Self-interest through delegation:

An additional rationale for the principal-agent relationship

John R. Hamman, George Loewenstein, Roberto A. Weber

Department of Social and Decision Sciences
Carnegie Mellon University
Pittsburgh, PA 15213

June 1, 2009[*]

Abstract

Principal-agent relationships are typically assumed to be motivated by efficiency gains from comparative advantage. However, principals may also delegate tasks to avoid taking direct responsibility for selfish or unethical behavior. We report three laboratory experiments in which principals repeatedly either decide how much money to share with a recipient or hire agents to make sharing decisions on their behalf. Across several experimental treatments, recipients receive significantly less, and in many cases close to nothing, when allocation decisions are made by agents.

Keywords: fairness, principal-agent, experiments
JEL codes: C91, D03, D64

Self-interest through delegation: An additional rationale for the principal-agent relationship

Principal-agent relationships are typically assumed to be motivated by efficiency gains from comparative advantage. However, principals may also delegate tasks to avoid taking direct responsibility for selfish or unethical behavior. We report three laboratory experiments in which principals repeatedly either decide how much money to share with a recipient or hire agents to make sharing decisions on their behalf. Across several experimental treatments, recipients receive significantly less, and in many cases close to nothing, when allocation decisions are made by agents.

In the standard economic analysis of the principal-agent relationship, principals are assumed to hire agents because delegation confers efficiencies, as the agent either possesses special ability or has a lower opportunity cost of time or effort. The central focus of the literature on principal-agent relationships has been on how to design monitoring and incentive schemes that enable these advantages to be realized despite the fact that agents typically face different incentives and possess different information than the principals who employ them (Paul R. Milgrom and John Roberts 1992; Patrick Bolton and Mathias F. Dewatripont 2005).

This paper examines whether the principal-agent relationship might in some cases serve a function unrelated to efficiency. Specifically, a principal may hire an agent to take self-interested or immoral actions that the principal would be reluctant to take more directly. The principal may feel more detached, and hence less responsible, for such an action if it is delegated, while the agent may feel that he or she was "just carrying out orders" or merely fulfilling the requirement

1

of an employment contract. Through the use of agents, therefore, accountability for morally questionable behavior can become vertically diffused, with no individual taking responsibility.

While this function of agency has not been fully investigated in the economics literature, it is commonplace in popular accounts of behavior in domains as diverse as politics, business, war, and everyday social interaction. Companies are often accused of outsourcing production and other functions to outside firms that act less ethically than the company would act itself – for example, by treating workers less generously.[1] Executives and stockholders of the outsourcing company may turn a blind eye toward the actions of the outside firms, remaining deliberately uninformed, or at least pretending to be. Recent press stories have noted the increasing use of "firing consultants," who often serve little role other than carrying out the act of firing employees (Paul B. Brown 2007; Max Chafkin 2007). Moreover, firm managers who act unethically may justify such behavior through their role as shareholders' agents (Tara J. Radin and Martin Calkins 2006).

Within firms or other organizations, high-level decision makers are often accused of tacitly encouraging unethical or illegal behaviors, while shielding themselves from knowledge and hence responsibility. For example, top executives and board members of Enron may have tacitly encouraged the company's CFO and external auditors to engage in financial malfeasance on the firm's behalf, while limiting their own personal involvement in such matters (Daniel Eisenberg, 2002; Bethany McLean and Peter Elkind, 2003). Similarly, increased reliance on outsourced military operations over the last few decades – a practice also prevalent in previous centuries – may be linked to private entities' willingness to engage in acts beyond those

---

[1] Nike's alleged reliance on exploitative labor practices in outsourced plants in developing nations during the 1990's provides two examples. First, in response to complaints about labor practices in these factories Nike executives often claimed to have limited control over the actions of the outside manufacturing firms. Second, Nike further responded by outsourcing investigation of these factories to consultants who were allegedly instructed to produce biased and favorable reports (Dara O'Rourke, 1997).

permissible or politically feasible for conventional military forces, thus limiting public officials'
accountability and responsibility for casualties or heinous acts (Thomson, 1995; P. W. Singer,
2007).[2]

Our Study

       We test whether actions taken through agents are in fact more self-interested and less pro-
social than actions taken directly. We report results from three experiments in which principals
face a decision of how much money to share with a passive and anonymous recipient, an
experimental setup known as the "dictator game" (Robert E. Forsythe et al., 1994). In each
experiment, we contrast a *Baseline* condition in which principals (dictators) make the decision
themselves with *Agent* conditions in which principals select agents to make the decision on their
behalf.

       We predict and find that principals are significantly less generous toward recipients when
they select agents to allocate the money on their behalf than when they give directly. Principals
share significant amounts if choosing for themselves. But they replace agents who are no more
generous than they would be if they made the decision themselves with agents who share little or
nothing with the recipient. This drives sharing in our "agent" treatments significantly downward,
and in some cases almost to zero.

       Our experiments also explore conditions in which principals have the opportunity to
decide whether to act individually or through agents (Experiment 1) and in which agents
announce the amount they intend to share prior to principals making their agent selections
(Experiment 2). We find instances in both experiments in which sharing is virtually extinguished

---

[2] For example, much of the interrogation at Abu Ghraib prison in Iraq was conducted or overseen by independent
contractors (Renae Merle and Ellen McCarthy, 2004). Also, the government in Sudan is widely accused of relying
on outsourced militias in committing violent acts in Darfur (Colum Lynch, 2004).

– with less than 1 percent of the endowment shared on average – when allocations can be made through agents.

In addition to the effects on behavior, we predict and find that subjects acting individually in the baseline condition report feeling responsible for outcomes, but that principals acting through agents do not feel responsible for outcomes, even though their actions play a central role in producing such outcomes. We also find that, although principals acting through agents end up sharing substantially less, they nevertheless judge their behavior as being significantly fairer than do those acting alone.

Finally, we also explore the extent to which our results are driven by competition among agents, which is present in the first two experiments. We conduct an additional treatment in which each principal is paired with a single agent (Experiment 3) and can choose to "hire" or "fire" only that agent in each period. Consistent with the results above, and suggesting that the results are not driven by competition between agents, we find that principals generally prefer to act through agents and that such delegation results in very little sharing with recipients.

Related Literature

Extensive research in economics suggests that people not only maximize their own payoffs, narrowly construed, but also display a preference for social outcomes such as fairness or equality (Daniel Kahneman, Jack L. Knetch, and Richard H. Thaler, 1986; and see Colin F. Camerer, 2003 for a review). Research on the dictator game shows that even under conditions of complete anonymity dictators share significant amounts of money with unknown recipients (Elizabeth Hoffman et al, 1994). Such results have led to the development of several theoretical models that assume people derive utility from behaving generously, or disutility from unequal

outcomes (George Loewenstein, Leigh Thompson & Max H. Bazerman, 1989; Matthew J. Rabin, 1993; Ernst Fehr and Klaus M. Schmidt, 1999; Gary E. Bolton and Axel Ockenfels, 2000; James Andreoni and John H. Miller, 2002).

More recent work suggests that instead of valuing equitable outcomes, people may value the *perception* that their behavior is generous. For example, Jason Dana, Roberto A. Weber and Jason Kuang (2007) used a binary dictator game and varied the extent to which dictators were directly, uniquely, and knowingly responsible for outcomes. Experimental manipulations which gave dictators "moral wiggle room" caused them to behave less generously as compared to a baseline treatment.[3] Thus, when afforded the opportunity to implement self-interested and inequitable outcomes without feeling responsible for such outcomes, participants generally do so even though they are reluctant to implement such outcomes directly.[4]

One mechanism through which people may feel less responsibility for social outcomes is when others are involved in a decision. Studies of "diffusion of responsibility" find that individuals feel significantly less responsibility for their own actions when they are in a group than when they are alone (e.g. John M. Darley and Bibb Latane, 1968). These studies found, for instance, that individuals are less likely to aid someone in distress if there are others who could do the same, to the extent that less aid is rendered when more people are present (Latane and Darley, 1968, Darley, 2002). Building on the above work, one treatment from Dana et al (2007),

---

[3] See also Gary E. Bolton, Rami Zwick and Elana Katok 1998; James D. Konow 2000; Dana, Daylian M. Cain and Robyn Dawes 2007; Emily Haisley and Weber 2008. A separate line of research examines whether social behavior is best characterized by a preference for fairness or equality, or by a preference for some other social consideration such as efficiency (Gary B. Charness and Rabin, 2002; Dirk Engelmann and Martin Strobel, 2004) or equity (Alexander Cappelen, Astri D. Hole, Erik Ø. Sørensen, and Bertil Tungodden 2007).

