) results are broadly consistent with prior literature on dichotomous choice experiments, with 51.4 percent of subjects choosing the "generous" option in the dictator game and 58.3 percent opting for truth in the deception game.[11]

For the delegation treatment, we measure "truth" by delegation to a $1 Sender (vs. $2 Sender); similarly, we measure "generosity" by delegation to a $1 Sender in the parallel dictator game. Using these measures, we find that delegating subjects are significantly less truthful than are subjects in the direct treatment, consistent with Hypothesis 1. This difference is evident using both the full delegation sample and the restricted sample that excludes subjects who indicate a belief that $1 Senders tend to be less truthful. Delegation reduces the proportion of truthful subjects by 22.6 percentage points in the full sample and 25.7 percentage points in the restricted sample; both differences are statistically significant at 1 percent.

To compare the impact of delegation in the deception game to its impact in the parallel dictator game, we use two approaches. First is the difference-in-difference test that examines the difference between the delegation effect on truthfulness and the corresponding effect on generosity. As indicated in the third row of Table 1, difference-in-difference statistics are statistically significant and consistent with Hypothesis 2; delegation reduces the propensity for truthfulness by an estimated 22.8 to 25.4 percentage points over and beyond its impact on the propensity for generosity. Second is the Hurkens and Kartik (2009) test. Among Senders indicating a preference for the self-benefiting ($7-$3) allocation in the dictator game,

[11] While an exact parallel to our payoffs is hard to find in the literature, our percentages are roughly in line with other work dating back to the original dichotomous choice experiments of Kahneman, Knetsch and Thaler (1986). For example, with somewhat starker choices between "fair" ($8/$8) and "selfish" ($11/$2) allocations, Fershtman, Gneezy and List (2012) find that 72.5% of U.C. San Diego subjects chose the "fair" payoff in their baseline dictator games. Indeed, in the Wisconsin experiments of Andreoni and Miller (2002), the estimated percentage of equality-preferring (Leontieff) and social welfare maximizing subjects – those who we would expect to be generous in our dictator game – was 52.8%, almost exactly equal to our observed percentage of generous players. In his original deception game, Gneezy (2005) finds that 64% of his Israeli subjects were truthful with weaker incentives to lie than in our experiments and 48% were truthful with stronger incentives to lie. With similar payoffs to ours, Innes and Mitra (2013) obtain almost identical proportions of truthful subjects in their Arizona (58.8%) and California (57.7%) control experiments as we do in our baseline Gneezy game (58.3%).

what fraction make a deceptive choice in the Gneezy game? Specifically, do more of the "selfish" subjects prefer deception in the delegated treatment than in the direct/control treatment? The Hurkens-Kartik statistics in Table 1 indicate that the answer is "yes", consistent with Hypothesis 3. Delegation reduces the proportion of the selfish who are truthful by 22 percentage points (using the full delegation sample) and by 30.6 percentage points (using the restricted delegation sample). The differences are statistically significant at 10 percent and 1 percent (two-tail) in the respective cases.

Panel B of Table 1 presents regressions that control for course and gender effects.[12] Coefficients for the Delegate dummy indicate the effect of delegation in the deception game, first on the overall propensity for truthfulness (Model 1), second on the preference for truthfulness among the selfish (the Hurkens-Kartik Model 2), and third on the differential preference for truthfulness over generosity (the difference-in-difference of Models 3 and 4). In all cases, the results of Panel A are confirmed, with strikingly consistent parameter estimates. Overall, delegation is estimated to reduce truthfulness by roughly 22 to 25 percentage points.

Experiment 1 produces two tentative conclusions. First, *delegation reduces lying aversion for a significant fraction of subjects.* Second, however, *delegation does not eliminate lying aversion.* The second conclusion is indicated by the Hurkens-Kartik statistics in Table 1. Among the delegating selfish, a positive proportion are truthful and this proportion is significantly different from zero in both delegation samples; corresponding z-statistics (p-values) are 3.76 (p=0.0006, full sample) and 2.51 (p=0.0186, restricted sample).

Criticisms of Experiment 1. An advantage of Experiment 1 is a natural design that controls little about Sender perceptions of the potential $1 and $2 Sender delegates, other than their behavior in the dot experiment. However, the absence of control also raises some potential issues that motivate further investigation.

First, in the direct treatment, decisions made by subjects are carried out with 100 percent probability. If Message A is chosen, it is sent. In the delegation treatment, choice of a $2

[12] Coefficients are from OLS regressions with robust standard errors. Similar results are obtained from qualitative dependent variable models such as Probit.

versus a \$1 Sender leads to a different probability that the deceptive message is sent. For example, suppose that a subject believes there is a 70 percent likelihood that \$2 Senders are deceptive in the Gneezy game and a 45 percent likelihood that \$1 Senders are deceptive (the average beliefs in our sample). Then, by choosing a \$2 (vs. \$1) Sender, the subject increases the probability of deception by 25 percent. In this sense, the decision under delegation involves less control over the outcome than under the direct treatment. Although these probabilistic differences are also symptoms of actual delegation situations, they may cloud the identification of pure delegation effects on lying aversion.

Second, there are two possible criticisms of the dictator controls. Although a \$1 vs. \$2 choice in the dot experiment provides a clear signal for truthful vs. deceptive behavior, it is perhaps a less clear indicator of behavior in the dictator game.[13] In addition, our dictator controls are within-subject, which could produce "experimenter demand" or portfolio effects that promote different behavior across the two games.[14]

Our main Experiments 2 and 3 address these issues with (1) between-subject designs that (2) explicitly control for probabilities that choices are implemented and (3) a framework for delegation in the dictator game that is explicitly tied to agent behavior in that game.

4 Experiment 2: The Black Lie Experiment

The Black Lie Experiment. In Experiment 2, we implement a two-by-two between-subject design. There are two games – deception and parallel dictator – and two treatments, direct and delegation. Our baseline deception game is adapted from the Erat and Gneezy (2012) "white lie" experiments, again drawing on the Lopez-Perez and Spiegelman (2012) dot framework.

[13]To some extent, responses to the belief questions suggest otherwise, with only 25 percent of delegating subjects indicating that \$1 Senders are less likely to be generous than \$2 Senders, and less than 20 percent indicating so in the restricted sample. However, subjects may nevertheless find it more difficult to align preferences over allocations with agent behavior in the dot experiment. For example, even among participants with beliefs consistent with more truthful behavior by \$1 Senders (our restricted sample), mean predicted probabilities of selfish behavior is very similar for the \$1 Senders (58% selfish) as for \$2 Senders (66% selfish), while mean predicted probabilities of deceit in the Gneezy game are significantly different for the \$1 Senders (45%) than for the \$2 Senders (70%).

[14]This effect is mitigated by a design that only pays for one of the two games and exhorts students to "make decisions in each Situation as if it is the Situation that will determine payments."

Direct Deception Game. Each Sender observes a dot that is either blue or green. A message is sent to the Receiver on behalf of the Sender indicating the true color of the dot. Two possible messages could be sent:

> **Message GREEN**: I solemnly swear that the dot is GREEN.
>
> **Message BLUE**: I solemnly swear that the dot is BLUE.

