

Bryant University

## Bryant Digital Repository

---

Economics Faculty Journal Articles

Economics Faculty Publications and Research

---

6-2016

### It's Not Just the Thought That Counts: an experimental study on hidden cost of giving

Xiaofei Pan

Bryant University, [xpan@bryant.edu](mailto:xpan@bryant.edu)

Erte Xiao

Monash University, Australia, [erte.xiao@monash.edu](mailto:erte.xiao@monash.edu)

Follow this and additional works at: [https://digitalcommons.bryant.edu/econ\\_jou](https://digitalcommons.bryant.edu/econ_jou)

 Part of the [Economics Commons](#)

---

#### Recommended Citation

Pan, Xiaofei and Xiao, Erte, "It's Not Just the Thought That Counts: an experimental study on hidden cost of giving" (2016). *Economics Faculty Journal Articles*. Paper 37.  
[https://digitalcommons.bryant.edu/econ\\_jou/37](https://digitalcommons.bryant.edu/econ_jou/37)

This Article is brought to you for free and open access by the Economics Faculty Publications and Research at Bryant Digital Repository. It has been accepted for inclusion in Economics Faculty Journal Articles by an authorized administrator of Bryant Digital Repository. For more information, please contact [dcommons@bryant.edu](mailto:dcommons@bryant.edu).

## **It's not just the thought that counts: An experimental study on the hidden cost of giving**

Xiaofei Pan and Erte Xiao<sup>1</sup>

### **Abstract**

Receiving a gift can create an impulse to reciprocate, even when doing so may be inefficient and potentially harmful to a third party. This paper provides a theoretical framework for a pure gift effect on reciprocity impulses and experimental evidence that such an effect exists: that is, a gift receiver will favor an actual gift giver over an intended gift giver, even if the intended gift giver incurred the same costs and signaled the same intention to give. This result contrasts with the predictions of existing theories on social preferences. We also show that the pure gift effect is present even when it leads to a less efficient outcome, or when the gift is given without the expectation of future returns. Our findings suggest that when reciprocating a gift becomes socially inefficient, it may be more advantageous to guard against gift receiving or to keep donations “secret” than to try to control the intent to give.

*Keywords: gift exchange; intention; reciprocity; efficiency; externalities; social preferences*

*JEL: C91, D03, D73*

**Acknowledgements.** We thank Omar Al-Ubaydli, Marco Castillo, Daniel Houser, David Laibson, Joshua Miller and seminar participants at George Mason University, Peking University, University of Pennsylvania, University of Michigan, University of Siena, University of Sydney, and the University of the South, the ESA northern American conference 2013 and the Birmingham and CUFÉ workshop 2013 for helpful comments and discussions.

---

<sup>1</sup>Xiaofei (Sophia) Pan is an assistant professor at the University of the South: Sewanee, [xpan@sewanee.edu](mailto:xpan@sewanee.edu). Erte Xiao is an associate professor at Monash University, [erte.xiao@monash.edu](mailto:erte.xiao@monash.edu)

## 1. Introduction

Gift giving is an important social phenomenon and is ubiquitous in human society. The act of giving triggers a receiver's obligation to reciprocate (Mauss, 1954, Cialdini 2001). The receiver often feels compelled to return the favor, even when there is no monetary benefit to doing so. Thus, gift giving can be socially optimal, as it initiates social ties and facilitates cooperative relationships (Akerlof, 1982; Carmichael and Macleod, 1997; Falk, 2007). Nonetheless, there is also a dark side to gift giving: a receiver's reciprocal behavior may harm a third party's interest and even lead to inefficient social outcomes (Cialdini, 2001). For example, gift giving and reciprocity have been identified as playing an important role in sustaining vote buying and vote selling that could potentially undermine the desired effects of the democratic electoral process (Finan and Schechter, 2012). In this paper, we investigate the underlying mechanism responsible for a receiver's reciprocal behavior when there is negative externality.

Previous studies of *two-person* exchange environments absent externalities have identified intentions as one key to understanding reciprocal behavior (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Al-Ubaydli and Lee 2009). This literature argues that people are strongly disposed to infer the intentions behind others' actions, and to respond favorably to kindness and negatively to unkindness. People react to the same outcome differently if they infer different intentions motivating others' actions. The importance of intentions in two-person reciprocal relationships raises the question of to what extent the intention behind the gift predicts the receiver's reciprocal behavior towards the giver, even though such reciprocity harms a third party and reduces social welfare. Does the receiver favor the gift giver at the cost of third parties only because the gift giver has signaled a good intention, while other third parties have not? If a receiver demonstrates favoritism only because she<sup>2</sup> wants to reciprocate the gift giver's good intention, then a policy that forbids the recipient from receiving the gift will not help to mitigate favoritism. The reason is that the gift giver has already signaled good intentions by attempting to give. Our study aims at understanding the role of intentions in a gift exchange relationship with externalities.

We draw attention to the difference between receiving gifts and receiving intentions. The literature in philosophy on moral luck has noted a difference between a person who "tried hard to

---

<sup>2</sup>For simplicity, we use "she" to refer to the receiver and "he" to refer to the gift giver and the third party, rather than using he/she.

help but failed” and one who actually succeeded in helping, even if her success was the result of random luck (Nagel, 1979; Williams, 1981; Scanlon, 2008; Nelkin, 2013). This situation is directly relevant to controlling reciprocal behavior with negative externalities. For instance, in the case of vote-buying, electoral candidates offer cash or goods in exchange for the vote. The theory of moral luck suggests we will observe less vote-selling if voters are required to abstain from receiving gifts. However, the theory of intention may predict no change as long as a candidate has already conveyed the intention of giving. For instance, when multiple candidates compete for the vote from the same group of citizens (primarily those who have no particular affiliations with any candidate), moral luck theory predicts that voters would favor the candidate from whom they actually received the gift. And this is true even when voters know that other candidates would also be just as willing to offer the same (but not-received) bribe<sup>3</sup>. In naturally occurring environments, it is of course difficult to know the difference between receiving gifts and receiving intentions. The reason is that it is nearly impossible to know precisely what voters know about the intentions of all the candidates when examining their reciprocal behavior and to identify the underlying mechanism for such reciprocity. Laboratory experiments can control this information and help us better understand the mechanisms underlying an individual’s reciprocal behavior. They can also provide a theoretical foundation for designing policies to curb socially undesirable reciprocal behavior.

We develop and test the hypothesis that a receiver will favor the gift giver from whom she *receives* a gift over a third party who paid the identical cost to signal the intention of giving. In other words, we hypothesize that people feel in debt to *the gift* they receive. We call this a *pure gift effect*. To test this hypothesis, we design an experiment based on previous experiments used to study bribery and corruption<sup>4</sup>. In these experiments, a bribery relationship between the briber and the bribee is modeled using a gift exchange game with negative externality. The bribee’s reciprocal behavior is socially undesirable because it harms third parties (Abbink et al.,2002;

---

<sup>3</sup>We can think of bribery in this case as a competitive market where candidates compete to successfully bribe a voter. Candidates are willing to pay a voter an amount up to the value of the return to the bribe. Like any market, the success of winning a voter’s vote is often determined not only by the candidate’s own intentions but also by factors outside of his/her control. For example, a candidate may have a limited budget such that she is able to bribe only a proportion of voters.

<sup>4</sup> One may question to what extent the behavior observed in such experiments extends to the bribery and corruption in the naturally occurring environment. This is a common concern related to the external validity of laboratory experiment research. Several recent papers have provided evidences supporting the external validity of the findings from lab experiments (Dai, Galeotti and Villeval 2016; Herbst and Mas 2015; Kroll and Rustagi 2016).

Malmendier and Schmidt, 2012). However, in these experiments, the third party often cannot send any signal to the gift recipient. Thus, it is unclear whether the “corrupt” behavior observed in the experiment is due to the intention difference between the third party and the gift giver. A key feature of our experiment is that it minimizes the possibility that people will be treated differently by the gift receiver due to differences in their intentions to give (or to bribe). One goal of our study is to shed light on the role of intention in the “corrupt behavior” observed in the previous experiments.

In our experiment, participants play the role of either a Divider or a Receiver. The experiment consists of two stages. In the first stage, Dividers can decide whether to divide a fixed small amount of money with a Receiver equally or keep all the money and leave nothing to the Receiver. Dividers make this decision *prior to* being matched with a particular Receiver. After all Dividers make their decisions, their decision is randomly assigned to a Receiver and each Receiver receives the amount given by the randomly assigned Divider. In the second stage, each Receiver must decide how to allocate resources between two Dividers: the one whose decision affects her payoff in the first stage and the other whose decision does not affect her payoff in the first stage. The Receiver earns an extra fixed amount of money for completing the allocation decisions in the second stage. Using a strategy method, we focus on the Receiver’s allocation decisions when her randomly assigned (paired) Divider and the other Divider have indicated the same generosity towards her (in the experiment we call the former Divider P and the latter Divider N). Our design ensures that, in this case, the two Dividers have signaled the same intentions and have the same amount of earnings before the Receiver’s allocation decisions.

Our main research question is whether a Receiver, in the second stage, will show favoritism towards her Divider P (*i.e.*, the actual gift giver) at the cost of Divider N (*i.e.*, the third party) who would be just as generous. In addition, we design various conditions to address the following two questions. First, does such favoritism, if any, vary with the cost of social efficiency? To address this question, a Receiver is asked to make allocation decisions under several conditions where every dollar allocated to the gift giver costs the third party different amounts.

Second, does the favoritism towards the gift giver vary depending on whether the Receiver believes a gift giver expects a future return when he decides to give? Answering this question sheds light on whether the favoritism towards gift givers would occur even in altruistic gift giving situations, such as donation, where gifts are given without any intention to “bribe.” To address

this question, we compare two treatments: the *Expected reciprocity opportunity* treatment (henceforth, *Expected*) and the *Unexpected reciprocity opportunity* treatment (henceforth, *Unexpected*). The *Expected* treatment differs from the *Unexpected* treatment in that participants in the latter are not told about the details of the second stage when participating in the first stage (although they do know that the experiment consists of two stages).