[4] A plausible interpretation of these and other related findings can be found in models in which fair behavior is not motivated by a preference for fair or equitable outcomes, but is based on the desire to maintain a positive image, manifested through one's actions (J. Keith Murnighan, John M. Oesch, and Madan M. Pillutla, 2001; Drazen Prelec and Ronit Bodner, 2003; Roland J. M. Benabou and Jean M. Tirole, 2006; James Andreoni and B. Douglas Bernheim, in press). In these models, one avoids behaving unfairly because of what doing so implies about one's self. However, if the objectionable actions do not reflect directly on one 's self, then it is much easier to act self-interestedly.

in which more than one "dictator" was responsible for a recipients' outcome, demonstrated how diffusion of responsibility can lead to less fair behavior and outcomes.

Diffusion of responsibility not only affects perceptions of one's own behavior, but also how others judge such behavior. For example, recent studies find that people are less willing to punish individuals for delegating than for taking actions directly, even when the resulting outcome is the same (Björn Bartling and Urs Fischbacher 2007; Lucas Coffman 2009). These studies compare the punishment of dictators – by recipients or by neutral third-parties – when dictators either make choices themselves or delegate to another individual who has an incentive to share little with recipients. Dictators are punished significantly less when delegation produces unfair outcomes, than when they implement equivalent outcomes themselves.

Taken together, these findings suggest that the act of delegating might free dictators from blame or responsibility for unfair actions or outcomes. However, none of the research just summarized directly tests how delegation by a principal to an agent with no explicit interest in the outcome affects sharing or fair behavior. Moreover, none of this previous research examines delegation in settings where there are no strategic motives for delegating, such as avoiding punishment.[5]

Experiment 1 is presented in section 2. In section 3, we report findings from a second experiment in which the principal has ex ante information about agent behavior. We discuss the questionnaire results in section 4, and, in section 5, present the results of a third experiment that addresses the importance of agent competition. Section 6 concludes.

---

[5] Other studies explore the use of agents in bilateral bargaining, where an agent can allow strategic commitment (Chaim Fershtman, Kenneth L. Judd, and Ehud Kalai, 1991; Michael L. Katz, 1991; Kyle Bagwell, 1995;Eric van Damme and Sjaak Hurkens, 1997; Andrew Schotter, Wei Zheng, and Blaine Snyder, 2000; Fershtman and Uri Gneezy, 2001). For example, Fershtman & Gneezy (2001) study an ultimatum bargaining game in which players delegate their actions to an agent directly incentivized to earn them the highest returns. They find that delegation moves outcomes towards the Nash prediction, regardless of whether proposers or responders use agents. Unlike this work, however, our focus is on the use of agents in non-strategic settings.

I. Experiment 1

In Experiment 1, we compare allocations made in a *Baseline* treatment (the standard dictator game) with allocations produced in *Agent* treatments in which each principal selects an agent to make the decision on his or her behalf. We also explore what happens when subjects are allowed the *Choice* of whether to delegate or make the decision themselves.

A. Experimental Design

Subjects were recruited by email from a list of graduate and undergraduate students interested in participating in experiments at Carnegie Mellon and the University of Pittsburgh. Either 12 (Baseline) or 15 (Agent and Agent/Choice treatments) subjects participated in each session. Each participant received a $7 show-up fee plus any additional earnings from the experiment, which lasted about 30–45 minutes. All sessions took place at the Pittsburgh Experimental Economics Laboratory (PEEL) at the University of Pittsburgh. The experiment was conducted via computer interface, using z-Tree software (Fischbacher, 1999).

The experiment consisted of three conditions, described in Table 1. For consistency across conditions, we use the term "principal" to refer both to principals in the Agent and Agent/Choice conditions and to dictators in the Baseline condition. In each condition, 12 subjects played the dictator game – 6 as principals and 6 as recipients – for 12 rounds. One round of the dictator game was randomly selected to count for payment. The Agent and Agent/Choice conditions also included 3 subjects in the role of agent. Subjects randomly drew cards containing a letter (A, B, or C), which assigned them to their role, and a number that served to identify the subject for the experiment.

After role assignment, instructions detailing the rules of the game were distributed and read aloud (Appendix). In each round, principals and recipients were informed they would be randomly paired. They would know the ID number, but not the identity, of the person with whom they were paired. In each round, every principal received an allocation of $10 to be divided, in ten cent increments, between the principal and his or her randomly-paired recipient.

In the Baseline condition, principals made the allocation decision themselves by specifying an amount they wished to share with the recipient. The allocation was shown to the paired recipient, and the experiment proceeded to the next round.

In the Agent treatment, each allocation decision was made by one of the three agents. In the first round each agent was matched with two principals to ensure that all agents began with an equal number of decisions. In subsequent rounds (2-12), each principal selected one of the three agents to make the allocation decision on his or her behalf, by clicking on that agent's number. In all rounds, each agent made allocation decisions sequentially for each principal who had selected that agent, meaning that an agent selected by multiple principals could allocate different amounts on their behalf. Any agents not selected in a round saw a waiting screen while the other agents made allocation decisions. After all agents made their decisions, each resulting allocation was shown to the corresponding principal and recipient.

The Agent/Choice condition was identical to the Agent condition through Round 8, after which subjects received new instructions that applied to Rounds 9-12. For Rounds 9–12, in the Agent/Choice condition, each principal had the option of making the allocation decision on her own or continuing to select an agent to make the decision on her behalf.

In all conditions, once allocations were made in a round, subjects saw a results screen that showed the ID number(s) of other subjects with whom they were directly involved

(including that of the agent, if applicable), and the amount of money transferred from the principal to the recipient. Subjects recorded all results on their record sheets before continuing.

All subjects were told they would receive payment in cash at the conclusion of the experiment. Principals and recipients were paid for one randomly chosen round, drawn at the end of the experiment (plus the $7 participation fee). Agents, in contrast, cumulated payments received in each round. Each of the three agents began the experiment with $5 (plus the $7 participation payment) and received or lost additional payment in each round based on how many principals selected him or her to make the allocation decision. More precisely, agents' earnings in a round were based on the following payoff function:

$$\pi_i = -\$0.60 + \$0.30 n_i, \tag{1}$$

where n_i represents the number of principals (out of a possible 6) choosing agent i. This payoff function creates a setup in which agents, in effect, face a fixed cost of providing service ($0.60), zero marginal cost, and a constant marginal revenue ($0.30). The initial $5 can be thought of as starting capital. Agents break even in a round when they are "hired" by one-third – i.e., 2 – of the principals.[6]

Notice that an agent's earnings in a round are independent of any allocation decisions they make, and depend only on the number of principals selecting that agent. We chose this incentive structure for two reasons. First, it keeps the net total payment to all agents in a round fixed at $0 and makes the net surplus available to agents for the experiment equal to the per capita surplus available to principals and recipients ($5 per person, in addition to the $7

---

[6] We presented the payoff function to subjects in a slightly different format. We informed them that each agent would receive $0.20 for every principal who selected that agent, and would lose $0.10 for every principal that did not. When there are 6 principals (as in all rounds of the Agent condition and the first 8 rounds of the Agent/Choice condition), this is identical to the payoff function above. However, in the final four rounds of the Agent/Choice condition, this alternative format allows us to easily keep the zero-sum nature of the payoffs even with varying numbers of principals by having them only apply to principals not opting to make the choice for themselves.

participation fee), thus reducing net surplus and fairness considerations in agents' compensation and selection. Second, it avoids directly rewarding an agent for behaving in a fashion that maximizes the principal's payoff (which would make our hypothesized result more likely but less interesting), but instead allows the incentives for agents to arise more subtly, by principals' decisions of whether to retain an agent or switch agents across rounds. This is a "conservative" design since the agents must infer what decisions the dictators are likely to reward from their experience of being "hired" and "fired," and dictators can only communicate these preferences via their switching decisions.

Following 12 rounds of the game, but before learning the random round that would count for payment (for the principals and recipients), subjects filled out a brief questionnaire measuring perceptions of fairness, responsibility, and enjoyment of participating in the session (items provided 5 response choices ranging from "Strongly Disagree" (-2) to "Strongly Agree" (+2), with a "Neutral" (0) option). Subjects also answered some basic demographic questions.

Once subjects completed questionnaires, they saw their earnings and the randomly selected round. They then privately received payment.

B. Results

We conducted 19 sessions with a total of 264 subjects (see Table 1). We explore the amount shared with recipients by treatment in Experiment 1 to test our hypothesis that acting through agents facilitates the implementation of selfish goals. We defer analysis of the questionnaire data until after Experiment 2, when we pool the responses for both experiments.

Giving to Recipients

Table 2 shows the mean amount given to recipients in each round for all three treatments.[7] Recall that in Rounds 1-8, both Agent and Agent/Choice conditions were identical. Therefore, Table 2 presents both the pooled Agent conditions (for Rounds 1-8) and also each condition separately.