We specify "strong" messages in order to add saliency to the truthfulness of the Sender's choice. Based only on the message he or she receives, the Receiver must Report whether the dot is blue or green. Depending upon the Report, payments are made to the Sender and Receiver. The payments are as follows, depending upon whether the Receiver reports the true color of the dot or the incorrect color of the dot:

> If the Receiver reports *the true color* (say blue), then payments are $5 each to Sender and Receiver.
>
> If the Receiver reports *the incorrect color* (say green), then payments are $7 to the Sender and $3 to the Receiver.

The Sender earns more with an incorrect Receiver report, giving him/her a monetary incentive to lie by sending an untruthful message (Message Green in the case of a blue dot).

The Receiver is told that the Sender observes the payments associated with the Report, and nothing more about payments. To avoid strategic considerations, Senders (but not Receivers) are told that "in similar experiments, Receivers almost always report according to the Message they receive". This statement is based on the Erat (2013) experiments where 15 of 16 Receivers reported according to the message.[15] In our modified experiment, Receivers report according to the message in 80.4 percent of cases. 74.4 percent of Senders predict that their Receiver will follow their message.[16]

[15] In Erat (2013) and Erat and Gneezy (2012), the message and report reflect the outcome of a die roll, namely, a number from one to six. The expanded message space might be expected to increase Receivers' propensity to report according to the messages received, relative to our (two option) space; however, our stronger messages might also be expected to increase Receiver acceptance rates.

[16] Senders are paid $1 for a correct prediction of their Receiver's choice.

Senders participate in two “situations”, one of which is randomly selected for payment. The first situation – Situation K – is implemented with (2/3) probability and the second – Situation L – is implemented with (1/3) probability. In Situation K, the Sender chooses which Message to send, Message Green or Message Blue. In Situation L, the truthful Message (Message Blue, for example) is automatically sent on the Sender's behalf. This structure mimics the probabilities in the delegation treatment, as will become clear in a moment. It implies that a truthful decision in Situation K (Message Blue) prompts a truthful (blue) message with probability one; similarly, an untruthful decision (Message Green) leads to an untruthful message with (2/3) probability and a truthful message with (1/3) probability.

Delegated Deception Game. In the delegation treatment, Senders do not themselves choose the message. Instead, they choose between two Agents. Having observed the same color dot as the Sender, one of the Agents has chosen a Blue Message (truthful, say) and the other has chosen a Green Message (untruthful). The Sender chooses whether to delegate to the “blue” Agent or the “green” Agent (identified as Agent 1 and Agent 2 in the instructions, with labels randomly varied). Once an Agent is selected, the corresponding message choice of the Agent is sent to the Receiver with (2/3) probability and the truthful (blue) message is automatically sent with (1/3) probability.[17] This design implies that, if the Sender chooses the truthful (blue) Agent, the truthful (blue) message is sent with probability one; similarly, if the Sender chooses the untruthful (green) Agent, the untruthful (green) message is sent with (2/3) probability and the truthful (blue) message is sent with (1/3) probability. The probabilities of the different messages are identical to those under corresponding decisions in the direct treatment, as indicated in the summary Figure 2.

We adhere to a probabilistic effect of Agent selection on message choice in order to mimic the uncertainty that delegation yields in practice. In prior literature on delegation, the agent's decision is typically unknown and uncertain at the time a principal makes a decision. Our design aims to mimic some of this uncertainty without any differential loss of

[17]The Receiver instructions for the delegated deception game are slightly different than for the direct deception game. As the Message is not chosen directly by the Sender, the instructions indicate that the Message is chosen by another student who observes exactly the same color dot as the Sender and chooses a message to send on behalf of the Sender, knowing the payment options. The difference in instructions did not produce noticeable differences in Receiver choices.

control across the treatments.

Senders and Receivers are paid according to the same rules as in the direct deception game. Agents are not paid based on payoffs in the Sender-Receiver game. Rather they are paid according to the number of Senders who choose them as their Agent. If multiple Agents choose the same Message (say green), each has an equal probability of being designated as the corresponding (green) Agent in any one student's questionnaire.[18] Senders are told simply that "the Agents are not paid based on the payments attached to the Messages, but rather according to how many Senders (like you) choose them as their own Agent."[19]

Direct and Delegated Dictator Games. Parallel direct and delegated dictator games are also implemented, each with different subjects. Payoff options are identical to those described above: $5 each (Option A) vs. $7 for the Sender and $3 for the Receiver (Option B). For the delegated dictator game, the Sender/dictator chooses between two Agents, one of whom chose Option A (the "generous Agent") and the other of whom chose Option B (the "selfish Agent"). In both games, decisions are implemented probabilistically. With (2/3) probability, the Sender chooses the payment option (in the direct treatment) or the Agent (in the delegated treatment). With (1/3) probability, the equal-split ($5 each) option is automatically chosen. With this design, if a Sender chooses the generous option/agent, then Option A is selected with probability one. If a Sender chooses the "selfish" option/agent, then Option B ($7/$3 to Sender/Receiver) is selected with (2/3) probability and the other ($5 each) Option A is selected with (1/3) probability. In order to mimic Receiver acceptance decisions in the deception game (again following Gneezy, 2005), the "selected" option is implemented with 80 percent probability and the other option is implemented with 20 percent probability. For Experiment 2, the 80 percent probability corresponds exactly with actual Receiver following decisions.[20]

[18] In order to maximize the number of times that they are listed and selected on the Sender questionnaires, the Agents' monetary incentives (admittedly not our interest) are to choose (i) different actions than other Agents and (ii) actions that Senders are likely to prefer.

[19] Delegating dictators are given a similar statement: "The Agents are not paid based on the payment options, but rather according to how many Senders (like you) choose them as their own Agent."

[20] In Experiment 2, symmetric decisions produce identical probabilities of outcomes in the different treatments/games. A subject's decision to lie directly or through an agent yields identical probabilities that untruthful and truthful messages are sent (Figure 2). A subject's decision to be "selfish" in a dictator game

Logistics. Following the same general protocols as for Experiment 1, we conduct Experiment 2 in seven lower- and upper-division undergraduate economics classes at U.C. Merced. In total, there are 356 Sender-Dictator/Receiver pairs, with 87 Senders in the direct deception game, 94 in delegated deception, 88 in direct dictator, and 87 in delegated dictator.[21] Agents are from a small Political Science graduate class. We randomly vary dot colors (in the deception games), option labels (in the dictator games) and agent labels (in the delegation treatments).

Results from Experiment 2. Figure 3 and Table 2 present results from Experiment 2, indicating overall percentages of truthfulness in the deception games (direct and delegated) and generosity in the dictator games (direct and delegated). A truthful decision is indicated by a subject's choice of the truthful message in the direct treatment and by choice of the truthful agent in the delegated treatment. Overall, 74.7 percent of subjects were truthful in the direct deception game, compared with 48.9 percent of subjects in the delegated deception treatment and compared with 56.8 percent of subjects who were generous in the direct dictator game. Confirming the results of Experiment 1 with the new controlled design, we find that delegation reduces the propensity for truthfulness by 25.8 percentage points across our samples. Consistent with prior work (Gneezy, 2005; Hurkens and Kartik, 2009), we also find that lying aversion leads to a significantly higher level of truthfulness, vs. generosity, among subjects participating in the direct treatments.