We find strong evidence of the pure gift effect. When a Receiver receives a gift from a divider, she allocates more to this divider than to the third party who would have been just as generous to her but did not have the opportunity. Such favoritism towards the gift giver persists in both treatments, even when it is less legitimate because it entails efficiency costs, *i.e.*, the gain to the gift giver is less than the cost to the third party.

Our study provides important behavioral insights on previous experimental studies on corruption. Our results suggest that intention alone cannot explain the “corrupt” behavior observed in the previous experiments. Moreover, favoritism occurs even when the gift giver does not expect any future return (*i.e.* no intention to bribe). In practice, intentions are often viewed as an important determinant of socially harmful behavior, such as corruption.<sup>5</sup> However, our findings suggest that intent to bribe may not be required to trigger “corrupt behavior” that leads to the less socially efficient outcome. Given the difficulty in confirming the intent behind giving, the pure gift effect provides behavioral evidence for the rationale of controlling gifts *per se* when designing institutions to curb corruption<sup>6</sup>. Moreover, our results suggest that the ideal of a “secret donation” or a “secret ballot” (Ayres and Bulow, 1998)<sup>7</sup> should be applied not only to those who give or vote with the expectation of special access and influence, but also to those who do not have such expectations.

---

<sup>5</sup> For example, several pharmaceutical companies, such as GSK and Novartis, define bribery and corruption as “giving, offering or receiving an undue reward with the *intention* of influencing the behavior of someone in government or business to obtain a commercial advantage.” (see <http://www.gsk.com/content/dam/gsk/globals/documents/pdf/AntiCorruption-Booklet.pdf>). The US legal system also

requires proof of intent as a necessary element of the crime of bribery to determine the discharge of official duties regardless of whether the official has accepted the bribe (<http://bribery.uslegal.com/elements-of-offense/intent/>)

<sup>6</sup> Indeed, the UK bribery Act differs from US FCPA (Foreign Corrupt Practices Act) in that the Bribery Act makes no requirement for a “corrupt” or “improper” intent in relation to the bribery of a foreign public official, although the requirement remains for the general bribery offence (<http://www.nortonrosefulbright.com/knowledge/publications/52195/differences-between-the-uk-bribery-act-and-the-us-foreign-corrupt-practices-act>).

<sup>7</sup> Also see (<http://www.law.yale.edu/news/2021.htm>) for an interesting article by Ackerman and Ayres. They argue that it will be more difficult for parties to sell access or influence if we keep political donations secret, because in that case politicians are not able to determine who has given how much.

Our findings provide theoretical insights on modeling reciprocity when multiple parties are involved. Previous theoretical and empirical studies on two-person gift exchanges cannot inform whether the gift receiver's reciprocal behavior is influenced by the gift giver's intention or by the gift she receives. The reason is that reciprocating intentions are always perfectly correlated with reciprocating the gift in two person interactions (Strassmair, 2009; Stanca, 2010). We show that the Receiver's behavior observed in our experiment is inconsistent with the most prominent social preference theories, such as intention-based and outcome-based theories. Differentiating receiving a gift from receiving the intention of giving is important in predicting behavior when more than one party intends to give. Our data suggest that a Receiver distinguishes between the one that benefited her and the one who intended to benefit her but, *by chance*, was not able to do so.

The rest of the paper is organized as follows: Section 2 discusses recent research related to the negative externality of gift giving. Section 3 describes the experiment design. Section 4 presents a revised intention-based model incorporating moral luck. Section 5 presents our main experimental results. Section 6 discusses the findings and section 7 concludes.

## **2. Related Literature on Negative Externality of Gift Giving**

Previous experiments on gift exchange with externality are mainly motivated by the topics on corruption. Abbink et al. (2002) first designed the bribery game, which shows that a receiver reciprocates toward a briber at the cost of third parties. Abbink (2004) further studied how staff rotation influences corruption. Barr and Serra (2010) investigated the connection between culture and corruption. Van Veldhuizen (2012) examined whether higher wages contribute to reducing corruption. Currie et al. (2013) conducted a field experiment and found that gift giving creates positive (negative) externalities for third parties (not) associated with the gift giver. A common feature of these studies is that the third party is always in a passive role, and thus cannot signal any intentions to the receiver. Recent field studies on vote buying do not allow us to know how voters perceive the intentions of the third parties (Hicken et al 2014; Vincente 2014). As a result, we cannot learn from these studies whether the "harmful" reciprocal behavior is simply due to the intention difference between the gift giver and the third party.

Several recent studies have added treatments in which third parties can signal their intentions. Malmendier and Schmidt (2012) conducted an Incentive treatment in which a recipient could receive gifts from both parties. Compared to a treatment in which only the assigned

potential gift giver is allowed to give a gift, the favoritism towards the gift giver is reduced when the third party can also give a gift to the recipient. Gneezy et al. (2013) also found that subjects are less likely to favor one party over another if they take gifts from both. Nonetheless, these studies are unable to separate the intention effect from the pure gift effect due to the fact that the intent to give always results in actual gifts to the recipient. By contrast, our study separates these two effects: both the gift giver and the third party may signal their intention to give, but the recipient only receives a gift from the gift giver.

Studies on conflict of interest also suggest the potential negative externality of gift giving. Previous literature has studied two major types of conflicts of interest. The first type involves conflicts between the professional's *personal* (usually financial) interests and the interests of another party (Loewenstein et al. 2011; Bazerman and Tenbrunsel 2011). Another type of conflict involves competing loyalties. This latter type of conflict may be between one party (client or patient) and another party to whom the professional owes contractual duties. This could include, for example, sponsors of research, insurance companies or employers (Moore et al. 2006). In all these cases, conflict of interest is often defined as “[t]he situation in which a public official or fiduciary who, contrary to the obligation and absolute duty act for the benefit of the public or a designated individual, exploits the relationship for personal benefit, typically pecuniary.”<sup>8</sup> For instance, accounting firms that provide both auditing and consulting services to the same client may be biased in their auditing services due to the fact that an unfavorable auditing report of the client will risk their future earnings on consultant services for the same client (Moore et al. 2006).

In contrast, the gift receiver in our study does not gain any tangible benefit by favoring one party at the cost of another. Thus, strictly speaking, there is no “conflict of interest” in the receiver’s decision problem. Nevertheless, our study suggests that a profitable relationship that occurred in the past can still lead an agent to favor her previous benefactor at the cost of others, even when there is no monetary benefit to doing so. The benefactor’s intention is not the only cause for such favoritism.

### **3. Experiment Design**

Our experiment is built on previous experimental research on corruption and bribery when the gift exchange relationship harms third parties (see Abbink et al., 2002; Malmendier and Schmidt,

---

<sup>8</sup><http://legal-dictionary.thefreedictionary.com/conflict+of+interest>



2012)<sup>9</sup>. Figure 1 outlines the structure of the experimental paradigm. The key feature of our design is minimizing the difference between the gift giver and the third party. In particular, both the gift giver and the third party can signal generosity to the gift receiver. For the receiver, the only difference between the gift giver and the third party is that the gift she receives comes from the gift giver.

The experimental paradigm consists of two stages, with one decision task in each stage. In Stage I (see Figure 1a), Dividers are given \$4 and Receivers are given \$0. Prior to being matched with a particular Receiver, each Divider decides how to divide \$4 between himself and a Receiver. A Divider can either give \$2 to a Receiver and keep \$2 for himself or keep all \$4 to himself. After all Dividers have made their decisions, the computer randomly assigned each Divider's decision to a different Receiver. Next, the computer randomly assigns four participants into a group. Each group has two Dividers and the corresponding two Receivers whose Stage I payoffs are affected by the decision of each of the two Dividers in Stage II (Figure 1b).

Our design of Stage I ensures that within a group, both Dividers can signal intentions (generous or selfish) to each of the two Receivers. For example, if both Dividers choose to split the \$4 equally between them and a Receiver, the Receiver should view the two Dividers as equally kind. In addition, there is no wealth difference between the two Dividers when they make the same division decisions. For example, if both Dividers choose to give \$2 to a Receiver, both will earn \$2 in the first stage (although the \$2 given by the Dividers went to different Receivers). To minimize the perceived difference between the two Dividers, we also use neutral terms to name the two Dividers. In the experiment, for each Receiver, the Divider whose decision is assigned to the Receiver is named as "Divider P" and the other Divider is named as "Divider N." Likewise, for each Divider, the Receiver who was assigned his decision is named as "Receiver P" and the Receiver who was assigned the other Divider's decision is named as "Receiver N". As we explain more in the next paragraph, the grouping mechanism also ensures that from each Receiver's perspective, either Divider P or Divider N is in the identical position in that each

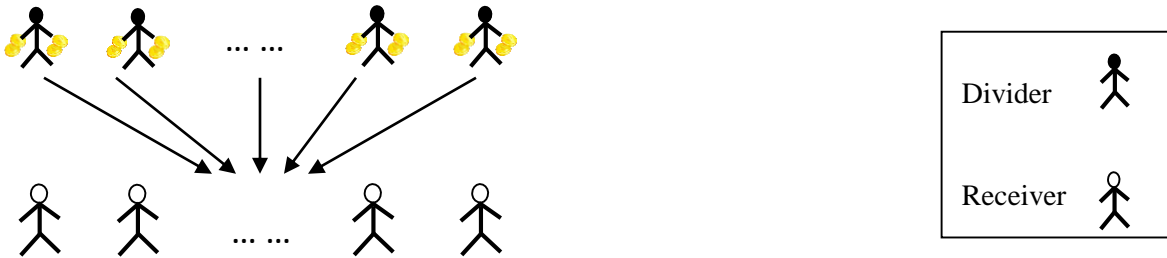
---

<sup>9</sup> To provide clean evidence for the role of intention in reciprocity with negative externality, we describe the experiment to the subjects in context-free neutral language. In particular, we do not use words such as "bribe," "corrupt," "harm," or "inefficient". Abbink and Henning-Schmidt (2006) conducted an experiment to compare the results where a corruption game using context-free wording (similar to our experiment) and another where the context is heavily loaded with negative ethical preconceptions. They did not find significant differences between the two games and conjecture that "*the experimental design transmits the essential features of a bribery situation already with neutral framing*".