In Rounds 1 through 4, sharing in the Agent conditions was not significantly lower than in the Baseline. In Round 12 the amount shared in the Agent condition, in which principals were required to delegate to agents, is quite high ($2.97).[8] However, in Rounds 5 through 11, the average amount shared was always lower in both Agent conditions than in the Baseline – usually by over $1, and always by at least $0.50.

To compare overall sharing between the Baseline and the pooled Agent conditions, we computed the average amount shared by each principal over the first 8 rounds, when the two Agent treatments were identical. In the Baseline, principals (dictators) on average shared $2.26 over the first 8 rounds, while principals in the Agent conditions shared only $1.70, which is significantly lower in a t-test ($t_{110}=2.53$, $p=0.01$) and a Mann-Whitney (rank-sum) test ($z=3.34$, $p=0.001$). While the difference of roughly $0.50 is not large, this is because sharing in the Agent conditions started at high levels, similar to those in the Baseline.[9] Comparing average sharing in Rounds 5-8, we observe much larger and statistically significant differences (Baseline: $2.32;

---

[7] We discovered after running one Baseline session that two subjects had previously participated in the experiment, so their results were dropped from the analysis. Since principals' choices in the Baseline condition are independent – principals never find out what other subjects chose – we retain the choices of the remaining 4 principals in that session. Therefore, the total number of principals in the Baseline condition is 40 instead of 42.

[8] We find high sharing in Round 12 across all of our agent treatments. It is driven largely by several agents (8 of 42 in this case) choosing to give away all $10 in the final round, when there are no future possible repercussions from principals. One possible interpretation of this behavior is of agents expressing displeasure at having aided in treating recipients unfairly in all previous rounds. While we find this feature of the data interesting, our primary focus is on behavior prior to such end-game effects, and we therefore exclude Round 12 from most subsequent analysis.

[9] In Rounds 1-4, sharing is almost identical in the Agent conditions and in the Baseline (Baseline: $2.20; Agent: $2.29) and this difference is not statistically significant ($t_{110}=0.30$; $z=0.10$).

Agent: $1.11; $t_{110}=5.08$, p<0.001; z=5.20, p<0.001). In Rounds 5 to 8, delegation to agents produced less than half as much sharing on average relative to the Baseline. However, the above individual-level tests may exaggerate statistical significance due to the fact that outcomes for principals in the Agent treatment are not entirely independent, as they draw from the same agent pool. A more conservative test can be obtained by calculating the average amounts shared for Rounds 5 through 8 for each session, and using each *session* as the unit of observation. Even using this highly conservative test, we find that sharing in our Agent conditions is significantly lower than in the Baseline condition (Mann-Whitney z=2.70, p=0.007).

Figure 1 presents the distributions (cumulative frequencies) of amounts shared across conditions in Rounds 5-8. While in the Baseline at least $1 was shared 62 percent of the time, this generosity was much less frequent in the pooled Agent treatments, where it occurred only 30 percent of the time. Similarly, the frequency with which at least $5 was shared was much higher in the Baseline (28 percent) than in the Agent conditions (11 percent).

As we predicted, the Agent treatments decreased sharing relative to the Baseline, though this difference took several rounds to emerge. Next, we explore whether this difference is due, as we predicted, to principals switching away from agents who shared large amounts.

Agent Selection

Beginning in Round 2, all principals in the Agent conditions could choose among the three agents in their session, including the option of keeping the same agent as in the previous round. Agent selection provides insight into the preferences of principals. If principals value fair outcomes, as one might infer from the Baseline, they should retain agents who share amounts comparable to those shared in the Baseline condition. However, if principals prefer agents who

12

implement selfish outcomes on their behalf, they should search for agents who give very little or nothing at all to recipients.

The data confirm that principals pursued selfish outcomes through their switching choices.[10] For example, in Rounds 1 through 7 of the pooled Agent conditions, agents shared between $0 and $0.90 on a principal's behalf 288 times. In 260 of these instances (90 percent), the principal chose the same agent again in the next round. When the agent shared exactly $0, this proportion was slightly higher (93 percent). However, the frequency with which principals retained the same agent decreased substantially with the amount shared. Of the 66 cases in which agents shared $5, the principal retained that agent only 26 times (39 percent).

Table 3 presents logistic regressions (with subject fixed-effects) testing the degree of switching in Round t as a function of the amount shared by the agent in Round $t – 1$. The data consist of Rounds 2 through 8 from the pooled Agent conditions. The large, highly significant coefficient on amount shared indicates that the more an agent shared with the recipient, the more likely that the principal switched to a different agent in the subsequent round.[11]

Agent Treatment

In the final rounds of the Agent treatment, average sharing continued to be generally lower than in the Baseline condition. In Rounds 9 through 11, recipients in the Agent condition received $1.38 on average, which is significantly lower than the $2.10 received in the Baseline condition over these rounds ($t_{80}=2.19$, $p=0.03$; $z=2.67$, $p=0.01$).

---

[10] An implicit assumption in principals' search is that agents' behavior exhibits reliability across rounds. To explore such reliability, we regressed the average amount shared by an agent in period t on the average amount shared by that agent in period $t-1$. The relationship is positive and statistically significant, and does not vary across rounds.
[11] The results are substantively identical if, instead of including subject fixed-effects, we cluster errors by subject. We also conducted additional regressions that included subject demographic characteristics (age, gender, race) and interactions between these variables and amount shared in previous round, replacing fixed effects with subject random-effects. None of the additional coefficients were statistically significant, nor did their inclusion change the significance of any coefficients reported in the table.

A closer look at the Agent treatment reveals some cross-session heterogeneity. Of seven sessions, two produced considerably higher sharing than the others. Across Rounds 1-11, the average amount shared in these two sessions ($3.48) was much higher than in the other five sessions ($0.91). This difference is driven mainly by fair and unpredictable agents in those two sessions, which also led to significantly greater switching by principals.[12] Principals' search process in these two sessions was similar to the one used by principals in other sessions. Thus, the relatively large amounts given to recipients in these two sessions are not due to principals' willingness to share more, but rather from agents' unwillingness to serve as reliable collaborators. This suggests that the number and kinds of agents available in a market may play an important role in determining whether egalitarian outcomes ultimately result.

Final Rounds in the Agent/Choice Treatment

Recall that in the final four rounds of the Agent/Choice condition, principals could opt to make allocation decisions themselves. We hypothesized that principals would continue to select agents to make the decision on their behalf when subsequently given the opportunity to make decisions themselves, and sharing would remain low. This prediction was partially supported. As Table 4 shows, roughly forty percent of subjects continued to select an agent, while the remaining sixty percent opted to make the decision themselves.[13]

---

[12] Unfair agents in these sessions were initially harder to find: agents in those two sessions shared much more in Round 1 ($4.25) than agents in the other five sessions ($1.28). Agents in these two sessions were also less reliable: among agents selected in consecutive rounds by a principal, only 22 percent shared the same amount again in these two sessions, while 75 percent did so in the other five sessions. Principals were less likely to retain their previous agent in these two sessions (58 vs. 86 percent). Of twelve principals in these two sessions, nine (75 percent) switched three times or more, while of the thirty principals in the other five sessions only five (17 percent) did so.

[13] To determine whether the agent vs. self decision was influenced by prior experience with agents, we estimated a probit model that regressed each principal's choice of self vs. agent in Round 9 on the average amount shared by that principal's agents over Rounds 5-8 (clustering errors by session to account for lack of independence within sessions). The coefficient on prior agent sharing is statistically insignificant (p=0.23), indicating that something other than previous agent behavior appears to drive principals' choices to make decisions themselves.

Interestingly, those dictators making their own decisions behaved almost identically to agents, sharing very little (roughly $0.14 on average). In total, principals chose to make the allocation decisions themselves 73 times in the last four rounds and in these instances shared less than $1 68 times (93 percent), and zero 56 times (77 percent). While principals opting to make their own allocation decisions essentially played the last four rounds in the Baseline condition, they shared significantly less than what was shared in the Baseline. One likely interpretation of this low sharing is that those subjects who opted to make the allocation decision themselves are most likely the ones who felt comfortable sharing very little. That is, the 40 percent of subjects opting for an agent might be those who would feel compelled to share in a condition such as the Baseline, without really wanting to, and might therefore rely on the agents to implement a selfish allocation, as we hypothesized. The remaining 60 percent might be those who would not share very much anyway and therefore see little reason to rely on an agent to do so (see Lazear, Malmendier and Weber, 2006).[14] Thus, selection by principals into self- or agent-produced allocations can yield outcomes, as in Rounds 9-11, in which almost nothing is shared.