Following our hypotheses in Section 2, we can gauge the statistical significance of the delegation effect on lying aversion in three ways. First is the raw difference between truth-

– whether directly or through an agent – produces identical probabilities that "selfish" and "generous" allocations are implemented. With an 80 percent probability that Receivers follow messages – the true probability in the experiment – a subject's decision to lie in the deception game also produces identical probabilities of "selfish" and "generous" allocations as does a "selfish" decision in the dictator game. In the on-line Appendix, we consider how heterogeneous Sender beliefs about the Receiver follow probability would affect inferences in the experiment, and describe plausible conditions under which difference-in-difference statistics (for Hypothesis 2) measure treatment effects on lying aversion in the presence of these heterogeneous beliefs.

[21]The proportion of male Sender subjects (our only available balancing variable) exhibits only minor variation across the treatments, ranging from a low of 51.5% (delegated deception) to a high of 57.9% (direct dictator), with 52.9% of subjects male in the other two treatments. The treatments are also roughly equally represented in all Sender classrooms. We can construct six pairwise Kolmogorov-Smirnov (KS) statistics for the null of a common distribution of observations across classrooms for the four different treatments; in each case, we order the seven classrooms in order to maximize the KS statistic; the smallest p-value for these statistics is 0.9849, indicating common distributions.

fulness in the direct and delegated deception games (Hypothesis 1). The test statistic for the null hypothesis of zero difference is 3.71 (p-value < 0.0001). Second is a difference-in-difference test (direct minus delegation, deception minus dictator) that controls for potential effects of delegation on preferences over the payoff allocations by netting out corresponding changes in dictator choices. Consistent with Hypothesis 2, the difference-in-difference is 20.7 percentage points; delegation reduces truthfulness by over 20 percentage points more than it reduces generosity. The test statistic for the null hypothesis of zero difference-in-difference is 2.02, significant at the (two-sided) 5 percent level. While we find that delegation increases "selfishness" in the dictator games, the effect is quite small (5 percent) and not significant. This result – not our focus, but an important control in our design – is loosely consistent with prior literature both in direction and magnitude.[22]

Third is a Hurkens-Kartik (2009) statistic for the hypothesis that the proportion of the "selfish" that are untruthful is higher under delegation than under direct choice (Hypothesis 3). For a between-subjects design like Experiment 2, the corresponding null hypothesis is that the ratio of proportions - the fraction of subjects who are untruthful in the deception game divided by the fraction of subjects who are "selfish" in the dictator game – is the same in the direct and delegation treatments. In the Appendix, we derive a test statistic that is approximately distributed as a standard normal under this null.[23] For our samples, this statistic is -2.242 (p-value 0.025), consistent with Hypothesis 3. We conclude that, for our payoffs, delegation leads to a significantly higher propensity for lies among the selfish.[24]

[22]Due to different experimental designs, we are reluctant to make direct comparisons between our dictator results and those from prior literature. This said, Bartling and Fischbacher (2012) find that the total rate of unfairness is a statistically insignificant 7.6 percentage points higher under delegation (59.1 percent) than it is with no delegation (51.5 percent) in their one-shot games with more extreme payoff choices than ours (z=0.90). Although the HLW design is very different with competing agents and a continuous choice of money-to-share they find no noticeable differences between baseline and agent treatments in the first several rounds, but significant differences in later rounds. To the extent that delegation promotes selfishness in one shot games, one would expect our design to produce smaller effects than in prior work because our Senders are confronted with precise information about what their chosen agent does on their behalf. Indeed, HLW (p. 1844) propose that in order to reintroduce the social pressure or obligation to behave altruistically, principals could be informed of their agents decisions and required to certify or override them, exactly as they do in our design. We believe our dictator results may be symptoms of this prescription.

[23]The statistic is slightly different than constructed by Hurkens and Kartik (2009), who assume that the true fraction of subjects preferring the selfish allocation is given by the corresponding sample proportion. The statistic derived in the Appendix is based on random sampling from all four distributions, deception and dictator, delegated and direct.

[24]In the delegated samples, the ratio of untruthfulness to selfishness, (0.5106/0.4828), is greater than one,

Panel B of Table 2 confirms conclusions of Panel A with regressions that control for course effects and gender. The regressions include all data from all four treatments. The endogenous variable is the zero-one indicator for a truthful or generous choice in, respectively, a deception and dictator game. The key regressors are the dummies for deception, delegation, and their interaction. The interaction captures the difference-in-difference effect of delegation on truthfulness vs. generosity. The models add controls as one moves from left to right, starting with the pure difference-in-difference Model 1, adding course effects in Model 2, gender effects in Model 3, and gender-treatment interactions in Model 4. In all models, the difference-in-difference interaction is statistically significant at five percent, with coefficients that are remarkably stable across models. The estimated delegation effect on truthfulness, over and beyond its effect on the propensity for generosity, is a reduction of between 20.7 percentage points (Model 1) and 21.7 percentage points (Model 2). Male gender has a significantly negative effect on truthfulness and generosity, but there is no evidence of significant differences in the gender effect across games and treatments.

The propensity for truthfulness is significantly greater than is the propensity for generosity, as indicated by coefficients on the Deception dummy. These coefficients confirm prior work documenting lying aversion in direct games. However, we cannot reject the hypothesis of a zero deception effect on the choices of subjects in the delegation treatments. Test statistics for the null of a zero sum of coefficients on the Deception dummy and the Deception-Delegation interaction have p-values ranging from 0.38 (Model 4) to 0.71 (Model 1). In fact, comparing our delegation samples, the propensity for truthfulness is less than the corresponding propensity for generosity, although the two are very close and not different in a statistical sense.

Overall, Experiment 2 confirms the conclusion of Experiment 1 that *delegation reduces lying aversion*. However, contrary to Experiment 1, Experiment 2 does not provide evidence that lying aversion persists under delegation. Experiment 3 provides some more direct evidence on both conclusions.

---

although the difference is not statistically significant. For this case, a potential alternative Hurkens-Kartik test would be for the null that the corresponding ratio for the direct samples equals one. The test statistic for this alternative null is -2.54 ($p<0.01$), giving rise to the same conclusion.

5 Experiment 3: The White Lie Experiment

The White Lie Deception Experiment. Experiment 3 mimics the design of Experiment 2 with one major difference: If the Receiver reports the incorrect color of the dot, both Sender and Receiver earn more money. Erat and Gneezy (2012) refer to this structure as a Pareto white lie situation (see also Gneezy, 2005). In our experiment, the specific payments are as follows:

> If the Receiver reports *the true color* (say blue), then payments are \$5 each to Sender and Receiver.
>
> If the Receiver reports *the incorrect color* (say green), then payments are \$6 each.

In Experiment 3, there is no reason for a subject to prefer the "truthful allocation" to the "untruthful allocation" because the latter is better for both players. However, lying aversion may nonetheless prevent subjects from sending an untruthful message that solemnly swears the dot is the incorrect color. The experiment identifies the extent to which delegation reduces lying aversion, promoting the delivery of more untruthful messages.

Logistics. Using the same broad protocols as in Experiments 1 and 2, we conduct Experiment 3 in three undergraduate economics classes at U.C. Merced, with 78 Sender-Receiver pairs, 39 Senders in the direct deception game and 39 in the delegated deception game. Direct and delegated treatments are equally represented in each classroom, with equally mixed questionnaires.[25] In the experiment, 86 percent of subjects predict that their Receiver will follow their message, supporting the design intent to avoid strategic considerations.