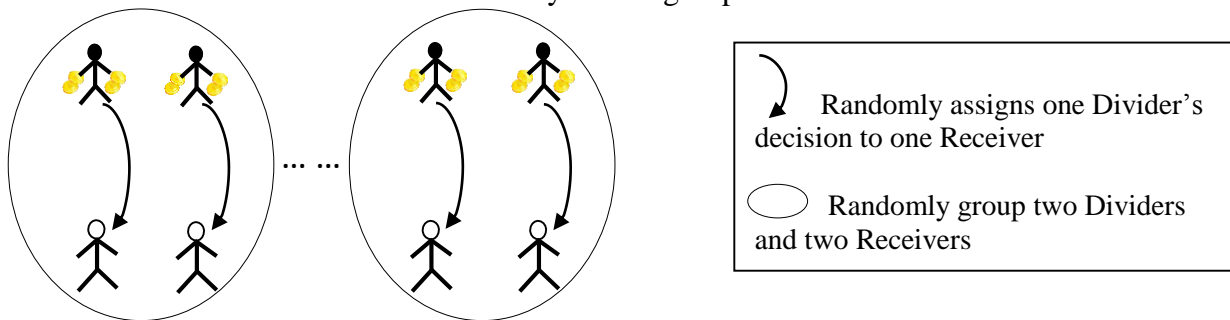
divider's payoffs in Stage II are determined by his Receiver P and Receiver N with an equal chance. Thus, for a particular Receiver, the only difference between the two Dividers is that only one of them is responsible for her earnings in Stage I as a result of random matching.

**Figure 1.** Experiment Paradigm

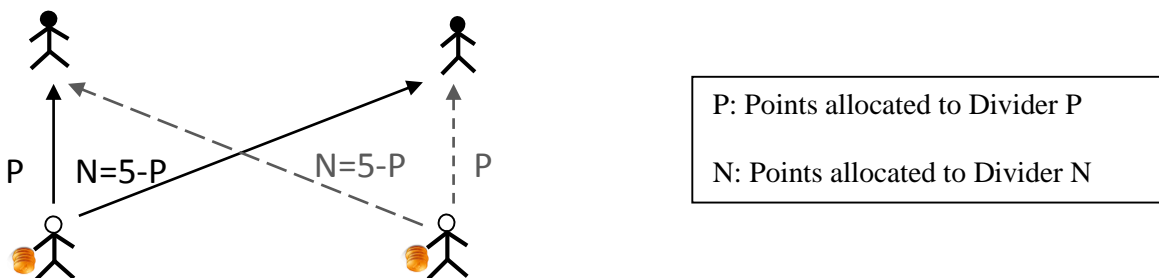
**a. Stage I.** Divider decides whether to give \$0 or \$2 to a Receiver



**b. Matching procedure.** Each Divider's decision is randomly assigned to a Receiver. Two Dividers and two Receivers randomly form a group.



**c. Stage II.** In each group, each Receiver simultaneously and independently decides how to allocate the five points between Divider P and Divider N. The decision of one of the Receivers is randomly selected to determine each Divider's payoff.



In Stage II, a Receiver receives 5 points. The Receiver is required to allocate *all* the points between Divider P and Divider N (Figure 1c). A Receiver’s allocation decisions do not affect her payoffs in Stage II. In particular, a Receiver earns a fixed \$11 in Stage II. We use strategy methods to elicit a Receiver’s allocation decisions in all four possible Stage I scenarios (listed in Table 1)<sup>10</sup>. Scenarios *a* and *c* were the symmetric outcome scenarios, where both Dividers had signaled identical intentions to the Receiver. We were particularly interested in Receivers’ decisions in scenario *a*, where both Dividers had signaled generosity to the Receiver. Therefore, in the strategy method we always presented scenario *a* first and randomized the order of scenarios *b*, *c* and *d*.

**Table 1:** Decision Scenarios for Receivers

- Four possible outcomes in Stage I

a. Your Divider P chooses Option A (\$2, \$2); Your Divider N chooses Option A (\$2, \$2).
b. Your Divider P chooses Option A (\$2, \$2); Your Divider N chooses Option B (\$4, \$0).
c. Your Divider P chooses Option B (\$4, \$0); Your Divider N chooses Option B (\$4, \$0)
d. Your Divider P chooses Option B (\$4, \$0); Your Divider N chooses Option A (\$2, \$2).

- Three Valuation Rules

Valuation rules	Divider P	Divider N
Rule 1	1 point worth \$1	1 point worth \$1
Rule 2	1 point worth \$1	1 point worth \$2
Rule 3	1 point worth \$2	1 point worth \$1

In addition, for each scenario, there were three possible valuation rules and each rule specified the dollar value of the point (either \$1 or \$2) to Divider P and Divider N (see Table 1). Participants were told that at the end of the experiment, the decision of one of the two Receivers

---

<sup>10</sup> Strategy method allows us to collect more data and conduct within subject comparison. The main disadvantage of the strategy method is that it may diminish the effect of emotions. However, if we assume that the positive emotion towards the actual gift giver is greater than the emotion towards the intended one, the strategy method would work against our hypothesis.

under one of the three valuation rules would be randomly selected. The selected Receiver's decision for the selected valuation rule and for the actual Stage I outcome would determine the two Dividers' payoffs in Stage II (see Appendix E for instructions and screenshots). Note that as only one (of the two) Receiver's decision will be applied to determine the two Dividers' earnings, this design minimizes the possibility that a Receiver's allocation decision may be affected by her beliefs regarding the other Receiver's allocation decision<sup>11</sup>.

As shown in Table 1, under Rule 1, each point is worth the same amount to the two Dividers. Thus, Rule 1 is efficiency neutral. We use Rule 1 to identify whether a Receiver favors one Divider at the cost of the other when there was no efficiency loss. The degree of favoritism can be measured by  $\Delta_{neutral} = (\$_{rule1}^P - \$_{rule1}^N)$  where  $\$_{rule1}^P$  is the earnings a Receiver assigned to Divider P and  $\$_{rule1}^N$  is the earnings a Receiver assigned to Divider N under Rule 1.

Under Rule 2, one point to Divider P is worth less than one point to Divider N. The opposite is true for Rule 3. That is, it is either inefficient to assign points to Divider P under Rule 2, or to assign points to Divider N under rule 3. Thus, by comparing the amount of earnings allocated to Divider P ( $\$_{rule2}^P$ ) under Rule 2 and to Divider N under Rule 3 ( $\$_{rule3}^N$ ), we are able to measure the degree of favoritism when favoritism not only harms the other party but also leads to efficiency loss:  $\Delta_{costly} = (\$_{rule2}^P - \$_{rule3}^N)$ . Assuming the less efficient the favoritism is, the less legitimate the favoritism is. The three valuation rules allow us to learn whether a Receiver showed favoritism toward Divider P at the cost of Divider N, and, if so, whether the favoritism persisted even when it was less legitimate in that it led to efficiency loss.

### *Treatments*

We conducted two treatments: *Expected* treatment and *Unexpected* treatment. The only difference between the two treatments was that in the *Expected* treatment, participants were given all the detailed information about the two stages at the beginning. By contrast, in the *Unexpected*

---

<sup>11</sup> One may argue that this design may still lead receivers to consider what the other receiver would do and thereby affect their behavior. First, without further assumptions, a receiver can form any type of beliefs regarding what the other receiver would do and there could be multiple equilibria. Our hypothesis of "pure gift effect" predicts a certain type of belief (a receiver should/would give more to her Divider P than Divider N) and thus selects certain equilibrium over others. Therefore, to argue that there might be a belief of favoring the actual gift giver is consistent with our hypothesis of "pure gift effect". Second, we asked receivers in a survey after the experiment whether they thought about what the other receiver might do when they were making the decisions. More than 70% receivers answered "No". Among those who answered "Yes", only 2 wrote that such a consideration affected their decisions.

treatment, participants were told that the session consisted of two experiments, but did not know what the second experiment was about until they had made their decisions in the first experiment. Note that we described the two parts as “Stage I” and “Stage II” in the *Expected* treatment, but use “Experiment I” and “Experiment II” to refer to these two parts in the *Unexpected* treatment. This is to better separate the two parts of the *Unexpected* treatment and minimize the possibility that Dividers may act strategically in the first part. When participating in Experiment I, all they knew was that their earnings in Experiment I would not be affected by Experiment II. This was common knowledge. We informed the subjects of the existence (although not the details) of the second stage to avoid concerns about deception, as Dividers’ decisions in the first stage could affect their earnings in the second one. Thus, *Unexpected*, like *Expected*, controls for intention differences between the gift giver and the third party, but further minimizes the intention of “bribery” of both the gift giver and the third party.

### *Survey*

After players finished making their decisions, and before seeing the outcome, they were also incentivized to predict how many points a Receiver would give to her Divider P and Divider N in each scenario under each valuation rule. Participants were paid \$1 for completing the survey. They were also told two questions would be randomly selected at the end of the experiment, and they would receive \$1 for each correct answer.

At the end of the experiment, we also asked Receivers why they thought Dividers might give \$2 to a Receiver and why they thought Dividers might keep all of the \$4. The answer to these questions allowed us to determine whether Receivers indeed perceived Dividers’ intentions differently between the *Unexpected* and *Expected* treatments.