There is also one other interesting pattern in the last few rounds of the Agent/Choice treatment. More subjects are willing to share nothing (or very little) with recipients, even without resorting to using agents, than in the Baseline. For example, in the Baseline only 58 of 160 choices (36 percent) made by dictators in Rounds 9 through 12 shared nothing with the recipient. But, in the Agent/Choice treatment, in 56 of 120 total choices (47 percent), the principal made the choice herself *and* shared nothing with the recipient. Similarly, the proportion foregoing the agent *and* sharing 50 cents or less in the Agent/Choice treatment (67 of 120, 56 percent) is

---

[14] In fact, these dictators, who want the selfish outcome and have no problem implementing it themselves, may dislike the uncertainty associated with allowing an agent to make the choice. This seems especially important given that many agents gave away large amounts of money – including the entire $10 – in the final round of both the Agent and Agent/Choice conditions.

greater than the proportion sharing 50 cents or less in the Baseline (64 of 160, 40 percent). That is, without even considering the principals in the Agent/Choice treatment who share very little through agents, we see that principals are more comfortable personally sharing little or nothing in later rounds of the Agent/Choice treatment than in the Baseline. This suggests that the previous 8 rounds, in which recipients received very little when choices were made through agents, possibly desensitized principals to the recipients' fate or established a norm of very little sharing, which then carried over to the final four rounds in which principals could make allocation choices themselves. Therefore, aside from decreasing sharing directly, through the use of agents, the Agent treatment might have had the perverse effect of extinguishing the intrinsic motivation to share by principals, even once they returned to making choices themselves.

To explore such desensitization, we conducted three additional Agent sessions in which, for the last four rounds, principals were required to make their own allocation decisions (i.e., an "Agent/Baseline" condition).[15] Average sharing in Rounds 9-11 of this new treatment was $0.55. This is much lower than the average amount shared in Rounds 9-11 of the Baseline ($2.10), but is not quite as low as in Agent/Choice ($0.13).[16] Out of 18 subjects, 14 (78 percent) shared $0 in all of the last four rounds. The results confirm that previous experience with delegation reduces the sharing motive when subjects subsequently play the Baseline dictator game.

II. Experiment 2: Advertising Intentions

The findings from Experiment 1 show that sharing declines when individuals delegate allocation decisions to agents. Principals generally switched away from agents who shared

---

[15] We thank an anonymous referee for suggesting this treatment.
[16] Mean sharing in Rounds 9-11 of the new Agent/Baseline condition is significantly lower than in those same rounds in the Baseline ($t_{56}=3.02$, p=0.004; z=3.43, p<0.001), but is not significantly different from giving in those rounds of the Agent/Choice condition ($t_{46}=1.60$, p=0.12; z=0.22, p=0.83).

substantial amounts with recipients to those who shared very little. However, the mechanism through which principals searched for desirable agents was quite crude. Principals had very little information about the likely behavior of agents, and had to sample particular agents in order to learn about their behavior. Experiment 2 was designed to determine how outcomes would differ if principals had better information about the likely behavior of agents.

In Experiment 2, principals received non-binding information from agents regarding their intended behavior. Prior to principals making their selections, agents sent messages to all principals stating the amount they would share if chosen. Thus, agents engaged in a very simple form of "advertising" in which they communicated an intended action. This additional variant allows us to explore the robustness of our previous results to different institutions – in this case, one with more information for principals. It also allows us to test two hypotheses to better understand why delegation to agents produced less sharing in Experiment 1.

One hypothesis is that the pre-play messages will increase the speed with which principals find agents to act selfishly on their behalf, thereby producing more rapid convergence to low levels of sharing. If principals don't feel responsible when someone else is making the allocation decision, and their main goal is to search for someone to share very little on their behalf, then we would expect agents' pre-play messages to decrease sharing almost immediately, and for sharing to remain low throughout the experiment.

However, it is also possible that this treatment might produce higher levels of sharing than in Experiment 1. For example, perhaps principals were comfortable with outcomes in Experiment 1 because they could claim that they didn't know what the agent would do. When agents announce their intentions, and a principal chooses an agent who proposes to give little or nothing to the recipient, it might be difficult for principals to rationalize that they are not

responsible for the final outcome. With announcements, principals should be less able to claim "deniability" of foreknowledge of agents' actions.[17]

A. Experimental Design

Experiment 2 consisted of two conditions. The *Baseline* condition was identical to the Baseline dictator game in Experiment 1.[18]

The *Announce* condition was identical to the Agent condition in Experiment 1, with two important differences. First, at the beginning of each round, agent sent messages to principals stating an amount of the endowment that they proposed to give to the recipient. Each agent entered an amount from $0 to $10, in $0.10 increments, in a box on their screen. Each dictator saw the messages sent by all three agents, but recipients did not see any of the messages. The messages were "cheap talk," meaning that agents, if chosen, could give any amount to recipients regardless of the message they sent. Second, principals were not randomly assigned to agents in Round 1. Beginning in Round 1, principals selected any of the three agents after viewing the messages.

All other procedures – subject pool, location, payment, etc. – were identical to those in Experiment 1. We conducted 10 sessions (5 for each treatment), using 135 subjects.

---

[17] Agents have little incentive to deceive principals. We explore the extent to which agents' messages are truthful when analyzing the results.

[18] We collected a new set of Baseline sessions because Experiment 2 data was collected roughly one year after Experiment 1. New baseline data eliminates the possibility of confounding environmental and temporal factors when comparing our new Agent (Announce) data with the old Baseline data.

B. Results

Giving to Recipients

Table 5 shows the mean amounts shared in the Baseline and Announce treatments. Mean giving was always lower in the Announce condition, and this difference is statistically significant in almost every round. Overall, the mean amount given to recipients across all 12 rounds was $1.71 in the Baseline and $0.65 in the Announce condition (t_{58}=6.09, p<0.001; Mann-Whitney z=4.82, p<0.001), meaning that 62 percent less was shared with recipients when the decisions were made through agents.[19] This difference is even greater when we omit Round 12, in which many agents gave all $10 to recipients as in Experiment 1.[20] For Rounds 1-11, mean Baseline sharing was $1.69, compared to $0.43 in the Announce condition (t_{58}=7.58, p<0.001; z=5.88, p<0.001), a decrease of 75 percent. As we did for Experiment 1, we can perform a more conservative test to account for possible interdependence among principals who select from among the same agents in a session. If we use the average amount shared for rounds 5-11 in a session as the unit of analysis, sharing in our Agent conditions is significantly lower than giving in the Baseline (Mann-Whitney z=2.30, p=0.02, two tailed).

Towards the end of the experiment (in Rounds 10 and 11), the mean amount given to recipients was almost zero in the Announce treatment. In Round 10 only one of 30 recipients received any positive amount ($1.60), while in Round 11 only two did so ($1 and $1.40).

---

[19] Average sharing in the Experiment 2 baseline condition ($1.71) is lower than in the corresponding condition for Experiment 1 ($2.18). Note that both are close to the usual average amount shared of around $2 (see Camerer, 2003), but the difference is statistically significant (t_{68}=2.85, p<0.01; z=2.74, p<0.01). The instructions and procedures were identical, and subjects were drawn from the same participant pool, so this difference is somewhat surprising. For our purposes, however, the key comparisons are of each Baseline with its corresponding experimental treatments, which were collected concurrently. Note that if we had relied on the Experiment 1 Baseline data for comparison with the Announce treatment of Experiment 2, our hypothesized results would be stronger.

[20] In Round 12, 13 of 30 amounts shared (43 percent) were higher than the amount indicated in the message. Across the entire experiment, there were 62 instances (of 360 possible cases, or 17 percent) in which the agent shared a different amount than indicated in the message. A majority of lies (73 percent) were unfavorable to the principal (i.e., more was shared than indicated in the message).

Requiring agents to state their intended action decreased the amount shared with recipients almost immediately, in contrast to Experiment 1 in which sharing was initially high and decreased over several rounds. This could be either because agents behaved differently from the outset in the Announce treatment or because principals found it easier to select low-sharing agents with the added information. We find evidence of both effects.

First, the mean amount announced by agents in Round 1 ($1.16) was significantly lower than the mean amount shared by agents in the first round of Experiment 1 ($2.59), and this difference is statistically significant ($t_{85}=2.05$, $p<0.05$; $z=1.97$, $p<0.05$).[21] Moreover, of the 15 agents in Experiment 2, only 3 proposed sharing an amount greater than $2 with the recipient in Round 1, and 10 proposed sharing less than $0.50. Agents, therefore, clearly behaved differently at the outset of Experiment 2 than in Experiment 1, proposing and implementing less sharing.