Results from Experiment 3. Figure 4 and Table 3 present results of the experiment. Two-thirds of subjects are truthful in the direct treatment, compared with 43.6 percent of subjects in the delegated treatment.[26] The 23 percentage point difference is large, statistically significant, and strikingly similar to delegation effects in Experiments 1 and 2. Panel

[25]The three Sender classrooms used in Experiment 3 are represented in the two (direct and delegated) samples with exactly the same proportions, roughly one-third each. The proportion of males in the direct and delegated samples are 46.1% and 43.6%, respectively.

[26]Perhaps surprising is the similarity between rates of deception in the "black lie" Experiment 2 and the "white lie" Experiment 3. There are offsetting effects: Experiment 3 (vs. 2) gives subjects a smaller own-benefit from lying, but gives Receivers a benefit (vs. cost) from the lie. Results from Erat and Gneezy

B of Table 3 presents regressions that control for course and gender effects, confirming the evidence from the coarse statistics.

Two conclusions emerge from Experiment 3. In the white lie context, *delegation again reduces lying aversion for a significant fraction of subjects.* In addition – and in contrast to Experiment 2 – Experiment 3 provides evidence that *delegation does not eliminate lying aversion.* Even with delegation, over 43 percent of subjects are truthful to their detriment, a percentage that is significantly different from zero (z-statistic 5.49, p-value < 0.001).

A Criticism of Experiment 3. We do not control for preferences over allocations in Experiment 3 because of the natural presumption that all will prefer the ($6,$6) payoff to the truthful ($5,$5) counterpart. However, perhaps delegation affects the *strength* of relative preference for the former over the latter. If subjects prefer to directly choose the superior allocation, rather than have it chosen indirectly via an agent, then our experimental results will understate the extent to which delegation reduces lying aversion; some subjects may then opt to tell the truth under delegation when they would not do so directly because they gain relatively less from the superior (untruthful) allocation. However, if subjects have a stronger preference for the superior allocation when it is obtained by a delegated choice than by a direct choice, then our experimental results will overstate the extent to which delegation reduces lying aversion.

We investigate this possibility with an additional experiment, the Plus-One design. In this experiment, there are two matched sets of Sender-Receiver pairs. In the first pair (players B and D), the Sender (player B) chooses either an initial equal-split allocation or a second/Plus-One allocation that adds $1 to the payoffs of both Sender (player B) and Receiver (player D). In the second pair of participants (players A and C, our interest), the Sender (player A) chooses between three alternatives: Option 1, implement the initial equal split; Option 2, implement the decision of the matched Sender B; or Option 3, directly add $1 to their own and player C's payoff, where the change is implemented with the same

(2012) suggest that the former (own payoff) effect should favor fewer lies in Experiment 3, whereas the latter (Receiver payoff) effect should favor more lies in Experiment 3. The small difference in lie rates between the experiments likely reflects this tradeoff.

probability as all player B's in the experiment choose the Plus-One allocation.[27] The order of the option choices is randomly varied.

In this experiment, indifference between the direct (Option 3) and delegated (Option 2) alternatives would produce an (approximately) equal fraction of subjects choosing each of the two options (2 and 3). If more subjects prefer the delegated implementation to the direct implementation of the $1 change, then we should see a larger proportion of subjects choosing the delegated Option 2 than choose the direct Option 3. Likewise, if more subjects prefer the direct to the delegated implementation, we should observe a larger proportion choosing the direct Option 3 versus the delegated Option 2.

The Plus-One experiment is implemented in one undergraduate economics class at U.C. Merced. Out of 48 "player A" subjects, 26 implement the Plus-One change directly and 11 choose to delegate (see Figure 5). The proportions (54.2% and 22.9%, respectively) are significantly different, indicating that more subjects prefer direct to delegated implementation.[28] This preference cannot be explained by probability differences, as the direct choice is implemented with the same probability as the delegated choice and choosers are informed of this probabilistic equivalence.

These results suggest that Experiment 3 *understates* the impact of delegation on lying aversion. For most subjects, delegation reduces the relative strength of preference for the superior allocation, which would reduce the propensity to lie in Experiment 3. The Plus-One results reflect the intuition that delegation is advantageous when it distances individuals from decisions that are harmful to others (and beneficial to oneself), but disadvantageous when it distances individuals from decisions that are beneficial to others. Here, the decision

[27] To implement the Plus-One experiment, all Sender players are given the three Options. Delegating (Option 2) players are each randomly matched with a non-delegating (Option 1 or 3) player and the delegate's (Option 1 or 3) choice is implemented for the delegating player. The set of non-delegating players represents the set of all "player B's" in the experiment for purposes of the probabilistic implementation of Option 3. Experimental instructions are consistent with this approach.

[28] The z-statistic for the difference in proportions is 3.95 (p-value < 0.001). Restricting the sample to the 37 subjects who do not choose the initial equal split, the z-statistic for the difference in proportions is 2.70 (p-value < 0.01). Eleven of 48 subjects choose the initial (low) equal split in the experiment. The latter (arguably surprising) choice was framed as a "no change" benchmark and likely represents a cognitive opt-out for some subjects.

is good news and the direct choice is preferred on average. This is the insight of Machiavelli (1532): "Princes should delegate to others the enactment of unpopular measures and keep in their own hands the means of winning favors" (see Bartling and Fischbacher, 2012).

6 Discussion and Conclusion

In our experiments, subjects are more willing to lie when the lie is made with an agent's message than when the deceptive message is sent directly by the subjects themselves. This is true even though outcomes from choices in the two treatments – deceptive agent vs. truthful agent in the delegation treatment and deceptive message vs. truthful message in the direct treatment – are exactly the same, and even though we control for delegation effects on preferences over the payoffs.

These results add to the growing literature on what determines individuals' propensities for truthfulness. To date, scholars have shown that lying aversion is affected by the consequences for both sides of the interaction (Gneezy, 2005; Gibson et al., 2013), social cues on how often others lie (Innes and Mitra, 2013), gender (Dreber and Johanneson, 2008), the extent of the lie (Lundquist et al., 2009; Fischbacher and Heusi, 2013), strategic considerations (Sutter, 2009), guilt aversion (Battigalli, Charness and Dufwenberg, 2013; Charness and Dufwenberg, 2010), team incentives (Conrads et al., 2013), and cooperation in prior play (Ellingsen et al., 2009) but not cooperative (vs. competitive) priming (Rode, 2010). Our results indicate that lying aversion is sensitive to delegation – that is, whether the decision is made directly or via choice of an agent.

This conclusion has implications for the use of delegation in markets, including vertical separation with outsourced suppliers and subcontractors. If delegation suppresses costly honestly – as we find – then firms have an added incentive to delegate, even beyond potential benefits of avoiding blame for decision that harm another party (Bartling and Fischbacher, 2012) and suppressing impulses for generosity that harm the principal's bottom line (HLW, 2010). Our results thus suggest an added economic motive for delegation, beyond the information, cost, strategic and commitment considerations identified in a rich literature on this subject (see, for example, Bolton and Dewatripont, 2005; Fershtman et al., 1991). From an empirical point of view, these benefits are likely to be particularly relevant in economic environments where deception is normal and advantageous, such as cultures with weak moral institutions. In these corrupt environments, both private economic benefits of dishonesty to contracting firms and opportunities for contracting with dishonest agents are likely to be greater. From a normative point of view, economic effects of contractual relationships are also likely to be more pernicious than they would be absent their impacts on lying aversion.