### *Experiment Procedure*

We conducted the experiment at Pittsburgh Experimental Economics Lab using z-tree (Fischbacher 2007). 124 subjects participated in our experiment, with 45% (56 subjects) men and 54% (68 subjects) women. The majority of the participants were Caucasian (74%), and the second and third largest groups were Asian (15%) and Black or African American (9%), respectively. Upon arriving, subjects were seated in a booth separated from each other so that they could not communicate or see other subjects during their decision making process. Before making decisions,

subjects were randomly assigned an ID number, which determined their roles. All decisions were made privately and anonymously through the computer interface z-tree. Each Receiver was asked to make multiple decisions in each of the four scenarios. To ensure receivers paid attention to each scenario, we always first presented the specifics of a scenario on the screen for 20 seconds and subjects could not skip the screen until the time was up. Then receivers made the allocation decisions under each of the three valuation rules with the outcome scenario remaining on top of the decision screen. After finishing all the decisions, receivers saw the calculated dollar value of their allocation decisions and were asked to confirm the decisions (see screenshots at the end of Appendix E). All sessions were finished within an hour. Subjects earned an average of \$16, including a \$5 show-up fee.

#### **4. Theoretic framework**

We show in Appendix A (A.1 – A.3) that the existing prominent theories on social preference, such as outcome-based and intention-based models, will not predict the pure gift effect: no favoritism towards the Divider P when both dividers chose to give \$2. Here we propose a theoretical framework to formalize the hypothesis of pure gift effect and illustrate the role of moral luck in gift exchange with externality. Our framework modifies Dufwenberg and Kirchsteiger (2004)'s intention-based model (henceforth, D&K 2004) by incorporating the concept of moral luck. Nagel (1979) defines moral luck as follows: “[w]here a significant aspect of what someone does depends on factors beyond his control, yet we continue to treat him in that respect as an object of moral judgment, it can be called moral luck” (also see Williams, 1981; Scanlon 2008; Nelkin 2013).<sup>12</sup> The theory of moral luck suggests that intention may not be

---

<sup>12</sup> An alternative relevant theory is the social ties model. While van Dijk and van Winden (1997) applied this model to public goods environments, Malmendier and Schmidt (2012) are closer to our study, in that they apply the model to explain corruption behavior in a variation of gift exchange games. They assume that initially, a decision maker is equally concerned about the welfare of all players. But once the gift is given, the gift creates a positive affective bond due to the fact that the gift giver chooses a strategy that gives the recipient a higher payoff than the expected payoff. Applying their model to our experiment, the question of how much a receiver values a divider's payoff would be determined by the “strategy chosen” by the divider. In particular, a receiver weighs the divider's payoff more into her utility function when the divider chooses to give \$2 than when the divider chooses to give \$0. Thus, if both dividers choose to give \$2, their model would predict that the receiver weighs the two dividers' payoff equally. However, if receiving the actual gift strengthens the social ties between the giver and the recipient to a greater degree than receiving the intentions of giving, the recipient will weigh the welfare of the actual giver more than that of the intended giver. In this case, the social ties model will also predict the pure gift effect.

sufficient to predict decisions in a world where outcomes are often determined not only by intentions, but also by elements that cannot be controlled. We demonstrate below that when we take into account the occurrence of moral luck in personal interactions, we generate different predictions of Receiver's behavior than traditional social preference models.

Consider our environment where we have Receivers and Dividers. Our focus is the Receivers' decisions in Stage II. Assume a Receiver's (denoted as R) utility contains two parts: her material payoff  $m^R$ ; and her aggregate reciprocal payoff with respect to each of the two Dividers (P and N).

Receiver's material payoff consists of the earnings from Stage I and Stage II in our environment. A Receiver earns a fixed payoff of \$11 in Stage II irrespective of her decisions or the Dividers' decisions. When a Divider P chooses to give \$2 in Stage I,  $m^R = \$13$ ; otherwise,  $m^R = \$11$ .

As in D&K's model, the Receiver's reciprocal payoff with respect to each Divider is a product of her interpretation of that Divider's kindness to her and her kindness to that Divider. Following D&K, we use  $\lambda_{RjR_{j \in \{P,N\}}}$  to represent the Receiver's perceived intention of Divider  $j$  toward her. In our environment,  $\lambda_{RjR_{j \in \{P,N\}}}$  is determined only by the difference between the payoff the Receiver received from Divider  $j$  (\$0 or \$2) and the equitable payoff she deserves. The equitable payoff of a Receiver can be specified as the average payoff she could have received given the strategy space of the Divider. Given that a Divider can either choose to give \$2 or \$0, the equitable payoff of a Receiver is \$1. Thus,  $\lambda_{RjR_{j \in \{P,N\}}} = 1$  or  $-1$ .

We refer to a Receiver's response to a Divider's intention specified in D&K (2004) as *pure intention effect*. We now consider that a Divider's good or bad intention does not necessarily result in *actual* benefit or harm to the Receiver due to random factors out of the Divider's control. Thus, it is important to distinguish between *pure intention* and *received* intention. As specified in Equation (1), we hypothesize that the Receiver reciprocates to the *received* intention (kindness or unkindness) rather than the perceived pure intention.

$$\text{Received intention} = f(w_1 * \lambda_{RjR}, w_2 * l_{Rj})_{j \in \{P,N\}} \quad (1)$$

where  $w_1 \in [0,1]$  is the weight assigned to  $\lambda_{RjR_{j \in \{P,N\}}}$  and  $w_2 \in [0,1]$  is the weight assigned to  $l_{Rj_{j \in \{P,N\}}}$ .  $l_{Rj_{j \in \{P,N\}}} \in [-1,1]$  is the proportion of realized intention of Divider  $j$ . When  $0 < l_{Rj_{j \in \{P,N\}}} \leq 1$ , this means that the Receiver receives all or some of the benefit Divider  $j$  intended

to send. When  $-1 \leq l_{Rj_{j \in \{P,N\}}} < 0$ , this means the Receiver receives all or some of the harms Divider  $j$  intended to impose. When  $l_{Rj_{j \in \{P,N\}}} = 0$ , this means the Divider  $j$ 's kindness/unkindness does not result in any actual benefit/harm to the Receiver. In our experiment,  $l_{RN} = 0$  regardless of Divider N's decisions as his decision never benefit or harm a Receiver.  $l_{RP} = 1$  when Divider P gives \$2 in Stage I and  $l_{RP} = -1$  when Divider P gives \$0. We assume that  $f(\cdot)$  is a concave, monotonically increasing function, with  $f(0,0) = 0$ .

Similar to D&K, we denote Receiver's kindness to Divider  $P$  (or  $N$ ) as  $g(k_{Rj})_{j \in \{P,N\}}$ <sup>13</sup>, with  $k_{Rj_{j \in \{P,N\}}} = m^j - e_j^R$ . Here,  $m^j$  is the earnings a Receiver assigns to Divider  $j$  and  $e_j^R$  is the Divider  $j$ 's equitable payoff given R choosing efficient strategies. Under Rule 1,  $e_{P,rule1}^R = e_{N,rule1}^R = \$2.5$ . Under Rule 2 or 3, the Divider who earned \$2 for each received point has an equitable payoff of \$5, while the Divider who earned \$1 for each received point has an equitable payoff of \$2.5. We assume that  $g(\cdot)$  is a concave, monotonically increasing function<sup>14</sup>.

In our revised intention-based model with moral luck, a Receiver's solves

$$\max_{m^j} U^R = m^R + \sum_{j \in \{P,N\}} f(w_1 * \lambda_{RjR}, w_2 * l_{Rj}). g(m^j - e_j^R) \quad (2)$$

$$\text{Rule 1: } s.t. m^P + m^N = 5$$

$$\text{Rule 2: } s.t. m^P + 0.5m^N = 5$$

$$\text{Rule 3: } s.t. 0.5m^P + m^N = 5$$

Our main interest is Receiver's allocation decision in the two scenarios when both Dividers have made the same decision in Stage I: 1) when both Dividers are kind,  $\lambda_{RjR} = 1$ ;  $l_{RP} = 1$ ;  $l_{RN} = 0$ ; and 2) When both Dividers are unkind,  $\lambda_{RjR} = -1$ ;  $l_{RP} = -1$ ;  $l_{RN} = 0$ . We derive the following predictions of these two scenarios (the predictions in the other two scenarios are detailed in Appendix A.4). To help differentiate the assigned dollar amount from the amount of points, below we use "\$" to denote the total dollar amount assigned to each Divider.

<sup>13</sup> One difference between our model and D&K (2004) is that we assume a Receiver's kindness to a Divider follows a concave rather than a linear function. The reason is that D&K (2004) applies to environments with mostly dummy decision variables, while a Receiver's decision in our environment can be any integer number between [0,5]. Therefore, it makes more sense to have a concave function for utility's components.

<sup>14</sup> Previous social preference models usually assume a linear utility function. Our prediction of favoritism would not change whether the utility function is concave or linear.



**Propositions**<sup>15</sup>: As long as  $w_2 > 0$ ,

*Under Rule 1*

When both Dividers have decided to give \$2,  $\$_{rule1}^N < 2.5 < \$_{rule1}^P$ .  $\$_{rule1}^P$  is increasing in  $w_2$  and is decreasing in  $w_1$ .

When both Dividers choose not to give,  $\$_{rule1}^P < 2.5 < \$_{rule1}^N$ .  $\$_{rule1}^P$  is decreasing in  $w_2$  and is increasing in  $w_1$ .

*Under Rule 2 and Rule 3:*

When both Dividers have decided to give \$2,  $\$_{rule2}^P > \$_{rule3}^N$ .  $\$_{rule2}^P$  is increasing in  $w_2$  and is decreasing in  $w_1$  and  $\$_{rule3}^N$  is decreasing in  $w_2$  and increasing in  $w_1$ .

When both Dividers choose not to give,  $\$_{rule2}^P < \$_{rule3}^N$ .  $\$_{rule2}^P$  is decreasing in  $w_2$  and increasing in  $w_1$  and  $\$_{rule3}^N$  is increasing in  $w_2$  and decreasing in  $w_1$ .

