



# Do you reward and punish in the way you think others expect you to?☆

Omar Al-Ubaydli<sup>a,\*</sup>, Min Sok Lee<sup>b</sup>

<sup>a</sup> Department of Economics and Mercatus Center, George Mason University, 4400 University Drive MSN 3G4, Fairfax, VA 22030, USA

<sup>b</sup> Citadel Group Foundation, Chicago, IL, USA

## ARTICLE INFO

### Article history:

Received 19 November 2010

Received in revised form 31 January 2012

Accepted 26 February 2012

### JEL classification:

D63

D64

D84

### Keywords:

Behavioral confirmation

Guilt

Intentions

Reciprocity

Reward

Punishment

## ABSTRACT

This paper addresses three questions: (1) When deciding on whether to reward or punish someone, how does how you think others expect you to behave affect your decision? (2) Does it depend upon whether others expect you to reward them vs. punish them? (3) What is the interpretation of such a causal effect? We investigate these questions using a modification of the lost wallet trust game (a simplified version of the trust game) that permits punishment. Like previous studies, we collect data on what second-movers think that first-movers expect them to do by directly eliciting the second-movers' expectations. Unlike previous studies, we ensure exogeneity of these expectations by instrumenting for them. The instrument is the expectations of neutral observers which are disclosed to second-movers prior to the elicitation of second-movers' expectations. We find that what you think others expect you to do has a zero causal effect on both reward and punishment decisions. We also find that it is important to instrument for second-order expectations because they are endogenous. We interpret these findings in terms of models of guilt-aversion and intentional reciprocity.

© 2012 Elsevier Inc. All rights reserved.

## 1. Introduction

Reward and punishment are critical to regulating economic relationships, even in one-shot settings. Investigating their determinants – especially the potentially controllable ones – is an important step toward understanding optimal incentive schemes. One such malleable determinant is what you think others expect you to do. In this paper, we address the following empirically driven questions.

1. In the trust game, when deciding on whether to reward or punish someone, how does how you think others expect you to behave affect your decision?
2. In the trust game, does it depend upon whether others expect you to reward them vs. punish them?

In addition to these primary goals, a secondary goal is to answer the following theoretical question.

3. In the trust game, if there is a causal effect of what you think others expect you to do on your behavior, what is its interpretation?

Previous studies include the lost wallet version of the trust game (Dufwenberg and Gneezy, 2000; see Berg et al., 1995 for the trust game), shown in Fig. 1.

The sender starts with \$10 and the responder starts with nothing. The sender can either keep the \$10, ending the game, or she can send the \$10 to the responder. If the sender decides on sending the \$10, they are tripled, and the responder unilaterally decides how much of the \$30 to return to the sender ( $y$ ).

Let  $y'$  denote the sender's expectation of  $y$ . Let  $y''$  denote the responder's expectation of  $y'$ , referred to as the responder's second-order expectation of  $y$ :  $y''$  is what the responder thinks that the sender expects the responder to send back. By eliciting  $y''$  in simple variants of the trust game, Dufwenberg and Gneezy (2000) and other studies find a positive relationship between  $y''$  and  $y$ .<sup>1</sup> This is an example of *behavioral confirmation*: you are more likely to behave in a certain way if you think that others expect you to behave in that way. This is to be contrasted with its obverse: *behavioral disconfirmation*.

☆ We wish to thank Morris Altman, Martin Dufwenberg, Uri Gneezy, Dan Houser, John List, Nat Wilcox and an anonymous referee for helpful comments, and Jason Aimone for help in running the experiment.

\* Corresponding author. Tel.: +1 703 993 4538; fax: +1 703 993 1133.

E-mail address: [omar@omar.ec](mailto:omar@omar.ec) (O. Al-Ubaydli).

<sup>1</sup> Guerra and Zizzo (2004), Charness and Dufwenberg (2006), and Bacharach et al. (2007).

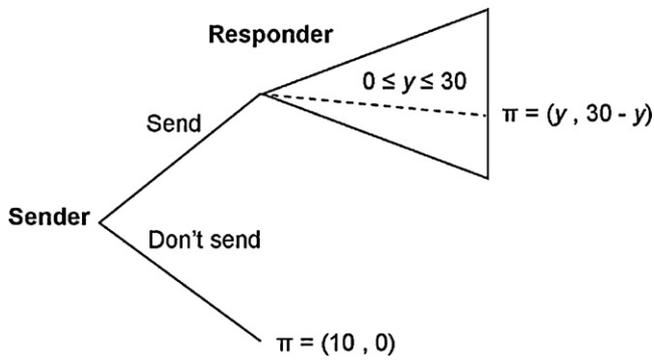


Fig. 1. The lost wallet game (Dufwenberg and Gneezy, 2000).

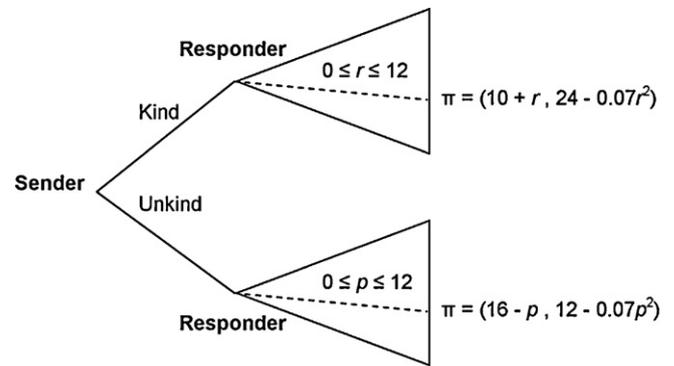


Fig. 2. The judgment game.