On the flip side, our results suggest that first party interactions are likely to be more truthful than are second party interactions. For example, private party sellers of used cars are more likely to be honest than are sales representatives in used car dealerships. This may help to explain (and justify) Gneezy's (2005) survey results indicating that students overwhelmingly have this belief.

In practice, the impact of agency on dishonesty is likely to be greater than suggested by our analysis. In the context of dictator games, Lazear, Malmendier and Weber (2012) show that the opportunity for exit – and the corresponding opportunity for selection into the dictator role – leads to selection in favor of more selfish dictators. While neither we nor they examine selection into the deception game, one might plausibly conjecture that self-selection will favor more dishonest agents (see Della Vigna et al., 2015, for a related study). Our results suggest an additional mechanism for the promotion of dishonesty, namely, the selection of agents by principals.

In addition, our Sender-Receiver experiments involve anonymous interactions with no opportunities for subsequent punishment. Audience effects (Andreoni and Bernheim, 2009) or opportunities for punishment (as in Bartling and Fischbacher, 2012, for example) might be expected to promote even greater effects of delegation on honesty by "shifting the blame."

Overall, our results speak to the importance of responsibility to social behavior, at least in the context of deception. Bartling and Fischbacher (2012) show that a measure of a subject's responsibility for a decision better explains punishment behavior of others at the receiving end than do measures of outcome-based or intention-based social preferences; that is, how others respond to our actions depends upon how responsible they think we are for those actions. Our results suggest that if a subject perceives a reduced responsibility for a lie, their own moral compass is altered – even absent any objective loss of true responsibility/control and no scope for response or full understanding by those at the receiving end of the lie. That is, responsibility alleviation (Charness, 2000) occurs when a lie is delegated even when control is not sacrificed.

7 References

Abbink, K., U. Dasgupta, L. Gangadharan, & T. Jain. 2014. "Letting the Briber Go Free: An Experiment on Mitigating Harassment Bribes." *Journal of Public Economics*, 111, 17-28.

Abeler, J., A. Becker, & A. Falk. 2014. "Representative Evidence on Lying Costs." *Journal of Public Economics*, 113, 96-104.

Andreoni, J. & B. D. Bernheim. 2009. "Social Image and the 50-50 Norm." *Econometrica*, 77(5), 1607-1636.

Andreoni, J. & J. Miller. 2002. "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica*, 70 (2) 737-753.

Bardsley, N. 2008. "Dictator Game Giving: Altruism or Artifact?" *Experimental Economics*, 11, 122-33.

Bartling, B., F. Engl, & R. Weber. 2014. "Does Willful Ignorance Deflect Punishment? An Experimental Study," *European Economic Review*, 70, 512-524.

Bartling, B. & U. Fischbacher. 2012. "Shifting the Blame: on Delegation and Responsibility." *Review of Economic Studies*, 79 (1), 67-87.

Basu, K. 2011. "Why, for a Class of Bribes, the Act of Giving a Bribe should be Treated as Legal." Working Paper, Government of India.

Battigalli, P., G. Charness, & M. Dufwenberg. 2013. "Deception: The Role of Guilt." *Journal of Economic Behavior & Organization*, 93, 227-232.

Bechara, A., H. Damasio, D. Tranel, & A. Damasio. 1997. "Deciding Advantageously Before Knowing the Advantageous Strategy." *Science*, 275, 1293-1295.

Bok, S. (1978). *Lying: Moral Choice in Public and Private Life.* New York. Pantheon Books.

Bolton, P., & M. Dewatripont. 2005. *Contract Theory.* MIT Press.

Cadsby, C., F. Song, & F. Tapon. 2010. "Are You Paying Your Employees to Cheat? An Experimental Investigation." *B.E. Journal of Economic Analysis & Policy*, 10(1), Article 35.

Charness, G. 2000. "Responsibility and Effort in an Experimental Labor Market," *Journal of Economic Behavior & Organization*, 42, 375-384.

Charness, G., R. Cobo-Reyes, N. Jimnez, J. Lacomba, & F. Lagos. 2012. "The Hidden Advantage of Delegation: Pareto Improvements in a Gift Exchange Game." *American Economic Review*, 102(5), 2358-2379.

Charness, G. & M. Dufwenberg. 2010. "Bare Promises: An Experiment." *Economic Letters*, 107, 281-283.

Charness, G. & M. Sutter. 2012. "Groups Make Better Self-Interested Decisions." *Journal of Economic Perspectives*, 26(3), 157-176.

Coffman, L. 2011. "Intermediation Reduces Punishment (And Reward)." *American Economic Journal: Microeconomics*, 3(4), 77-106.

Collier, P. 2007. *The Bottom Billion: Why the Poorest Countries are Failing and What Can Be Done About It.* New York. Oxford University Press.

Conrads, J., B. Irlenbusch, R.M. Rilke & G. Walkowitz. 2013 . "Lying and Team Incentives." *Journal of Economic Psychology*, 34, 1-7.

Cullen, R. 2008. *The Poverty of Corrupt Nations.* Toronto. Blue Butterfly Book Publishing.

Dana, J., D. Cain, & R. Dawes. 2006. "What You Don't Know Won't Hurt Me: Costly (But Quiet) Exit in Dictator Games." *Org. Behavior & Human Decision Processes*, 100(2), 193-201.

Dana, J., R. Weber, & J. Kuang. 2007. "Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness." *Economic Theory*, 33(1), 67-80.

Della Vigna, S., J. List, U. Melmendier, & G. Rao. 2015. "Voting to Tell Others." Working Paper, U.C. Berkeley.

DePaulo, B., D. Kashy, S. Kirkendol, N, Wyer, M. & J. Epstein. 1996. "Lying in Everyday Life." *Journal of Personality & Social Psychology*, 70(5), 979-995.

Dreber, A. & M. Johannesson. 2008. "Gender Differences in Deception." *Economics Letters*, 99, 197-199.

Dreze, J. 2011. "The Bribing Game." *Indian Express*, April 23.

Drugov, M., J. Hamman, & D. Serra. 2014. "Intermediaries in Corruption: An Experiment." *Experimental Economics*, 17, 78-99.

Dufwenberg, M. & G. Spagnola. 2015. "Legalizing Bribe Giving." *Economic Inquiry*, 53, 836-853.

Eisenkopf, G., & U. Fischbacher. 2015. "Nave Responses to Kind Delegation." *Managerial & Decision Economics*, forthcoming.

Ellingsen, T., M. Johannesson, J. Lilja, & H. Zetterqvist. 2009. "Trust and Truth." *Economic Journal*, 119, 252-276.

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

Erat, S. & U. Gneezy. 2012. "White Lies." *Management Science*, 58(4), 723-733.

Falk, A., & M. Kosfeld. 2006. "The Hidden Costs of Control." *American Economic Review*, 96(5), 1611-1630.

Fershtman, C., & U. Gneezy. 2001. "Strategic Delegation: An Experiment." *Rand Journal of Economics*, 32(2), 352-368.

Fershtman, C., U. Gneezy, & J. List. 2012. "Equity Aversion: Social Norms and the Desire to be Ahead." *American Economic Journal: Microeconomics*, 4(4), 131-144.

Fershtman, C., K. L. Judd, & E. Kalai. 1991. "Observable Contracts: Strategic Delegation and Cooperation." *International Economic Review*, 32: 551-559.