**Proof:** See appendix A.4 for details.

Our framework predicts favoritism toward Divider P both when it is efficiency neutral and when it leads to efficiency loss ( $\Delta_{neutral} > 0$  and  $\Delta_{costly} > 0$ ), as long as  $w_2 > 0$ . Our general framework does not produce a unique prediction on how the degree of favoritism may differ when it is efficiency costly than when it is efficiency neutral in our setting. Whether efficiency cost mitigates the favoritism depends on: the specifications of  $g$  function; how much a Receiver cares about pure intention ( $w_1$ ); and how much she cares about to what extent the intention *actually* benefits or harms her ( $w_2$ ). In Appendix A.4, we provide some examples to illustrate the comparisons of favoritism under different efficiency cost conditions.

Our model also predicts that favoritism appears in both *Unexpected* and *Expected* treatments and its magnitude does not vary between these two treatments.

## 5. Results

---

<sup>15</sup> Here we provide general predictions assuming a continuous choice space. In the experiment, subjects were allowed to allocate only integer points. In Section 6.2, we report data from another robustness check treatment where decimal points are allowed.

We obtained 64 participants for the *Expected* treatment and 60 participants for the *Unexpected* treatment. Given that we were interested in Receivers' allocation decisions when both Dividers had identical intentions, we focused on Receivers' decisions when both Dividers gave \$2 (or \$0). We first report the decisions under Rule 1 and then those under Rule 2 and Rule 3. We also report Dividers' decisions in Stage I at the end of the Results section. In Appendix B, we report Receivers' allocation decisions when the two Dividers made different decisions in Stage I.

### 5.1. Receiver's decision when both Dividers decided to give \$2

In both treatments, under Rule 1, we observe that on average Receivers gave significantly more to Divider P than to Divider N, even though Divider N was just as generous as Divider P (*Expected*: 3.4 vs. 1.6; *Unexpected* 3.1 vs. 1.9;  $p < 0.01$ , two-sided Wilcoxon signed-rank test<sup>16</sup>). We can calculate the degree of favoritism for each Receiver  $\Delta_{neutral} = \$_{rule1}^P - \$_{rule1}^N$ , which is the difference between the dollar amount she assigned to Divider P and to Divider N under Rule 1. We find no significant difference in the degree of favoritism ( $\Delta_{neutral}$ ) between the two treatments (1.8 in *Expected* vs. 1.3 in *Unexpected*,  $p = 0.11$ , two-sided Mann-Whitney test).

In Figure 2, we report the distribution of  $\Delta_{neutral}$ . Figure 2 shows that for both the *Expected* and *Unexpected* treatments, there is a higher proportion of positive  $\Delta_{neutral}$  than negative ones, which indicates favoritism towards Divider P. Such differences are statistically significant (*Expected*: 97% vs. 3%; *Unexpected*: 87% vs. 13%,  $p < 0.01$ , two-sided M-W test). This result supports our predictions and indicates that most Receivers care to what extent they actually benefited from the intention rather than pure intention alone. The distribution data also suggests that in both treatments there is a non-trivial proportion of Receivers (about 20% in *Unexpected* treatment and 34% in *Expected* treatment) showing a high degree of favoritism toward Divider P ( $\Delta_{neutral} \geq 3$ ).

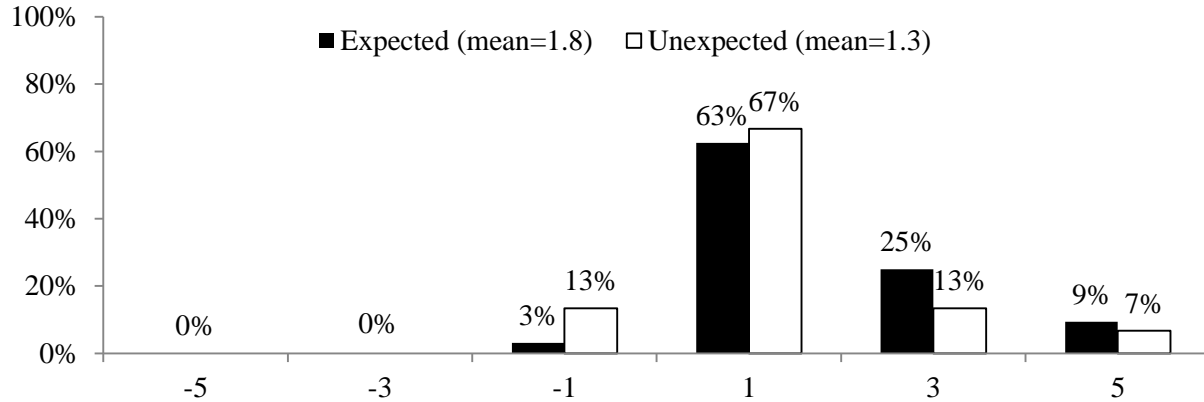
Next, we compare the dollar amount Divider P received under Rule 2 and the dollar amount Divider N received under Rule 3 to determine whether the favoritism towards Divider P persists when it leads to less efficient outcomes. We continue to find that Divider P received significantly more under Rule 2 than Divider N under Rule 3 (3.7 vs. 2.0,  $p < 0.01$ ) in the *Expected* treatment. The same is true for the *Unexpected* treatment (3.4 vs. 2.1,  $p < 0.01$ ). Again,

---

<sup>16</sup> Henceforth, unless otherwise notified, p-values are reported using a two-sided signed-rank Wilcoxon test.

for each Receiver we can calculate the degree of favoritism as  $\Delta_{costly} = \$_{rule 2}^P - \$_{rule 3}^N$ , which is the difference between the dollar amount a Receiver assigned to Divider P under Rule 2 and the amount assigned to Divider N under Rule 3. Similar to the findings under Rule 1 ( $\Delta_{neutral}$ ), we find  $\Delta_{costly}$  does not differ between the two treatments (1.7 vs. 1.3,  $p = 0.46$ , two-sided M-W test).

**Figure 2:** Distribution of  $\Delta_{neutral} (\$_{rule 1}^P - \$_{rule 1}^N)$  when both Dividers give \$2.

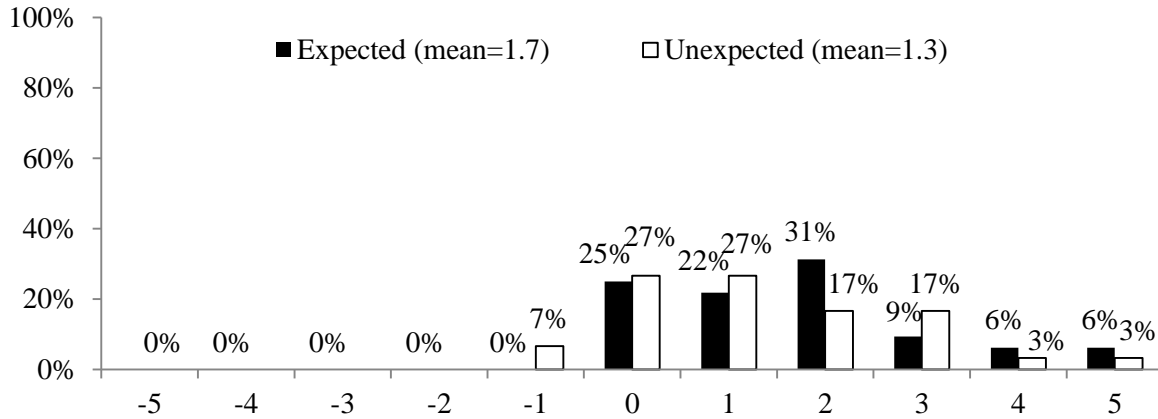


Note: There are two 0% bars underlying  $\Delta_{neutral} = -5$  and  $-3$ ; for simplicity, we only note one 0%. We apply the same notation for the following figures.

Figure 3 reports the distribution of favoritism ( $\Delta_{costly}$ ) in both treatments. Supporting our predictions, Figure 3 shows that there is a significantly higher proportion of positive  $\Delta_{costly}$  than negative ones (75% vs. 7% in *Expected*,  $p < 0.01$ ; 66% vs. 0% in *Unexpected*,  $p < 0.01$ , two-sided M-W test). Similar to Rule 1, we again observe that in both treatments, about 20% of Receivers display a high degree of favoritism toward Divider P ( $\Delta_{costly} \geq 3$ ).

Comparing  $\Delta_{neutral}$  and  $\Delta_{costly}$  in each treatment, our data show that the degree of favoritism does not vary with its consequences on efficiency in either the *Expected* or *Unexpected* treatments (1.8 vs. 1.7 in *Expected* treatment,  $p = 0.25$ ; 1.3 vs. 1.3 in *Unexpected* treatment,  $p = 0.77$ ).