Second, principals in Experiment 2 tended to select those agents in Round 1 who proposed sharing very little. In every session of Experiment 2, there was some heterogeneity in Round 1 agent proposals, with some agents proposing to share higher amounts. Among the 30 principals, 20 selected the agent in their session who proposed sharing the least in Round 1, while only 5 selected the agent who proposed sharing the most. Even those 5 principals who selected the agent with the highest proposals did so only in sessions where the high proposals were $0.20 and $0.40. There were three sessions in which an agent proposed sharing more than $2 in Round 1; none of the 15 principals in those three sessions selected those agents.

Figure 2 presents the distributions of amount shared with recipients in the Baseline and Announce conditions in Rounds 5 through 11. As in Experiment 1, much smaller amounts are

---

[21] These comparisons use the 15 Round 1 agent announcements from Experiment 2 and the 72 Round 1 choices made by 36 agents on behalf of 2 principals each from Experiment 1. If, for Experiment 1, we instead use the average amount shared by each agent (thus treating each of the 36 agents as an observation), the statistical significance is slightly stronger. We do not use the mean amount shared per agent for Experiment 2, as several agents were not selected to make an allocation decision even in Round 1 and this selection was not random.

shared with the recipient when decisions are made through agents than when principals make the decisions themselves. For example, sharing $1 or more occurs 45 percent of the time in the Baseline, but only 11 percent of the time in the Announce treatment.

Agent Selection

Principals were very sensitive to the announcements by agents. In the first round, 20 of 30 principals (67 percent) selected the agent in their session who proposed sharing the least. In the remaining rounds, 88 percent of principals selected the agent proposing to share the least, and this percentage was never below 80 percent in any round.

Table 6 presents estimates from logistic regressions, with subject fixed effects, of switching as a function of amount shared in the previous round.[22] As in Experiment 1, dictators generally switched away from agents who shared with recipients. The large positive coefficient on amount shared shows that principals switched away from agents who shared sizeable amounts with recipients.[23] Round number, unlike in Experiment 1, had a significant and negative effect on the propensity to switch agents, as seen in Model 2. This suggests that subjects settled in with their preferred agent after a few rounds and stopped switching, as we discuss above, while in Experiment 1 they were more likely to continue searching later into the experiment (especially in the two Experiment 1 sessions in which agents were unreliable and shared large amounts).

---

[22] The results are substantively identical if, instead of including subject fixed-effects, we cluster errors by subject. We also estimated additional models, with subject random-effects, including additional independent demographic variables and their interaction with amount shared. In one model females are slightly less likely to switch (p=0.09). No other coefficients for demographic variables or their interactions are statistically significant, nor does their inclusion change any of the results in Table 6. We also estimated models with binary variables indicating whether the agent lied to the principal in the previous round (sent a different amount than what was announced). We included variables for all lies, favorable lies, and unfavorable lies. The coefficients are never statistically significant, though the occurrence of these events is rare (see footnote 20).

[23] As for Experiment 1, we conducted a regression of the average amount shared by an agent in period t on the average amount that agent shared in period t-1. We again find a positive and statistically significant relationship. However, unlike in Experiment 1, the relationship becomes stronger across rounds.

Agents who developed a reputation for reliably sharing zero were highly successful in attracting principals. There were four agents, each in a separate session, who always announced that they would share either $0 or $0.10 and who always shared $0.[24] In every such case, there was at least one round in which that agent was selected by *every* principal in the session. That is, in four of five sessions, we observed an outcome in which the only agent "hired" was the one with the strongest reputation for sharing $0. These were the only agents in these sessions who were ever hired by more than 2 principals in any round after the first. Moreover, in the remaining session, Rounds 4 through 11 had exactly the same outcome: two agents attracted all the business (with one being selected by 4 principals and the other by 2), with both agents announcing they would share $0 and doing so. Reliable selfishness on principals' behalf was thus highly effective for obtaining "market power" as an agent.

III. Questionnaire Analysis

Since the basic structure of Experiments 1 and 2 is quite similar and the results of our questionnaire analysis do not differ much by experiment, we pool the analysis across experiments (we note where the results differ for the two experiments).

Table 7 displays mean responses, by role, for several questionnaire items. These items address perceived fairness, responsibility for the recipients' outcomes, and overall enjoyment of the experiment. In describing the questions, we replace the actual labels used in the experiment to describe a role (e.g. "Group A") with more informative labels ("principals").

As predicted, principals in the pooled Agent conditions (Agent and Agent/Choice from Experiment 1 and Announce from Experiment 2) felt significantly less responsible for the

---

[24] The $0.10 announcement corresponds to only one of these agents who announced this amount in Round 1 but then shared $0. Otherwise all four of these agents always announced and shared exactly $0.

amount given to recipients than did Baseline condition principals. Agent condition principals also felt they behaved more fairly than did Baseline condition principals, despite the more disparate earnings between themselves and recipients. Agent condition recipients, on the other hand, reported receiving significantly less fair treatment than recipients report in the Baseline condition.

Recipients in the Agent conditions judged the behavior of agents slightly more harshly than that of principals (-1.04 vs. -0.83). However, this difference is not statistically significant. When we consider this difference separately for each experiment, it is larger and statistically significant for Experiment 1 (Mann-Whitney z=1.88, p=0.06), but smaller and insignificant for Experiment 2. This is not surprising, as principals in Experiment 2 saw the proposed amounts to be shared by agents and then selected one of these agents. This foreknowledge of likely outcomes made it harder for them to avoid responsibility. Thus, similarly to Bartling and Fischbacher (2007), we find that recipients were more likely to blame agents than principals, but mainly in Experiment 1 when principals had less knowledge of what agents would do.

Table 8 shows several ordered probit regressions with questionnaire responses as dependent variables. The independent variables are average earnings across all rounds (recall that subjects completed the questionnaire before finding out which round counted for payment) and a binary variable indicating treatment (0 = Baseline, 1 = Agent, Agent/Choice, or Announce). Principals' perceptions of the fairness of their actions were heavily influenced (negatively) by how much they earned, but also by the condition – principals in the Agent conditions believed they behaved more fairly, after controlling for earnings. Conversely, recipients believed that they were treated less fairly in the Agent conditions, controlling for their average earnings. Principals' feelings of responsibility were also strongly influenced by the

23

experimental condition, with principals in the Agent conditions viewing themselves as significantly less responsible than those in the Baseline. Finally, with respect to enjoyment, we find no effect of treatment for principals, but recipients' enjoyment was lower in the Agent treatments.[25]

IV. Experiment 3: Eliminating Agent Competition

In Experiments 1 and 2, agents were forced to compete with each other for selection. Such competition might have played a role in the low sharing we observed. An important question, then, is whether delegation produces low sharing even without such competition among agents. To address this, we designed a new Agent condition in which each principal was assigned to a single, separate, agent for the duration of the experiment. Thus, in each session, there were 6 principals, 6 agents, and 6 recipients. Principals and recipients were randomly re-paired in every round, but each principal remained paired with the same agent throughout the experiment. In each round, principals decided whether to make allocation decisions individually or through their assigned agent.

A. Experimental Design

Procedures and software were similar to those used in Experiments 1 and 2. We conducted one treatment (Fixed-Agent), which we compare to the results from Experiment 2.

In each round, principals chose whether or not to delegate their decision in that round to their paired agent after receiving a message from the agent indicating the amount the agent

---

[25] We also estimated models including binary demographic variables (age, race, gender). While some of these variables have statistically significant coefficients (non-white principals report lower feelings of responsibility and older participants judge their own behavior as more fair), their inclusion does not change the significance of the variables we report in the table.

proposed to share (as in the Announce condition in Experiment 2). If selected, an agent received $0.50 for that round and made the allocation decision for the principal as in Experiments 1 and 2. If the principal chose not to select the agent in a round, the $0.50 payment that the agent would have received was divided equally among the other five agents and the principal made her own allocation choice in that round. This payment design allowed us to preserve the incentive structure of the other Agent conditions while maintaining constant total revenue in the agent pool.

Note that this design retains one key feature of our previous experiments – the agent wants to be selected by the principal in each round – while eliminating competition among agents since each principal/agent pair is independent of all other agents.

B. Results

We conducted two sessions of this new treatment. The average amounts shared by round are depicted in Figure 3, along with results from Experiment 2, which also contained messages from agents. Overall, the average amount shared is significantly lower in the Fixed-Agent than in the Baseline ($0.80 vs. $1.71, $t_{40}=3.65$, $p<0.001$; Mann-Whitney $z=3.27$, $p=0.001$). This difference is even stronger if we look only at Rounds 5-11, in which the average amount shared is $0.37 in the Fixed-Agent condition and $1.61 in the Baseline ($t_{40}=4.72$, $p<0.001$; $z=4.76$, $p<0.001$).