Fischbacher, U. & F. Föllmi-Heusi. 2013. "Lies in Disguise - An Experimental Study on Cheating." *Journal of the European Economic Association*, 11(3), 525-547.

Friedman, D., S. Huck, R. Oprea, & S. Weidenholzer. 2015. "From Imitation to Collusion: Long-run Learning in a Low-Information Environment." *Journal of Economic Theory*, 155, 185-205.

Friedman, E. 2011. "Romney Explains Record of Hiring Illegal Immigrants as Lawn Keepers." *ABC News*, October 18.

Gibson, R., C. Tanner & A.F. Wagner. 2013. "Preferences for Truthfulness: Heterogeneity among and within Individuals." *American Economic Review*, 103(1), 532-548.

Gneezy, U. 2005. "Deception: The Role of Consequences." *American Economic Review*, 95, 384-394.

Grossman, Z. 2015. "Strategic Ignorance and the Robustness of Social Preferences." *Management Science*, 60(11), 2659-2665.

Gurdal, M., J. Miller, & A. Rustichini. 2013. "Why Blame?" *Journal of Political Economy*, 121(6), 1205-1235.

Haisley, E. & R. Weber. 2010. "Self-Serving Interpretations of Ambiguity in Other-Regarding Behavior." *Games & Economic Behavior*, 68(2), 614-625.

Hamman, J., G. Loewenstein, & R. Weber. 2010. "Self-Interest Through Delegation: An Additional Rationale for the PrincipalAgent Relationship." *American Economic Review*, 100(4), 1826-1846.

Henrich, J., J. Ensminger, R. McElreath, A. Barr, C. Barrett et al. 2010. "Markets, Religion, Community Size, and the Evolution of Fairness and Punishment." *Science*, 327, 1480-1484.

Hurkens, S. & N. Kartik. 2009. "Would I Lie to You? On Social Preferences and Lying Aversion." *Experimental Economics*, 12(2), 180-192.

Innes, R. & A. Mitra. 2013. "Is Dishonesty Contagious?" *Economic Inquiry*, 51(1), 722-734.

Josephson Institute of Ethics. 2012. *2012 Report Card on the Ethics of American Youth.* Los Angeles: Josephson Institute.

Kahneman, D., J. Knetsch, & R. Thaler. 1986. "Fairness as a Constraint on Profit Seeking: Entitlements in the Market." *American Economic Review*, 76(4), 728-741.

Lazear, E., U. Malmendier, & R. Weber. 2012. "Sorting in Experiments with Application to Social Preferences." *American Economic Journal: Applied Economics*, 4(1), 136-163.

List, J. 2007. "On the Interpretation of Giving in Dictator Games." *Journal of Political Economy*, 115, 482-93.

Lopez-Perez, R. & E. Spiegelman. 2013. "Why do People Tell the Truth? Experimental Evidence for Pure Lie Aversion." *Experimental Economics*, 16(3), 233-247.

Lundquist, T., T, Ellingsen, E. Gribbe & M. Johannesson. 2009. "The Aversion to Lying." *Journal of Economic Behavior & Organization*, 70 (1-2), 81-92.

Machiavelli, N. 2003. *The Prince.* London: Penguin Classics.

McLean, B., & P. Elkind. 2003. *The Smartest Guys in the Room: The Amazing Rise and Scandalous Fall of Enron.* New York: Portfolio.

Muehlheusser, G., A. Roider & N. Wallmeier. 2015. "Gender Differences in Honesty: Groups versus Individuals." *Economics Letters*, 128, 25-29.

Nelkin, D. 2004. "Moral Luck." In: *The Stanford Encyclopedia of Philosophy*, E. Zalta, ed., Stanford University.

Oexl, R. & Z. Grossman. 2013. "Shifting the Blame to a Powerless Intermediary." *Experimental Economics*, 16: 306-312.

O'Rourke, D. 1997. *Smoke from a Hired Gun: A Critique of Nikes Labor and Environmental Auditing in Vietnam.* Transnational Resource and Action Center. San Francisco, November.

Rode, J. 2010. "Truth and Trust in Communication: Experiments on the Effect of a Competitive Context." *Games & Economic Behavior*, 68, 325-338.

Rosenbaum, S., S. Billinger & N. Stieglitz. 2014. "Lets be Honest: A Review of Experimental Evidence of Honesty and Truth-Telling." *Journal of Economic Psychology*, 45, 181196.

Sachs, J. 2011. *The Price of Civilization: Reawakening American Virtue and Prosperity.*. New York. Random House.

Sánchez-Pagés, S. & M. Vorsatz. 2007. "An Experimental Study of Truth-telling in a Sender-Receiver Game." *Games & Economic Behavior*, 61, 86-112.

Sánchez-Pagés, S. & M. Vorsatz. 2009. "Enjoy the Silence: an Experiment on Truth-Telling." *Experimental Economics*, 12: 220-241.

Sutter, M. 2009. "Deception Through Telling the Truth? Experimental Evidence from Individuals and Teams." *Economic Journal*, 119: 47-60.

60 Minutes. 2015. "Lumber Liquidators Linked to Health and Safety Violations." *CBS News*, March 1.

Appendix: The Hurkens-Kartik Statistic for Experiment 2

Let p_k = mean of Bernoulli distribution k, q_k = sample mean from k, and n_k = number of observations for sample k. We have four independent random samples, $k = 1$ (direct deception), $k = 2$ (delegated deception), $k = 3$ (direct dictator), and $k = 4$ (delegated dictator). Each Bernoulli observation takes a value of one if the choice is untruthful (selfish), and zero if the choice is truthful (generous). The null hypothesis is that the ratio of untruthful to selfish for delegated decisions equals the ratio of untruthful to selfish for direct decisions:

$$H_0 : (p_1/p_3) - (p_2/p_4) = 0 \longleftrightarrow (p_1p_4) - (p_2p_3) = 0; \qquad H_1 : (p_1p_4) - (p_2p_3) \neq 0$$

A few preliminary observations aid derivation of the test statistic for this null.
(i) $E(q_kq_m) = p_kp_m$ and $E(q_k^2q_m^2) = E(q_k^2)E(q_m^2)$ for $k \neq m$ by independence of observations;
(ii) $E(q_k^2) = E\{[\sum_i x_i^2 + \sum_i \sum_{j\neq i} x_ix_j]/n_k^2\} = (1/n_k)[p_k + (n_k - 1)p_k^2]$ where x_i and x_j are (Bernoulli) observations from sample k and the second equality follows from independence of observations ($E(x_ix_j) = p_k^2, i \neq j$) and $E(x_i^2) = p_k$;
(iii) for $k \neq m$, and using properties (i) and (ii),

$$\begin{aligned} V(q_kq_m) &= E\{(q_kq_m - E(q_kq_m))^2\} = E\{(q_kq_m)^2\} - E\{q_kq_m\}^2 = E(q_k^2) \cdot E(q_m^2) - p_k^2p_m^2 \\ &= p_kp_m\{1 + (n_k - 1)p_k + (n_m - 1)p_m - (n_k + n_m - 1)p_kp_m\}/(n_kn_m) \end{aligned}$$

(iv) by independence of samples, the variance of the difference in sample mean products is:

$$V(q_1q_4 - q_2q_3) = V(q_1q_4) + V(q_2q_3)$$

By these properties, the Central Limit Theorem and the Law of Large Numbers, the following statistic is approximately distributed as a standard normal under the null:

$$t_0 = \{(q_1q_4) - (q_2q_3)\}/\{[W(q_1q_4) + W(q_2q_3)]^{1/2}\} \overset{a}{\sim} N(0,1) \text{ under } H_0$$

where W = estimated $V = q_kq_m\{1 + (n_k - 1)q_k + (n_m - 1)q_m - (n_k + n_m - 1)q_kq_m\}/(n_kn_m)$ for $(k = 1, m = 4)$ and $(k = 2, m = 3)$.