**Figure 3:** Distribution of  $\Delta_{costly} = (\$_{rule 2}^P - \$_{rule 3}^N)$  when both Dividers give \$2.

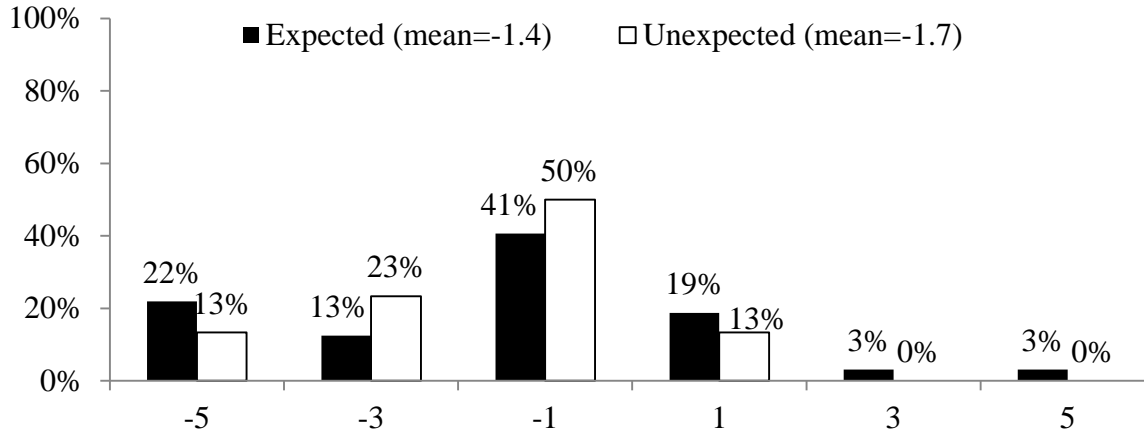


### 5.2. When both Dividers decide not to give

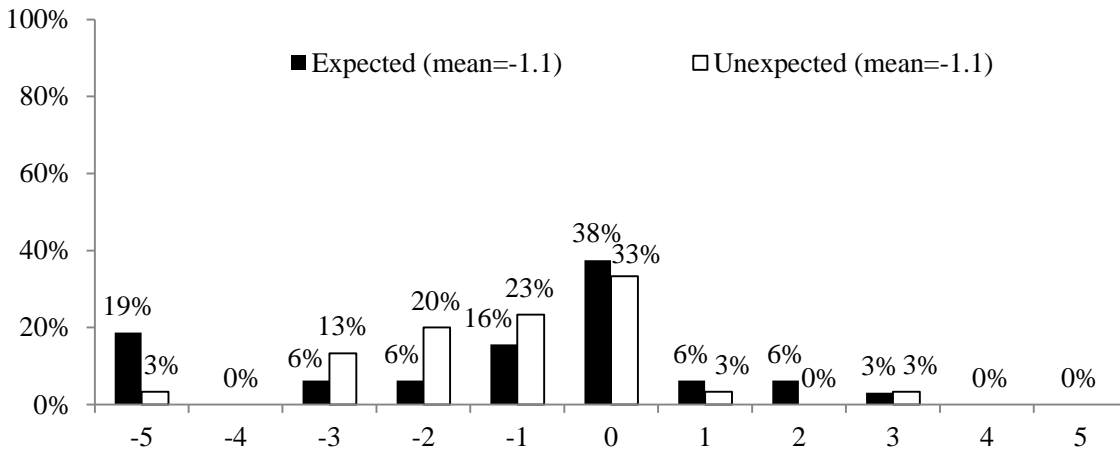
Our data suggest that a Receiver not only cares about whose kindness actually benefited her, but also whose selfish intention essentially harmed her. In both the *Expected* and *Unexpected* treatments, when both Dividers give \$0, we observe that a Receiver gives significantly less to Divider P than to Divider N under Rule 1 (*Expected*: \$1.8 vs. \$3.2,  $p < 0.01$ ; *Unexpected*: \$1.6 vs. \$3.4,  $p < 0.01$ ). We find no significant difference in the degree of bias against Divider P between the two treatments ( $\Delta_{neutral}$ : -1.4 in *Expected* vs. -1.7 in *Unexpected*;  $p = 0.63$ ). Figure 4 reports the distribution of  $\Delta_{neutral}$  in each treatment. Supporting our predictions, we observe a significantly higher proportion of negative  $\Delta_{neutral}$  than position ones (*Expected*: 75% vs. 25%; *Unexpected*: 86% vs. 13%,  $p < 0.01$ , two-sided M-W test).

We further examine what happens when bias against Divider P also leads to efficiency loss. We continue to find that Divider P receives significantly less earnings under Rule 2 than Divider N under Rule 3 (*Expected*: 2.3 vs. 3.4,  $p < 0.05$ ; *Unexpected*: 2.3 vs. 3.4,  $p < 0.01$ ). Like with Rule 1, we observe no difference in the degree of bias against Divider P between the two treatments ( $\Delta_{costly}$ : -1.1 in *Expected* vs. -1.1 in *Unexpected*,  $p = 0.58$ , two-sided M-W test). Figure 5 further reports the distribution of  $\Delta_{costly}$ . Supporting the predictions, the proportion of negative  $\Delta_{costly}$  is significantly higher than positive ones (*Expected*: 47% vs. 15%,  $p < 0.01$ ; *Unexpected*: 59% vs. 15%,  $p < 0.01$ , two-sided M-W test).

**Figure 4:** Distribution of  $\Delta_{neutral}$  ( $\$_{rule1}^P - \$_{rule1}^N$ ) when both Dividers give \$0.



**Figure 5:** Distribution of  $\Delta_{costly}$  ( $\$_{rule2}^P - \$_{rule3}^N$ ) when both Dividers give \$0.



Similar to the case when both Dividers give \$2, we find no significant difference between  $\Delta_{neutral}$  and  $\Delta_{costly}$  (*Expected*: -1.4 vs. -1.1,  $p=0.18$ ; *Unexpected*: -1.7 vs. -1.1,  $p=0.06$ ). Our data also show that overall a majority of those receivers who favor Divider P when both dividers give \$2 also display the negative bias toward Divider P when both give \$0. First, we examine the case under Rule 1 ( $\Delta_{neutral}$ ). In the *Expected* treatment, among the 31 receivers who display favoritism toward Divider P ( $\Delta_{neutral} > 0$ ), when both dividers give \$2, 23 (74%) of them also hold a negative bias toward Divider P ( $\Delta_{neutral} < 0$ ) when both give \$0. This percentage is 84% (22 out of 26) in the *Unexpected* treatment. Next, we examine the case when the bias leads to efficiency losses ( $\Delta_{costly}$ ). There are 24 receivers in the *Expected* and 20 in the *Unexpected*

treatment who favor Divider P ( $\Delta_{costly} > 0$ ) when both dividers give \$2. Among these receivers, 46% of them (11 out of 24) in the *Expected* and 60% (12 out of 20) in the *Unexpected* also hold a negative bias toward Divider P ( $\Delta_{costly} < 0$ ); 33% of them (8 out of 24) in the *Expected* and 35% (7 out of 20) in the *Unexpected* treat the two dividers equally ( $\Delta_{costly} = 0$ ) when both give \$0.

### 5.3. Dividers' behavior and belief

We find that a significantly greater proportion of Dividers gave \$2 in the *Expected* than in the *Unexpected* treatment (31% vs. 6.7%,  $p < 0.05$ , two-sided M-W test), suggesting that Dividers expect their Receiver to allocate more to them if they act generously. As mentioned in Section 3, we elicited Dividers' beliefs regarding Receivers' allocation decisions in Stage II for each scenario of Stage I. Again, here we focus on the case when both Dividers make the same decisions in Stage I. As we report below, the survey data suggests that Dividers expect positive bias toward Divider P when both Dividers are generous, but fail to predict negative bias against Divider P when both are selfish. Again, we observe no difference in the belief data between the *Expected* and *Unexpected* treatments.

For the scenario in which both Dividers chose to give \$2 in Stage I, on average, Dividers expected Receivers to give significantly more to Divider P who chose to equally split the \$4 than to Divider N, even though Divider N was just as generous as Divider P under Rule 1 (*Expected*: 2.9 vs. 2.1; *Unexpected*: 3.0 vs. 2.0;  $p < 0.01$ ). Dividers also correctly predicted that the favoritism would persist even when it led to efficiency losses: Dividers expected that Divider P would receive significantly more earnings under Rule 2 than Divider N under Rule 3 (*Expected*: 3.0 vs. 2.2,  $p < 0.01$ ; *Unexpected*: 3.2 vs. 2.2,  $p < 0.01$ ).

Interestingly, Dividers do not seem to expect significant negative bias against Divider P in the scenario where both Dividers act selfishly. Under Rule 1, on average, Dividers expected no significant differences in the amount allocated to Divider P and to Divider N (*Expected*: 2.3 vs. 2.8,  $p = 0.41$ ; *Unexpected*: 2.4 vs. 2.6,  $p < 0.83$ ). Similarly, Dividers did not expect significant difference in the allocation amount to Divider P under Rule 2 and to Divider N under Rule 3 (*Expected*: 2.3 vs. 2.7,  $p = 0.98$ ; *Unexpected*: 2.4 vs. 2.6,  $p = 0.78$ ).

We also calculated Dividers' belief of  $\Delta_{neutral}$  and  $\Delta_{costly}$  to measure the degree of Dividers' expected bias and provided statistic comparisons between the two treatments (see Table

4). As shown in Table 4, there is no difference between the *Expected* and *Unexpected* treatment in any cases.

**Table 4:** Dividers’ belief

Stage I	$\Delta_{neutral}$			$\Delta_{costly}$		
	Expected	Unexpected	M-W test	Expected	Unexpected	M-W test
Both Dividers gave \$2	0.8	1.1	P = 0.44	0.8	1.0	P = 0.39
Both Dividers gave \$0	-0.5	-0.2	P = 0.67	-0.2	-0.3	P = 0.87

## 6. Discussions

### 6.1 Expected vs. Unexpected

We provide experimental evidence supporting the pure gift hypothesis. Receivers treat the dividers differently depending on whether the dividers’ kindness actually affected the receivers’ payoffs. Interestingly, we do not observe significant differences between the *Expected* and *Unexpected* treatments. This result is in keeping with some of the previous literature on related topics. For example, Strassmair (2009) finds that in a modified trust game, the trustee’s behavior is not affected by the investor’s expected probability of being reciprocated. One implication of our finding is that pure gift effect occurs regardless of whether the gift giver is purely altruistic or acts kindly for future returns.

Yet, one may argue that our manipulation of the *Expected* and *Unexpected* treatments may not be effective in that it does not lead receivers to interpret the giving behavior differently. To examine this possibility, we analyzed the receiver’s belief of the intention of the gift. In the survey at the end of the experiment, we asked each Receiver to answer two questions: 1) “Please tell us what you think why some Dividers might choose Option A (to give \$2 to a Receiver)?”; and 2) “Please tell us what you think why some Dividers might choose to give \$0.” Receivers’ answers provided insights on whether they interpreted Dividers’ behavior differently in the *Expected* and *Unexpected* treatments.