In the Fixed-Agent condition, it is possible that low sharing might be driven entirely by principals themselves choosing to make their own allocation decisions. However, the data do not support this interpretation; over all twelve rounds, principals chose to delegate their decision 66 percent of the time. Looking only at Rounds 5-11, in which sharing is lowest, principals

delegated even more frequently, choosing to have their agent make the decision 79 percent of the time.[26] Questionnaire responses again reveal that Fixed-Agent principals feel less responsibility than Baseline dictators (z = 2.27; p = 0.023).

As in the Announce condition of Experiment 2, we also see a decrease across rounds in the proposed amounts sent by the agents. In Round 1, the average proposed amount was $1.95, which dropped to $0.64 in Round 5 and to $0.13 in Round 8.

V. Conclusion

The research presented here suggests that principal-agent relationships can serve functions beyond those usually attributed to them in the economics literature. Specifically, the intermediation of agents increases the extent to which principals pursue self-interested outcomes. Sharing decreases considerably when decisions are made through agents, and those agents willing to act most self-interestedly on principals' behalf are able to attract most – and in some cases, all – of the principals' business. Moreover, acting through agents allows principals to maintain positive impressions of their own behavior and role in determining outcomes.

Agents serve this function through a subtle interplay of psychological factors. Principals do not feel that they are behaving unfairly because they do not directly take immoral actions; they simply hire agents. They also do not feel responsible for the ultimate outcomes. Such positively-biased perceptions about their own behavior are reflected in their questionnaire responses. Agents, who are randomly-selected from the same population as baseline dictators, behave much more self-interestedly on principals' behalf than they would likely behave if they were dictators themselves. Given the incentives they face, they may feel they have no choice: "I

---

[26] In Rounds 5-11, principals who made the allocation decision themselves shared slightly more on average ($0.45) than those who delegated the decision ($0.35). This difference is not statistically significant.

was just following orders," "I'm just trying to make a living," or "if I didn't do it he would have hired someone else to do it," are emblematic of the types of phrases one often hears, after the fact, in investigations of atrocities, business fraud, and other forms of immoral behavior. In the real world, these psychological factors often interact with others that reinforce the evasion of responsibility – for example, self-serving biases, self-deception, and willful disregard of information (Linda Babcock & Loewenstein, 1997; Niklas Karlsson, Loewenstein & Duane Seppi, forthcoming; Dana, Weber & Kuang, 2007; Haisley & Weber, 2008). It is also worth again highlighting an interesting finding across our experiments: many agents "retaliate" in the final period of the experiment by giving most or even all of the money to the recipient. This suggests that the psychological forces at work in the agency relationships we study here can also backfire against the principals.

These results may, if anything, understate the impact of agency on selfishness because our study is conservative on several dimensions. In our experiments, a principal could select from at most three agents, which introduced a significant likelihood that principals would be unable to find *any* agent who would act selfishly on their behalf. Indeed, as we discussed previously, in two Agent sessions of Experiment 1 principals repeatedly sought out agents who would act selfishly, but were unable to find them, and the result was very high sharing. In many economic settings there are likely to be larger pools of available agents from which to choose, and agents are likely to have well established reputations, which, consistent with the Announce results of Experiment 2, would tend to strengthen the effects observed in our experiments. Moreover, agents in real-world environments are unlikely to be representative members of the population, and may in fact self-select into that role because of their willingness to act immorally (cf. Edward P. Lazear, Ulrike M. Malmendier & Weber 2007).

A limitation of the study is that it examined a principal-agent relationship with only one kind of contract, whereby agents receive a fixed payment. In real economic relationships, there are many kinds of possible contracts, with varying incentives for the agent. The selection of contract, in some cases, is endogenous to the relationship. It would be interesting to explore how behavior and outcomes change under different kinds of contracts. However, given the relatively neutral monetary incentives provided in the contract in our setting, it seems likely that an incentive contract designed by the principal would only serve to further decrease sharing.

A general take-away message from the current research is that agents may not always serve the function that they are ostensibly hired to perform. Our study suggests that agents may sometimes be used to avoid taking responsibility for personally desirable but unethical outcomes, but similar functions in other domains may also be possible. For example, the more thrifty member of a couple may propose to hire a financial advisor, not truly to provide investment advice, but to convey to the spendthrift partner that he or she needs to cut back on spending. Similarly, a restaurant may provide diners with the opportunity to delegate their dietary intake to someone likely to enhance flavor at the expense of a healthy diet, while the customer maintains some degree of uncertainty or distance regarding the exact ingredients going into the food.

At a practical level, our work also points to potential institutional arrangements that could reduce immoral behavior in hierarchical situations. The two high-sharing sessions in the Agent treatment of Experiment 1 show that when the set of potential agents is sufficiently small and those agents are socially conscious (evidenced by the high amounts of sharing in the first period), principals' search for self-interested outcomes may be hindered. Therefore, regulating how many and what kinds of individuals act as agents in some contexts may be socially beneficial. Conversely, the more rapid decrease towards low sharing in Experiment 2 relative to Experiment

1, when all that differed was agents' ability to send messages, suggests a role for regulating the kinds of advertising employed by agents in some contexts.

Our experiments suggest that simply removing principals from the act of determining allocations produces more immoral behavior. A new treatment could further address this issue: Once the agent makes the decision, a principal could be forced to see the decision made on her behalf and have the option of overriding it or certifying that they favor implementing the action. Such a mechanism could potentially reintroduce the social pressure or obligation to behave altruistically. Indeed, a mechanism somewhat along these lines was part of the Sarbanes-Oxley bill that was enacted in the wake of Enron and other corporate scandals. It requires CEOs of companies to personally vouch for the accuracy of the books, thus undermining the common excuse of those at the top that they were unaware of, or otherwise not responsible for, malfeasance implemented by those acting on their behalf.

References

Andreoni, James and Bernheim, B. Douglas. (in press). "Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects." *Econometrica*.

Andreoni, James, and Miller, John H. (2002) Giving According to Garp: An Experimental Study of Rationality and Altruism, *Econometrica*, 70, 737-753

Babcock, Linda, & Loewenstein, George (1997) Explaining bargaining impasse: the role of self-serving biases. *Journal of Economic Perspectives*. 11, 109-126.

Bagwell, Kyle (1995) Commitment and Observability in Games, *Games and Economic Behavior*, 8, 271 – 280.

Bartling, Björn, and Fischbacher, Urs (2007) Shifting the Blame: On Delegation and Responsibility, University of Zurich working paper.

Benabou, Roland J. M., and Tirole, Jean M. (2006) Incentives and Prosocial Behavior, *American Economic Review*, 95 (5), 1652-1678..

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

Bolton, Patrick. and Dewatripont, Mathias (2005). Contract Theory. Cambridge: MIT Press.

Brown, Paul B. (2007) "Letting a Stranger Do the Firing," New York Times, November 10.

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

Cappelen, Alexander, Hole,Astri D, Sørensen, Erik Ø., and Tungodden, Bertil (2007) The pluralism of fairness ideals: An experimental approach, *American Economic Review*, 97, 818-827.

Chafkin, Max (2007) "Meet Rebecca. She's Here to Fire You," Inc. Magazine, November.

Charness, Gary. and Rabin, Matthew J. (2002) Understanding Social Preferences with Simple Tests, *Quarterly Journal of Economics,* 117(3), 817-869.

Coffman, Lucas (2009) Intermediation Reduces Punishment, Harvard University working paper.

Dana, Jason, Cain, Daylian M., Dawes, Robyn (2005) What you don't Know Won't Hurt me: Costly (but quiet) Exit in a Dictator Game, *Organizational Behavior and Human Decision Processes*, 100(2), 193-201.

Dana, Jason, Weber, Roberto A., and Kuang, Jason Xi (2007). "Exploiting 'moral wriggle room': Experiments demonstrating an illusory preference for fairness." *Economic Theory* 33(1): 67-80.

Darley, John M. (2002) Crowded Minds: The Implicit Bystander Effect, *Journal of Personality and Social Psychology*, 83, 843-853.

Darley, John M., and Latane, Bibb (1968) Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, 8, 377-383.

Eisenberg, Daniel (2002). "Is Ken Lay a Victim Too?" Time Magazine, Sunday, February 3.

Engelmann, Dirk and Strobel, Martin (2004) Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments, *American Economic Review*, 94, 857-869.

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

Fershtman, Chaim and Gneezy, Uri (2001) Strategic Delegation: An experiment, *RAND Journal of Economics*, 32, 352-368.

Fershtman, Chaim, Judd, Kenneth L., and Kalai, Ehud (1991) Observable Contracts: Strategic Delegation and Cooperation, *International Economic Review*, 32, 551-559.