Figure 1. Experiment 1 Results

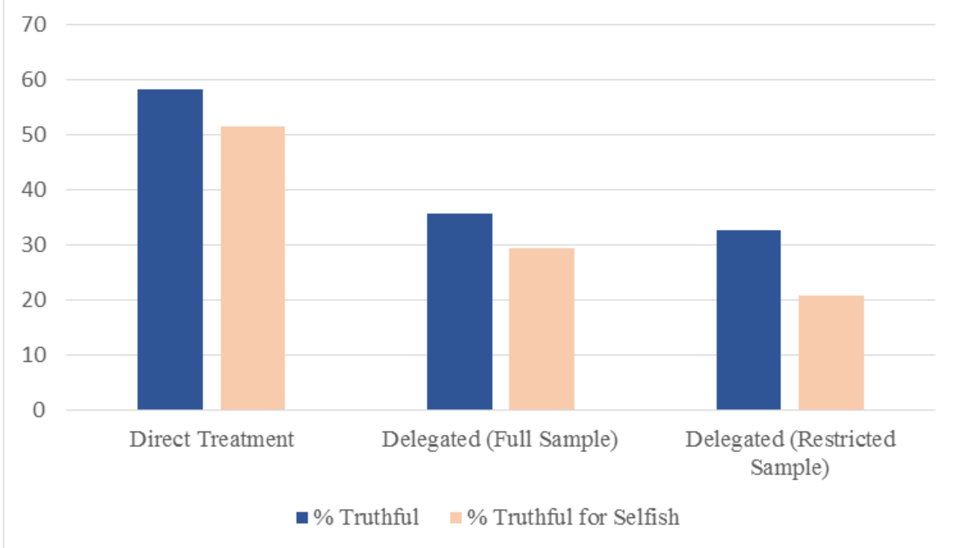

Figure 2. Deception Game Tree for Experiment 2 (When Blue is True)