We note that when Dividers choose to give \$2, their intention can be interpreted as either generous or selfish depending on whether there is a possibility of future returns. In contrast, it is natural to assume that such contingent interpretations may not be applied to “stingy” behavior (giving \$0). Indeed, we find that most Receivers, in their answers to the second question, mentioned things like “greedy,” “want money,” “don’t trust receivers,” “don’t believe that Receivers will give more,” “don’t see long term benefit,” etc. All of these indicate perception of selfish motives. It is therefore unsurprising that we do not observe a significant difference in Receivers’ behavior between the two treatments when both Dividers give \$0. Our analysis reported below focuses on the answers to the first question.

To provide a formal analysis, we applied a method developed in Houser and Xiao (2011). Specifically, for each set of the answer messages we obtained in each treatment, we recruited another 5 subjects to code the answers. Each coder was instructed to classify every message into three categories: Self-interest; Nice; or Others. For each message, coders could choose more than one category.<sup>17</sup> Each coder earned \$10 (including a \$5 show-up bonus) for completing coding. At the end of the session, two messages were randomly chosen. For each of the two messages, if their code matched with the most common codes by other coders, they would receive another \$5.<sup>18</sup>

The inter-rater reliability is quite satisfactory, with the combined *kappa* statistics at  $k = 0.83$ ,  $z = 18.7$  in the *Unexpected* treatment,  $k = 0.77$ ,  $z = 11.8$  for the *Expected* treatment.<sup>19</sup> In particular, 71.67% (43 messages) reached 100% agreement among the 5 coders, while 26.67% (16 messages) reached 80% agreement, and only one message had an agreement rate of 60%.

The coding results suggest a difference in Receivers’ perceived intentions of Dividers between the two treatments. In particular, those who believed that a Divider gives because he is nice increased significantly from 12.9% in the *Expected* treatment to 51.7% in the *Unexpected* treatment (two-sided M-W test,  $p < 0.01$ ). On the other hand, those who believed a Divider gives for self-interest decreased significantly from 83.9% in the *Expected* treatment to 37.9% in the

---

<sup>17</sup> When coding a message solely in the “Other” category (8 out of 270 codes), coders were also asked to add a comment. 6 (out of the 8) codes are related to “Self-interest” (e.g. desire, strategizing). We treat these same as coding “Self-interest.” 2 other comments are related to “confused”.

<sup>18</sup> When a coder classified a message into more than one category, we pick the one that receives most agreement among the other four coders.

<sup>19</sup> One of the coders was dropped due to 3 missing observations (where he chose a category not included in the instruction but did not provide his personal explanation). Including him will not change the significance,  $k = 0.62$ ,  $z = 13.2$  (assuming that is a category different from other coders).



*Unexpected* treatment (two-sided M-W test,  $p < 0.01$ ). The percentage of those who believed that a Divider’s giving might be due to a mixed motive of both was low in both *Expected* treatment (3.2%) and *Unexpected* (10.3%, two-sided M-W test,  $p = 0.29$ ). These results exclude the possibility that the non-significant treatment difference is due to the inefficient treatment manipulation.

We also examined whether there is any significant behavioral difference between those who believe a Divider gives out of “nice” motives and those who believe a Divider gives for “self-interest” reasons. We group Receivers according to their interpretations of Dividers’ intentions and calculate the degree of favoritism ( $\Delta_{neutral}$  and  $\Delta_{costly}$ ) within each coding category. To ensure we have a sufficient number of observations in each category, we pool the two treatments together. We find that on average,  $\Delta_{neutral}$  does not differ between the two categories (“nice” vs. “self-interest”) (1.63 vs. 1.54,  $p = 0.85$ , two-sided M-W test). The same result holds when favoring Divider P leads to efficiency loss  $\Delta_{costly}$ : 1.63 vs. 1.56,  $p = 0.86$ , two-sided M-W test). This is consistent with the result that the degree of favoritism is independent of whether the gift giver is purely altruistic or acts kindly for future returns.

One explanation for the result of non-significant treatment effect is that in our experiment, the receivers’ decision is to allocate points between Divider P and Divider N, rather than to each divider separately. Thus, the decisions should be determined by the *differences* between the two Dividers. Our design ensures that when both dividers give \$2, the receiver should interpret the intentions of the two dividers in the same way, be it kindness (in the *Unexpected* treatment) or selfish (in the *Expected* treatment). Regardless of the treatment, from the perspective of the receiver, the difference between the two Dividers is only from whom she actually receives the gift. Our results suggest that favoritism can be driven by reciprocity to the actual gift regardless of the underlying intention of the gift. This result is also consistent with the prediction of our model in Section 4.

## **6.2. Robustness checks**

In the experiment we reported above, receivers can only allocate an integer number of points. Thus, under Rule 1, receivers have to choose one Divider to favor<sup>20</sup>. This feature of the experiment is similar to some situations in naturally occurring environments, such as the political campaign where voters can vote only one candidate. However, in many other cases, a gift recipient may have the opportunity to treat both the actual gift giver and the intended gift giver equally. For example, a government official may have multiple contracts to allocate among the (actual/intended) gift givers. It is important to examine whether pure gift effect would still present when the receivers have the option to make an equal allocation between the two dividers. We thus conducted another experiment with the only change being that we allowed Receivers to assign decimal points in both the *Expected* and the *Unexpected* treatments. Thus, in this new experiment, an equal split was possible under Rule 1. We recruited another 124 participants from the same subject pool (32 pairs in the *Unexpected* treatment and 30 pairs in the *Expected* treatment). We continued to observe a bias towards Divider P in the new experiment. Below we report the results when both dividers gave \$2. The results of the negative bias toward Divider P when both gave \$0 are reported in Appendix C.

On average, receivers gave significantly more to Divider P than to Divider N under Rule 1 (*Expected*: 3.2 vs. 1.8; *Unexpected*: 2.8 vs. 2.2;  $p < 0.01$ ). Divider P also received significantly more under Rule 2 than Divider N under Rule 3 (*Expected*: 3.4 vs. 2.1; *Unexpected*: 3.3 vs. 2.7;  $p < 0.01$  for both). The favoritism remained even when it reduced efficiency. Receivers allocated significantly more to Divider P under Rule 2 than Divider N under Rule 3 (*Expected*: 3.4 vs. 2.1, *Unexpected*: 3.3 vs. 2.7,  $p < 0.01$ ). We also calculated  $\Delta_{neutral}$  and  $\Delta_{costly}$  to measure favoritism for each receiver. The distributions of  $\Delta_{neutral}$  and  $\Delta_{costly}$  are plotted in Figure D1 and D2 in Appendix D. We find 74% receivers displayed favoritism when it was efficiency neutral ( $\Delta_{neutral} > 0$ ) in the *Expected* treatment and 44% did so in the *Unexpected* treatment. When favoritism reduced efficiency, about 64% receivers in the *Expected* and 50% in the *Unexpected* treatment favored Divider P ( $\Delta_{costly} > 0$ ).

---

<sup>20</sup> Note that this is not the case when we compare between Rule 2 and Rule 3. While the first experiment does not allow  $\Delta_{neutral} = 0$ ,  $\Delta_{costly}$  can still be equal to zero. Thus, in principle, allowing equal splits in the second experiment may reduce the proportion of positive  $\Delta_{neutral}$ , but it should not affect the distribution of  $\Delta_{costly}$ . In practice, however, subjects in the first experiment may infer that they should treat the two dividers differently when equal split is not allowed. If so, allowing equal splits can also have an impact on  $\Delta_{costly}$  in the second experiment.

The data from this new experiment show that the overall majority of receivers still favored Divider P even when they could treat the two dividers equally. On the other hand, compared with treatments where equal splits were not allowed, the percentage of efficiency-neutral favoritism ( $\Delta_{neutral} > 0$ ) was significantly lower when we allowed equal splits (*Expected*: 74% vs. 97%; *Unexpected*: 44% vs. 87%. Z-tests,  $p \leq 0.01$ ). This result suggests that some receivers would treat the two dividers equally if possible, but they would favor Divider P when an equal treatment is not allowed. Note that both the previous experiment and this new experiment allow  $\Delta_{costly} = 0$ . It is not surprising that the differences in the proportion of positive  $\Delta_{costly}$  between these two experiments are much smaller and not significant (*Expected*: 64% vs. 75%; *Unexpected*: 50% vs. 66%. Z-test,  $p > 0.10$ )<sup>21</sup>.

In sum, the data from the second experiment suggests that the pure gift effect exists even when the recipients are given the option to treat the gift givers equally<sup>22</sup>.

## 7. Conclusion

Gift giving is a widespread phenomenon that has an important impact on how decisions are made in politics, business, the legal system, the health care system, sports, education, and many other domains with substantial economic consequences. In this paper, we design an experiment to understand the driving force underlying the gift recipient's reciprocal behavior toward the gift giver when such behavior harms a third party's interest. We show that intentions alone cannot explain reciprocity with a negative externality. We introduce the ideal of moral luck in the existing intention-based model and provide both a theoretical framework and experimental evidence for a pure gift effect: a gift receiver will favor the gift giver from whom she *receives* a gift over a third party who paid the identical cost to signal the *intention* to give. Such favoritism persists even when it leads to less efficient social outcomes, and even under circumstances when the recipients know that the gift giver does not expect any return.

---

<sup>21</sup> The results of the comparisons between the second and the first experiments are the same when we compare the magnitude of  $\Delta_{neutral}$  and  $\Delta_{costly}$ .