Fischbacher, Urs (1999) z-Tree - Zurich Toolbox for Readymade Economic Experiments - Experimenter's Manual, Working Paper Nr. 21, Institute for Empirical Research in Economics, University of Zurich.

Forsythe, Robert E., Horowitz, Joel L., Savin, N.E., and Sefton, Martin. (1994) Fairness in Simple Bargaining Experiments, *Games and Economic Behavior*, 6, 347 - 369.

Haisley, Emily and Weber, Roberto A. (2006). "Favorable Impressions of Ambiguity and Unstable Preferences for Fairness." Working paper.

Hoffman, Elizabeth, McCabe, Kevin A., Shachat, Kariv, Smith, Vernon (1994) Preferences, property rights and anonymity in bargaining games, *Games and Economic Behavior*, 7, 346-380.

Kahneman, Daniel, Knetch, Jack L., Thaler, Richard H. (1986) Fairness and the assumptions of economics, *Journal of Business*, 59, 285-300.

Karlsson, Niklas, Loewenstein, George, and Seppi, Duane J. (forthcoming). The 'Ostrich Effect': Selective Attention to Information about Investments, *Journal of Risk and Uncertainty.*

Katz, Michael L. (1991) Game-Playing Agents: Unobservable Contracts as Precommitents, *RAND Journal of Economics*, 22, 307-328.

Konow, James D. (2000) Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions, *American Economic Review*, 90 (4), 1072-1091.

Latane, Bibb and Darley, John M. (1968) Group inhibition of bystander intervention, *Journal of Personality and Social Psychology*, 10, 215-221.

Loewenstein, George, Thompson, Leigh, & Bazerman, Max H. (1989) Social utility and decision making in interpersonal contexts. *Journal of Personality and Social Psychology*, 57, 426-441.

Lazear, Edward P., Malmendier, Ulrike M., Weber, Roberto A. (2005) Sorting in Experiments. Stanford working paper.

Lynch, Colum (2004). Rights group says Sudan's government aided militias. *Washington Post*, Tuesday, July 20, A12.

McLean, Bethany. and Elkind, Peter (2003). The Smartest Guys in the Room: The Amazing Rise and Scandalous Fall of Enron. Penguin: New York.

Merle, Renae and McCarthy, Ellen (2004). "6 Employees from CACI International, Titan Referred for Prosecution." Washington Post, Thursday, August 26; page A18.

Milgrom, Paul and Roberts, John (1992). *Economics, Organization, and Management*. Englewood Cliffs: Prentice Hall.

Murnighan, J. Keith, Oesch, John M., and Pillutla, Madan M. (2001). Player Types and Self Impression Management in Dictator Games: Two Experiments, *Games and Economic Behavior*, 37(2), 388-414.

O'Rourke, Dara (1997). "Smoke from a Hired Gun: A Critique of Nike's Labor and Environmental Auditing in Vietnam as Performed by Ernst & Young," Transnational Resource and Action Center report: San Francisco, November 10, http://nature.berkeley.edu/orourke/PDF/smoke.pdf.

Prelec, Drazen and Bodner, Ronit (2003) "Self-signaling and self-control,"' Time and Decision, G. Loewenstein, D. Read, and R.F. Baumeister (eds.) Russell Sage Press, New York, 2003.

Rabin, Matthew J. (1993) "Incorporating Fairness into Game Theory and Economics," *American Economic Review,* 83, 1281-1302.

Radin, Tara J. and Calkins, Martin. (2006). "The struggle against sweatshops: Moving towards responsible global business," *Journal of Business Ethics*, 66: 261-272.

Schotter, Andrew, Zheng, Wei, and Snyder, Blaine (2000) Bargaining through agents: An experimental study of delegation and commitment, *Games and Economic Behavior*, 30, 248-292.

Singer, P. W. (2007) "The dark truth about Blackwater," Salon.com, October 2.

Thomsom, Janice E. (1995). *Mercenaries, Pirates, and Sovereigns*. Princeton: Princeton University Press

van Damme, Eric and Hurkens, Sjaak (1997) Games with Imperfectly Observable Commitment, *Games and Economic Behavior*, 21, 282-308.

Table 1: Description of Experiment 1 by Condition (each cell indicates who made the allocation decision)

	Baseline	Agent	Agent/Choice
Rounds 1 – 8	Self	Agent	Agent
Rounds 9 – 12	Self	Agent	Agent/Self option
Sessions	7	7	5
N *(principals, agents, recipients)*	84 (42, n/a, 42)	105 (42, 21, 42)	75 (30, 15, 30)

Table 2: Mean amount given to recipient by treatment (Experiment 1)

Round	Baseline (n = 40)	Pooled Agent conditions (n=72)	Statistical comparison with Baseline (a)	(b)	Agent (n = 42)	Statistical comparison with Baseline (a)	(b)	Agent/ Choice (n = 30)	Statistical comparison with Baseline (a)	(b)
1	$2.59	$2.59	0.27	0.00	$2.13	0.72	0.84	$3.23	1.45	1.08
2	$1.87	$2.83	1.24	1.72*	$3.16	1.00	1.94*	$2.37	1.15	0.95
3	$2.31	$2.02	1.06	0.58	$2.35	0.72	0.07	$1.54	1.14	1.48
4	$2.05	$1.74	1.42	0.60	$1.57	1.74*	0.86	$1.96	0.56	0.14
5	$2.40	$1.45	1.62	1.90*	$1.34	1.84*	1.92*	$1.60	0.83	1.29
6	$2.38	$1.10	2.21**	3.01***	$1.24	2.08**	2.23**	$0.90	1.66*	2.90***
7	$2.26	$0.91	2.81***	3.70***	$0.91	2.79***	3.20***	$0.90	1.92*	2.78***
8	$2.25	$0.98	2.90***	3.03***	$1.18	2.42**	2.23**	$0.71	2.54***	3.01***
9	$2.23				$1.73	1.64*	0.96	$0.13	4.51***	5.60***
10	$2.12				$1.38	2.02**	1.50	$0.18	4.44***	4.84***
11	$1.95				$1.04	2.13**	2.01**	$0.08	4.12***	4.53***
12	$1.81				$2.97	0.43	1.62	$0.59	2.55***	2.68***

a: Mann-Whitney rank-sum (z); b: t-statistic
* = p<.10; ** = p<.05; *** = p<.01; all two-tailed

Table 3: Logistic regression of switching in round *t* based on amount given in round *t – 1* (Experiment 1, data pooled from two agent conditions, Rounds 2 – 8)

	Model 1	Model 2
Amount shared by agent in previous round	0.53^{***} (0.08)	0.49^{***} (0.08)
Round		-0.09 (0.08)
N	336 (48)	336 (48)
Log-likelihood	-88.66	-87.90

Subject fixed-effects; 24 subjects dropped due to all positive or all negative outcomes
* - p<.10; ** = p<.05; *** - p<.01; all two-tailed

Table 4: Results of Rounds 9 – 12 in Agent/Choice condition of Experiment 1 (Number of allocation decisions in parentheses)

	Allocation Decision		*Average Amount Shared*		
Round	*By Agent*	*By Self*	*Agent*	*Self*	*Baseline*
9	43% (13)	57% (17)	0.17^{***}	0.10^{***}	$2.23
10	40% (12)	60% (18)	0.08^{***}	0.19^{***}	$2.12
11	40% (12)	60% (18)	0.08^{***}	0.08^{***}	$1.95
12	33% (10)	67% (20)	$1.51	0.13^{***}	$1.81

Statistical comparison with Baseline (Mann-Whitney rank-sum): * - p<.10; ** = p<.05; *** - p<.01

Table 5: Mean amount given to recipient by treatment (Experiment 2)

Round	Baseline (n = 30)	Announce (n=30)	Statistical comparison with Baseline (a)	(b)	Amount announced by agents (n=15)
1	**$1.86**	**$0.77**	2.18**	2.26**	$1.16
2	**$2.00**	**$0.52**	2.27**	3.19***	$1.25
3	**$1.98**	**$0.45**	2.34**	3.19***	$1.04
4	**$1.45**	**$0.15**	3.61***	3.73***	$0.83
5	**$1.89**	**$0.27**	3.60***	3.44***	$1.18
6	**$1.80**	**$0.23**	2.85***	3.53***	$0.67
7	**$1.73**	**$0.79**	2.90***	1.68*	$1.42
8	**$1.50**	**$0.47**	2.86***	2.10**	$0.31
9	**$1.39**	**$0.87**	2.25**	0.90	$0.53
10	**$1.82**	**$0.05**	4.26***	4.31***	$1.02
11	**$1.16**	**$0.08**	3.49***	3.23***	$1.11
12	**$1.92**	**$3.13**	0.71	1.33	$0.62

a: Mann-Whitney rank-sum (z); b: t-test, two-tailed
* = p<.10, ** = p<.05, *** = p<.01. All tests using two-tailed analysis