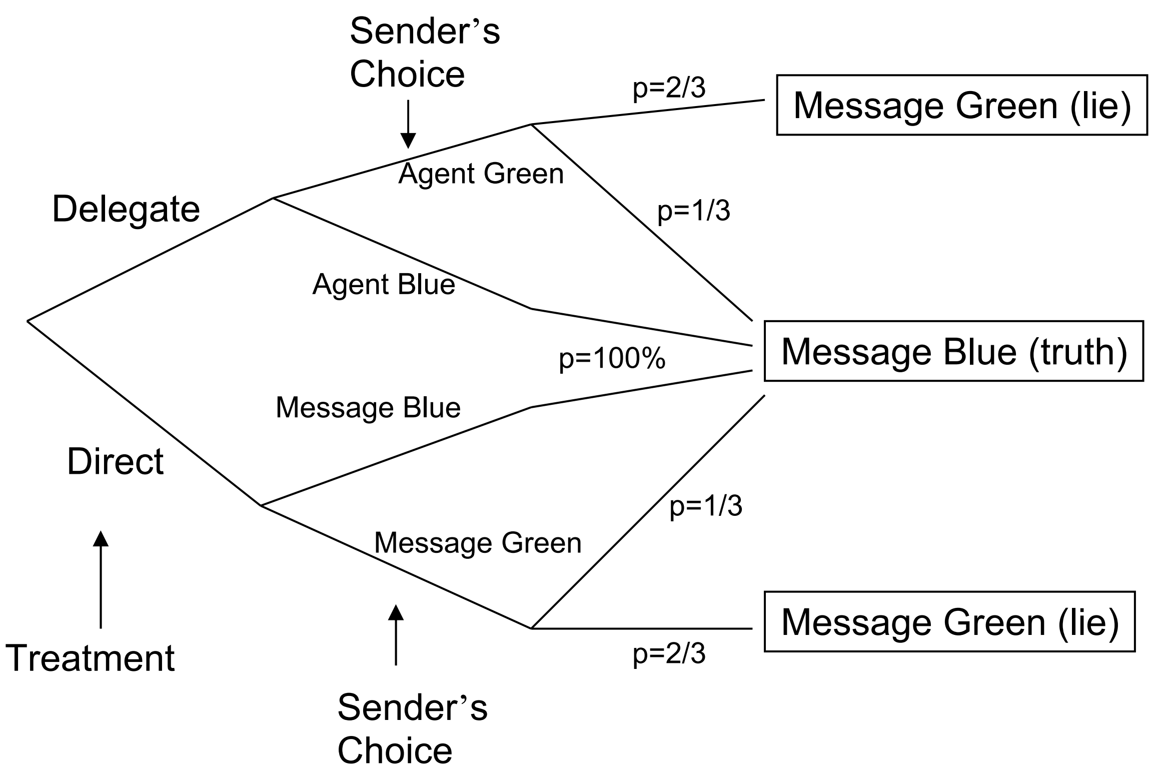

Figure 3. Experiment 2 Results

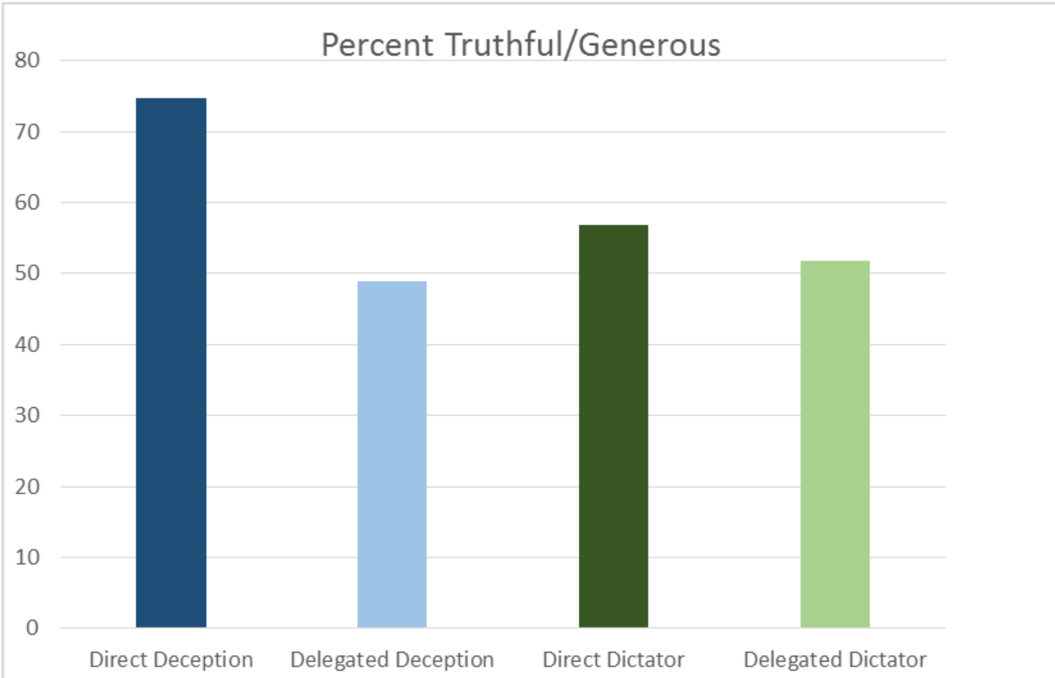

Figure 4. Experiment 3 Results

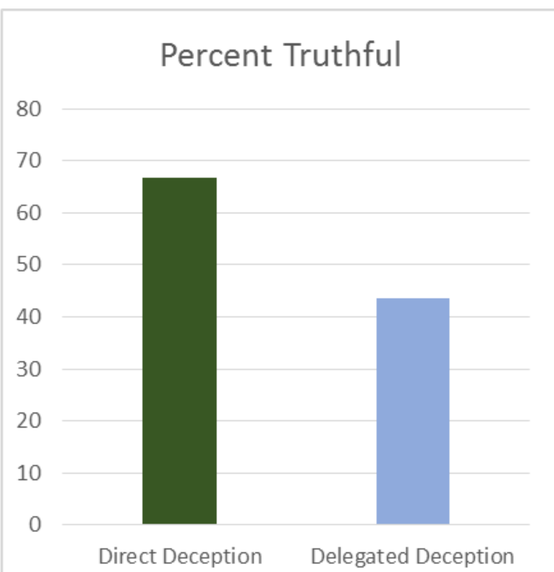

Figure 5. Plus One Results

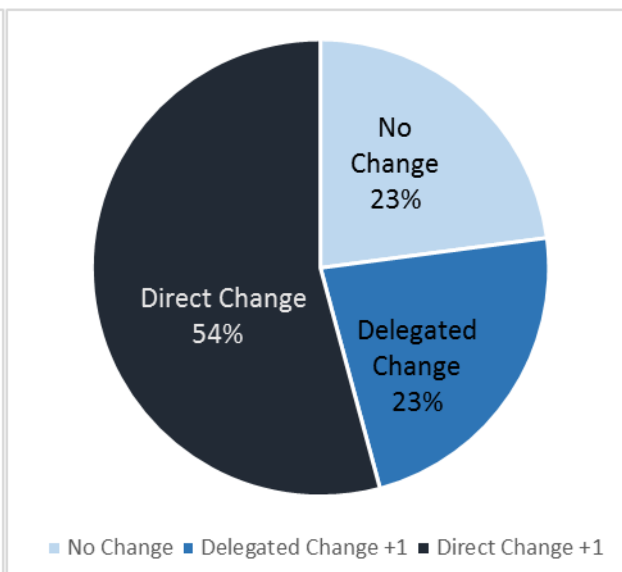

Table 1. The Dot Delegate Experiment

A. Summary and Difference Statistics

	Direct Treatment	Delegate Treatment Full Sample	Delegate Treatment Restricted Sample[A]	Difference (z-statistic) Full Sample	Difference (z-statistic) Restricted Sample[A]
% Truthful	58.33% (N=72)	35.71% (N=70)	32.65% (N=49)	22.62% (2.773)***	25.68% (2.828)***
% Generous	51.39% (N=72)	51.43% (N=70)	51.02% (N=49)	-0.04% (0.005)	0.37% (0.040)
Difference-in-Difference	--	--	--	22.66% (2.064)**	25.31% (2.189)**
Hurkens-Kartik: % Truthful for Selfish	51.43% (N=35)	29.41% (N=34)	20.83% (N=24)	22.02% (1.913)*	30.60% (2.585)**

B. Regressions[B]

	Overall Truthfulness	Hurkens-Kartik % Truthful for Selfish	Difference-in-Difference	
Dependent Variable →	Truthful	Truthful	Truthful minus Generous	
	(1)	(2)	(3)	(4)
	Marg. Effect (N=142)	Marg. Effect (N=69)	Marg. Effect (N=142)	Marg. Effect (N=142)
Delegate	**-0.2524 (0.0833)*****	**-0.2181 (0.1200)***	**-0.2241 (0.1093)****	**-0.2549 (0.1098)****
Gender (Male)	-0.0461 (0.0856)	-0.0885 (0.1215)	No	-0.0401 (0.1104)
Course Effects	Yes	Yes	Yes	Yes
R^2	0.0703	0.0639	0.0517	0.0615

Notes: *,**,*** denotes significant at 10%, 5% or 1% (two tail). [A] Restricted Sample: Subjects with reported beliefs consistent with more truthful behavior by $1 (vs. $2) Senders. [B] OLS regressions with robust standard errors in parentheses.

Table 2. The Black Lie Experiment

A. Summary and Difference Statistics

Treatment	Number of Observations	Percent Truthful/Generous
Direct Deception	87	74.71%
Delegated Deception	94	48.94%
Direct Dictator	88	56.82%
Delegated Dictator	87	51.72%

	Difference	z-statistic
Direct Deception – Delegated Deception	25.77%	3.71***
Direct Dictator – Delegated Dictator	5.10%	0.68
Difference-in-Difference	20.67%	2.02**

	Percent Untruthful / Percent Selfish		
	Direct	Delegated	Difference (z-statistic)
Hurkens-Kartik Statistic: Difference in Ratios	0.5857	1.0576	-0.4719 (-2.24)**

Table 2. The Black Lie Experiment (continued)

B. Regressions[A]

Dependent Variable: Truthful/Generous
N=357

	Model			
	(1) Marginal Effect (Standard Error)	(2) Marginal Effect (Standard Error)	(3) Marginal Effect (Standard Error)	(4) Marginal Effect (Standard Error)
Deception Dummy	0.1789 (0.0708)**	0.1830 (0.0692)***	0.1733 (0.0685)**	0.1296 (0.0871)
Delegate Dummy	-0.0509 (0.0756)	-0.0498 (0.0741)	-0.0596 (0.0726)	-0.0506 (0.0918)
Deception *Delegate	**-0.2068 (0.1030)****	**-0.2167 (0.1011)****	**-0.2101 (0.0994)****	**-0.2079 (0.0998)****
Course Effects	No	Yes	Yes	Yes
Gender (Male)	No	No	-0.1823 (0.0514)***	-0.2128 (0.0911)**
Male*Deception	No	No	No	0.0795 (0.1006)
Male*Delegation	No	No	No	-0.0200 (0.1003)
R^2	0.0410	0.0880	0.1200	0.1217

Notes:

[A] OLS regressions with robust standard errors.

*,**,*** denotes significant at 10%, 5% or 1% (two tail).

Table 3. The White Lie Experiment

A. Summary and Difference Statistics

	Percent Truthful	Number of Observations
Direct Sample	66.67%	39
Delegated Sample	43.59%	39
Difference (z-statistic)	23.08% (2.11)**	

B. Regressions

Dependent Variable: Truthful
N=78

	(1) Marginal Effect (Standard Error)	(2) Marginal Effect (Standard Error)
Delegate	**-0.2308** **(0.1110)****	**-0.2343** **(0.1105)****
Gender (Male)	No	-0.1388 (0.1100)
Course Effects	Yes	Yes
R^2	0.0770	0.0954

Notes:

[A] OLS regressions with robust standard errors.

*,**,*** denotes significant at 10%, 5% or 1% (two tail).