<sup>22</sup> It is interesting to note that, similar to the first experiment, the degree of favoritism is lower in the *Unexpected* treatment than in the *Expected* treatment. However, once again we find non-significant treatment differences in all conditions except the case where receivers have to make allocation decisions under Rule 1 when both Dividers give \$2 (44% vs. 73%, Z-test,  $p = 0.02$ ). A joint test shows that there is no significant difference between the *Expected* and the *Unexpected* treatments overall ( $p > 0.10$ , F test). Yet it is possible that the non-significance is due to the lack of power of our sample size.

Our findings contribute to the existing social preference literature by shedding light on behavior in gift exchange environments with negative externalities. We revise the intention-based model by taking into account that whether the intended action can be fully realized is often beyond the actor's control. People care not only about intentions, but also to what extent intentions contribute to ultimate outcomes. We show that our model predicts the pure gift effect. To provide clean evidence on the pure gift effect, we design the experiment such that it is transparent to the receiver that the two potential gift givers have exactly the same intention. In the naturally occurring environments, it is of course often hard to verify the true intentions independent of the outcomes. Humans might be wired to react first to tangible realizations while beliefs regarding intentions may be important as second-order considerations.

Moreover, our model applies not only to positive intentions, but also to negative ones. Our experiment shows that when both Dividers act selfishly, the Receiver's allocation decision is biased against the paired Divider regardless of whether such an allocation decision is socially optimal. This result complements a recent study on delegation where principals blame agents for bad outcomes beyond their control (Gurdal et al. 2013). Our study suggests that the blameworthiness of negative intentions can be reduced if, by chance, the intentions did not lead to bad outcomes.

Our results provide behavioral insights for recent research on corruption. Hicken et al (2015) compared effectiveness of two intervention mechanisms on controlling vote switches due to vote-buying activities. In one mechanism, voters promise not to accept any gift from the candidates; in the other mechanism, voters can accept gifts, yet promise to vote according to their conscience. Our model predicts that the former mechanism will be more effective than the latter for the reason that in the former mechanism, by not accepting the gift, voters receive only the intention from the candidate. By contrast, in the latter mechanism, the voters receive not only the intention but also the actual gift from the candidate. This prediction is consistent with their findings.

We show that the favoritism does not vary with the purpose of giving, whether it was to influence the recipient's decisions or to be purely kind to the recipient. For example, even altruistic donation may lead recipients to favor donors at a cost to others. In situations where recipients' favoritism toward donors may result in less efficient social outcomes, it might be

socially desirable to keep anonymity of all donors, both those who expect reciprocity and those who do not have such expectations.

To provide clean evidence for the pure gift effect, in our study, the gift recipient is “forced” to receive the gift. In naturally occurring environments, an agent can often decide whether to accept the gift. Recent studies have shown that an option to choose the gift does have an impact on receiver’s decisions (Gneezy et al 2013). It seems important to investigate to what extent an intended recipient may resist the temptation to accept a gift, and how a recipient’s allocation decisions may differ depending on whether she has that choice. Our own continuing research targets these questions.

## References:

- Abbink, K. (2004) Staff Rotation as an Anti-corruption Policy: An Experimental Bribery Game. *Journal of Law, Economics, and Organization*. 18, 428-454.
- Abbink, K, Irlenbusch, B., and E., Renner. (2002) An Experimental Bribery Game. *Journal of Law, Economics, and Organization*, 18, 428-454.
- Abbink and Henning-Schmidt (2006) Neutral versus loaded instructions in a bribery experiment. *Experimental Economics*. 9, pp. 103–121
- Akerlof, G. (1982) Labor Contracts as Partial Gift Exchange. *The Quarterly Journal of Economics*, 97 (4): 543-69.
- Al-Ubaydli, O. and Lee, M. (2009) An experimental study of asymmetric reciprocity. *Journal of Economic Behavior & Organization*, 72, 738-749.
- Andreoni, J., and Miller, J. (2002). Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism. *Econometrica*, 70, 737-753.
- Ayres, I., and Bulow, J. (1998). The Donation Booth: Mandating Donor Anonymity to Disrupt the Market for Political Influence. *Stanford Law Review*, 50, 837-891.
- Barr, A., and Serra, D. (2010). Corruption and culture: An experimental analysis. *Journal of Public Economics*, 94(11-12), 862–869.
- Bazerman, M., and Tenbrunsel, A. (2011) Blind spots: Why we fail to do what's right and what to do about it. *Princeton University Press*.
- Bolton, G. and Ockenfels, A. (2000). ERC: A Theory of Equity, Reciprocity and Competition. *The American Economic Review*. 90, 166-193.
- Carmichael, L. and Macleod, B. (1997) Gift Giving and the Evolution of Cooperation. *International Economic Review*, 38 (3), 485-509.
- Charness, G and Dufwenberg, M. (2006) Promises and Partnership. *Econometrica*, 74, 1579-1601.
- Charness, G. and Levine, D. (2007) Intention and Stochastic outcomes: An Experimental Study. *The Economic Journal*, 117, 1051-1072.
- Charness, G. and Rabin, M. (2002) Understanding Social Preferences with Simple Tests. *The Quarterly Journal of Economics*. 117, 817-869.
- Cialdini, R. (2001) Influence: science and practice. 4<sup>th</sup> edition. Person Education; Needham Heights, MA.

- Currie, J., Lin, W. and Meng, J. (2013) Social networks and externalities from gift exchange: Evidence from a field experiment. *Journal of Public Economics*, 107, 19-30.
- Dai, Z. Galeotti F. and Villeval M. (2016) Cheating in the Lab Predicts Fraud in the Field: An Experiment in Public Transportations. *GATE working paper 5824*.
- Dufwenberg, M. and Kirchsteiger, G. (2004). A Theory of Sequential Reciprocity. *Games and Economic Behavior*, 47, 268-298.
- Falk, A. (2007) Gift Exchange in the Field. *Econometrica*, 75, 1501-1511.
- Fehr, E., and Schmidt, K. (1999) A Theory of Fairness, Competition and Cooperation. *The Quarterly Journal of Economics*, 114, 817-868.
- Fischbacher, U. (2007) z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*.10:171-178.
- Finan, F. and Schechter, L. (2012) Vote-buying and Reciprocity, with Laura Schechter, *Econometrica*, 80(2): 863-882.
- Geanakoplos, J., Pearce, D. and Stacchetti, E. (1989) Psychological games, *Games and Economic Behavior*, 1, 60-79.
- Gul, Faruk and Pesendorfer, Wolfgang (2010) The canonical type space for interdependent preferences, working paper of Princeton University.
- Gneezy, U., Saccardo, S., Veldhuizen, R. (2013) Bribery: Greedy versus Reciprocity. Working paper.
- Gurdal, M., Miller, J. and Rustichini, A. (2013) Why Blame? *Journal of Political Economy*, 121 (6), 1205-1247.
- Herbst, D and Mas A. (2015) Peer effects on worker output in the laboratory generalize to the field. *Science*, 350 (6260), 545-549.
- Hicken, A., Leider, S., Ravanilla, N., Yang, D. (2015) Measuring Vote-Selling: Field Evidence from the Philippines. *American Economic Review: Papers & Proceedings*, 105 (5): 352-356.
- Houser, D. and Xiao, E. (2011) Classification of Natural Language Messages Using A Coordination Game. *Experimental Economics*, 14 (1), 1-14.
- Kroll, M and Rustagi D. (2016) Got Milk? Motivation for Honesty and Cheating in Informal Markets: Evidence from India. *SAFE working paper*, No 127.

- Levine, D. (1998) Modeling Altruism and Spitefulness in Experiments. *Review of Economic Dynamics*, 1, 593-622.
- Huang, P. (2000). Reasons within passions: emotions and intentions in property rights bargaining. *Oregon Law Review*, 79, 435-478.
- Lowenstein, G., Cain, D. and Sah, S. (2011) The limits of transparency: Pitfalls and potential of disclosing conflict of interest. *American Economic Review: Papers and Proceedings* 101:3, 423-428.
- Malmendier, U. and Schmidt, K. (2012) You Owe Me. NBER working paper No. 18543.
- Mauss, M. (1954) *The Gift: Forms and Functions of Exchange in Archaic Societies*. By *Cohen & West*.
- Moore, D., Tetlock, P., Tanlu, L. and Bazerman M. (2006) Auditor Independence: Moral seduction and strategic issue cycling. *Academy of Management Review*, 31 (1), 1-000-000.
- Nagel, T. (1979) *Mortal Questions*. *New York: Cambridge University Press*, 1979
- Nelkin, K., "Moral Luck", *The Stanford Encyclopedia of Philosophy* (Summer 2013 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2013/entries/moral-luck/>
- Rabin, M. (1993) Incorporating Fairness into Game Theory and Economics. *The American Economic Review*.83, 1281-1302.
- Scanlon, T.M. (2008) *Moral Dimensions: Permissibility, Meaning, Blame*. *Cambridge, MA: Belknap Press*.
- Small, D. and Loewenstein, G. (2003). Helping *a* Victim or Helping *the* Victim: Altruism and Identifiability. *The Journal of Risk and Uncertainty*, 26:1, 5-16.
- Stanca, L. (2010) How to be kind? Outcomes versus intentions as determinants of fairness. *Economic Letters*. 106: 19-21.
- Strassmair, C. (2009) *Can Intentions Spoil the Kindness of a Gift? – An Experimental Study*, Munich Discussion Paper No.2009-4.
- Van Dijk, F. and Van Winden, F. (1997) Dynamics of social ties and local public good provision. *Journal of Public Economics* (64) 323-341.
- Van Veldhuizen, R. (2012). The Influence of Wages on Public Officials' Corruptibility: a laboratory investigation. *Tinbergen Institute Discussion Paper TI 2012-038/1*, 1–29.



Vincente, P. (2014) Is vote buying effective? Evidence from a field experiment in west Africa.  
*Economic Journal*, 124, F356-F387.

Williams, B. "Moral Luck." *Moral Luck*. Cambridge: Cambridge University Press, 1982. pp. 20–39.