Table 6: Logistic regression of switching in round t based on amount given in round $t – 1$ (Experiment 2, Rounds 2 – 12)

	Model 1	Model 2
Amount shared by agent in previous round	0.75*** (0.22)	0.61*** (0.20)
Round		-0.18*** (0.07)
N	165 (15)	165 (15)
Log-likelihood	-62.40	-58.37

Subject fixed-effects; 15 subjects dropped due to all positive or all negative outcomes
* - p<.10; ** = p<.05; *** - p<.01

Table 7: Questionnaire responses (Experiments 1 and 2 combined)

Question	Principals		Agents	Recipients		Significant Difference (rank-sum)
	(A)	(B)	(C)	(D)	(E)	
	Baseline (n=70)	*Agent (n=102)*	*Agent (n=51)*	*Baseline (n=70)*	*Agent (n=102)*	
I feel responsible for how much the [recipients] were allocated.	0.83	-0.09	0.71	--	--	A/B, p<.001 B/C, p<.001
I {made fair decisions / was treated fairly} in this experiment.	-0.09	0.32	0.00	-0.80	-1.28	A/B, p=.04 D/E, p<.001
[Principals], in general, made fair decisions in this experiment.	--	--	--	--	-0.83	
[Agents], in general, made fair decisions in this experiment.	--	--	--	--	-1.04	
I enjoyed this experiment.	0.83	0.73	0.08	-0.03	-0.61	A/C, p<.01 B/C, p<.01 D/E, p<.001

Five-item scale, from strongly disagree (-2) to strongly agree (+2), with a neutral option (0).

Table 8: Ordered probit regressions of questionnaire responses (Experiments 1 and 2 combined)

	Dictators / Principals			Recipients	
	Fair Decisions	Responsi-bility	Enjoyment	Fair Treatment	Enjoyment
Treatment (Agent = 1)	0.50*** (0.18)	-0.77*** (0.15)	-0.24 (0.18)	-0.42** (0.18)	-0.38*** (0.11)
Average Earnings	-0.27*** (0.06)	-0.04 (0.07)	0.13** (0.07)	0.22*** (0.08)	0.26*** (0.05)
N	172	172	172	172	172
Log likelihood	-254.98	-258.37	-223.44	-212.78	-243.41

Standard errors in parentheses (clustered by session)
* - p<.10; ** = p<.05; *** - p<.01

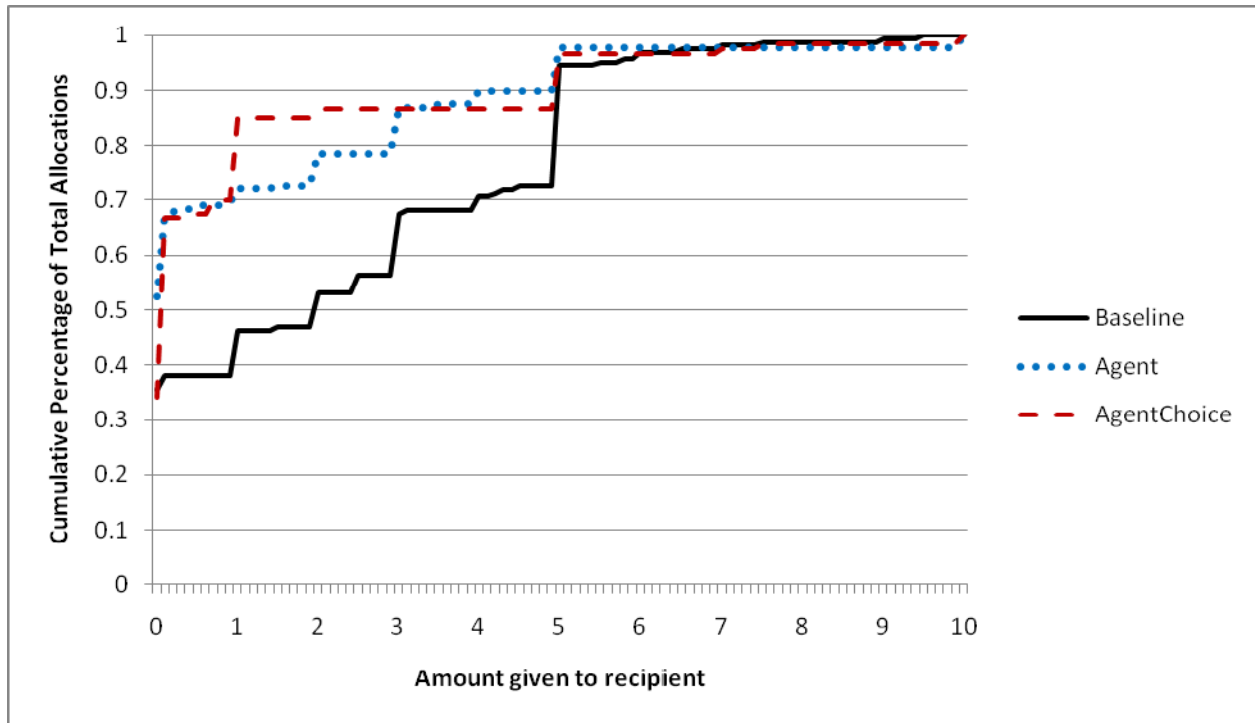**Figure 1: Distributions of amount shared by condition (Experiment 1, Rounds 5 – 8)**

Figure 2: Distributions of amount shared by condition (Experiment 2, Rounds 5 – 11)
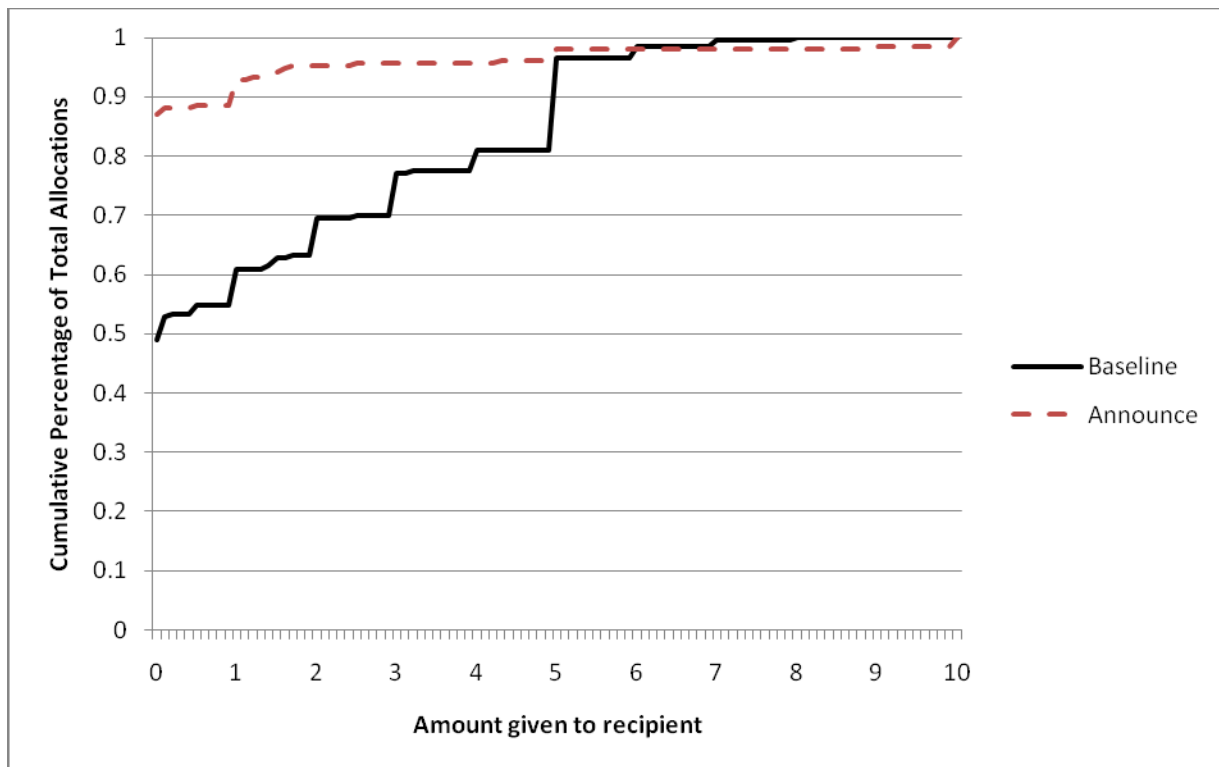
Figure 3: Mean amount shared by round (Pooled Baseline (Experiments 1 & 2), Announce (Experiment 2), and Fixed Agent (Experiment 3) conditions)