Behavioral confirmation and disconfirmation are descriptions of the causal effect of second-order expectations on behavior. The aforementioned studies interpret this instance of behavioral confirmation as resulting from guilt-aversion (which is what their designs are seeking to test): when  $y''$  exceeds  $y$ , the responder is failing to fulfill what she believes to be the sender's expectations. The responder feels guilty about disappointing the sender, and will feel guiltier the larger the difference between  $y''$  and  $y$ . A guilt-averse responder will therefore send back more money when she thinks that the sender expects more back.

In the trust game, the responder's only alternative to a materialistic best response is to reward the sender.<sup>2</sup> What if the responder is also allowed to punish the sender? In punishment decisions, do we expect behavioral confirmation or disconfirmation? We here consider an extended form of the trust game that allows for punishment.

In addition to extending the scope from reward to punishment decisions, we also investigate the interpretation of the data. In particular, there are theories that predict behavioral disconfirmation in the trust game, such as the model of intentional reciprocity.<sup>3</sup> Observing behavioral confirmation in the trust game is evidence in favor of guilt-aversion. However we explore why observing behavioral confirmation in the trust game is not necessarily evidence against models that predict behavioral disconfirmation in the trust game, and, crucially, why observing neither in the trust game (which we find) is not necessarily evidence against guilt-aversion.

A final issue that we explore is how second-order expectations are observed. Dufwenberg and Gneezy (2000) elicit second-order expectations directly from the responders. Since this explanatory variable is not randomly induced by the experimenter, the design risks endogeneity bias; Charness and Dufwenberg (2006) refer to this possibility as a form of the false-consensus effect. To estimate the endogeneity bias, we transmit the expectations of non-playing observers to responders and then elicit the second-order expectations of responders. We then use the expectations of the observers as instruments for the second-order expectations of the responders. In a final treatment, to permit comparability with Ellingsen et al. (2010) and Reuben et al. (2008), we transmit the expectations of each sender to her partner prior to the responder making her choice.

<sup>2</sup> Rewards are deviations from best responses that increase the sender's payoff. Punishments are deviations from best responses that decreases the sender's payoff.

<sup>3</sup> The intentional reciprocity model explains reward decisions as resulting from a desire to reciprocate kind actions, and punishment decisions as resulting from a desire to reciprocate unkind actions. See Schopler and Thompson (1968), Pruitt (1968), Tesser et al. (1968), Greenberg and Frisch (1972), Blount (1995), McCullough et al. (2001) and Ames et al. (2004) as well as Rabin (1993), McCabe et al. (2003), Dufwenberg and Kirchsteiger (2004), Falk and Fischbacher (2006) and Cox et al. (2007).

Our results are as follows. When using observer expectations as an instrument for second-order expectations, second-order expectations have no effect on both reward and punishment behavior, i.e., we find neither behavioral confirmation nor behavioral disconfirmation. Similar results are obtained when we transmit expectations to each sender's partner. When we elicit expectations directly, we find behavioral confirmation in both reward and punishment decisions, suggesting that elicited expectations are endogenous.

## 2. Experimental design

### 2.1. The judgment game

The judgment game is shown in Fig. 2. It is a non-linear, continuous version of the games in Dufwenberg and Gneezy (2000) and Offerman (2002). It is a variant of the trust game where the responder can punish as well as reward.

The sender starts with \$16 and the responder with \$12. If the sender plays *unkind*, the payoffs are unchanged. The responder can then punish the sender by reducing the sender's payoff by  $p \in [0, 12]$ . This costs the responder  $0.07p^2$ .<sup>4</sup>

If the sender plays *kind*, then she transfers \$6 to the responder, which are then doubled. The responder can then reward the sender by increasing the sender's payoff by  $r \in [0, 12]$ . This costs the responder  $0.07r^2$ .

The quadratic cost of reward (punishment) means that the marginal cost of reward (punishment) rises from \$0 to \$1.7 as  $r$  goes from 0 to 12 ( $p$  goes from 0 to 12). We selected a quadratic cost as it implies an interior solution under conventional models of behavioral preferences, such as Dufwenberg and Kirchsteiger (2004) and Falk and Fischbacher (2006).

Let  $r''$  be the responder's second-order expectation of  $r$ , i.e., what value of  $r$  she thinks that the sender expects her to pick. Similarly, let  $p''$  be the responder's second-order expectation of  $p$ .

**Research question 1:** In the judgment game, what is the relationship between reward and second-order expectations of reward,  $\partial r / \partial r''$ ? What is the relationship between punishment and second-order expectations of punishment,  $\partial p / \partial p''$ ?

**Research question 2:** In the judgment game, do  $\partial r / \partial r''$  and  $\partial p / \partial p''$  differ?

<sup>4</sup> In the experiment, we used a neutral frame. Senders were *Blues* and responders were *Reds*. *Kind* was *Dash* and *unkind* was *Solid*. Reward was *Increase Blue's earnings* and punish was *Decrease Blue's earnings*. Also, payoff consequences of actions were not expressed as the sender starting with an amount that she can choose to send to the responder. Rather the sender was choosing between two actions each with a certain payoff consequence. See the instructions in Appendix.

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات