Journal of Behavioral and Experimental Economics 000 (2016) 1–10

Contents lists available at ScienceDirect

Journal of Behavioral and Experimental Economics

journal homepage: www.elsevier.com/locate/jbee

The effect of third party intervention in the trust game

Marina Fiedler[a,1], Ernan Haruvy[b,*]

[a] Institute for Management, People and Information, University of Passau, Innstr. 27, 94032 Passau, Germany
[b] Jindal School of Management, University of Texas at Dallas, 800 W. Campbell Rd., Richardson, TX 75080

ARTICLE INFO

Article history:
Received 29 February 2016
Revised 29 October 2016
Accepted 31 October 2016
Available online xxx

JEL classification:
C93
C99
D63

Keywords:
Trust game
Investment game
Punishment
Reward
Reinforcement

ABSTRACT

The literature suggests that in the trust game setting a third party may have an impact on trust and trustworthiness. This may be done through monitoring alone, as well as through punishment or reward. We examine the impact of these three factors in both fixed and random partner settings. We find evidence that third party intervention is sensitive to participant actions and can result in changes in investment and return behavior, although not necessarily in the intended direction. Despite this individual impact, adding punishment or reward capabilities to the third party monitor has no significant effect in the aggregate. The increase in contributions in the presence of a third party can primarily be attributed to third party monitoring.

1. Introduction

We examine whether third party monitoring can result in greater efficiency and equity in the trust game, – an economic game intended to abstract investment opportunism. And, if so, whether giving this third party additional reinforcement tools would aid this third party towards these goals. Opportunism is behavior intended to benefit one firm at the expense of an investing partner firm in violation of the investing firm's trust (Williamson, 1993a, b, 1998) through expropriation of investment profits by the receiving firm (Rokkan et al., 2003). This is a common concern in business partnerships and relationships (see Geyskens et al., 2006 and Rindfleisch and Heide, 1997). The literature proposes two sets of governance mechanisms: (1) economic incentives by a (third party) government agency to discourage opportunistic behavior (Anderson and Weitz 1992; Williamson, 1993b) and (2) transparency, leading to better monitoring and disapproval, also through (third party) government agencies (Antia et al., 2013). The current research, in the abstract environment of the trust game, addresses the relative impact and effectiveness of these two sets of mechanisms.

A third party in the present context is a party who can affect the payoffs of others but whose own monetary payoff is affected only by its own actions; this party has monetary disincentives to intervene. The question we ask is whether it is helpful to provide this monitoring third party with additional tools such as punishment and reward capabilities. Specifically, we investigate whether the third party reward and punishment effects add value to the monitoring effect.

The setting for this investigation is the trust game. We are interested in disentangling the effects of monitoring, punishment and reward incentives on trust, as captured by investment levels, and trustworthiness, as captured by the proportion returned. We first address this question by examining treatment effects in experimental manipulations that vary the nature of the feedback, the presence of monitoring, and the punishment and reward capabilities.

Further, we are interested in mapping the feedback loop between third party and primary agents' actions. That is, we expect third party actions to be affected by the observed actions of the primary agents in the current period and to – in turn – impact the next period behavior by these same agents. We seek to know the extent of each party's responsiveness to the other parties, in the hope that this mapping would shed light on the aggregate differences or lack thereof.

In the aggregate analysis, we find that, in the trust game with random partners, the aggregate effect of adding a third party mon-

* Corresponding author. Fax: +972 883 6727.
E-mail addresses: Marina.Fiedler@uni-passau.de (M. Fiedler), eharuvy@utdallas.edu (E. Haruvy).
[1] Fax: +49 851 509 37 2490.

http://dx.doi.org/10.1016/j.socec.2016.10.003
2214-8043/

itor with punishment or reward capabilities is largely due to monitoring. That is, on aggregate, third party monitoring with punishment or reward has no additional effect on the aggregate, compared to third party monitoring alone.

In individual analysis, examining the responsiveness of individual players to incentives, we find that while monitors with the power to punish and reward are responsive to other participants' actions, investors and responders may not respond to punishment and reward as intended. In punishment conditions with repeated observations, investors and responders perceive an absence of punishment to unkind actions as positive reinforcement. Thus, the possibility of punishment followed by the absence of such punishment appears to provide positive reinforcement to unkind behavior if the third party monitor fails to follow up such behavior with punishment. Likewise, in reward conditions, investors react as if they perceive reward to indicate that they had over-invested. Thus, they reduce investment in response to reward, which would lead to the observed detrimental effect of rewards.

2. Background

To clearly define the scope of our investigation, we must first distinguish between affected party and third party monitoring and actions, – a distinction also made in Salmon and Serra (2014). An affected party is a player who is directly monetarily affected by the actions of the other players, while a third party monitor is a player who is not directly monetarily affected by the primary players' actions. There is an established literature that examines the effects of affected party monitoring (e.g., Dugar, 2013; Masclet et al., 2003; Noussair and Tucker, 2005), as well as reward and punishment (Camerer and Fehr, 2004; Reuben and Riedl, 2009; Sefton et al., 2007; Fehr and Fischbacher 2004b; Nese and Sbriglia 2009). With some exceptions (Salmon and Serra, 2014), it could be argued that when an affected party has greater power to monitor and react, it can improve its position. Salmon and Serra find that third party monitoring (with social disapproval) can be more effective than affected party monitoring. Our investigation focuses solely on third party monitoring, punishment, and reward.

Third party monitoring. Even the mere cue or perception of being monitored– can drastically alter actions. In a seminal work, Haley and Fessler (2005) showed that placing eye-like stimuli in participants' environment during a Dictator Game (DG) caused participants to be more generous towards an anonymous other person. While the 'watching eye' effect has been replicated elsewhere in the DG game, Fehr and Schneider (2010) found no such effect in the trust game, the game we study here. They conclude that eye cues have lower impacts on strong reciprocity. While visual cues are known to be powerful, de facto monitors are also major influences. Gneezy and List (2013) provide substantial field evidence about the power of observation by strangers or acquaintances to increase social behavior, including charitable giving. The roles of third parties are also important. When parties in positions of leadership or authority have the ability to monitor, reward, or punish, these effects are often observed to be greater (Cardenas and Carpenter, 2008; Ostrom, 2002; Kosfeld and Rustagi, 2011).

Third-party reward and punishment. In recent research, it has been suggested that a separate third party could act effectively in the role of punisher. Charness et al., (2008) analyze the effects of the possibility of third-party intervention (punishment as well as reward) on behaviors in a variant of the investment game (Berg et al., 1995). They find that first mover transfers are more than 60% higher when there is the possibility of third party punishment, and that there is very little difference in first movers' behavior if the third party can also choose to reward the first mover. Further, they find that introducing a third party increases the receiver's responsiveness to the first mover's transfer.

Sutter et al., (2009) obtain a similar finding in a prisoner's dilemma game in that they find that third party rewards have positive effects on cooperation. However, they find that while third party rewards can be effective, most of that impact can be attributed to the monitoring by the third party rather than the reward mechanism. That is, the mere observation of players' actions through an unaffected third party raises cooperation rates by about 50%. The reward mechanism contributes only an additional 10% to cooperation rates. We call this the monitoring effect.

The threat (promise) effect. There are two elements to the effect of punishment or reward. The first element is the force of the threat (promise) of possible punishment (reward). The mere threat of being sanctioned or rewarded sustains and enhances cooperation (Sefton et al., 2007). Reuben and Riedl (2009) showed, with a public goods game, that the possibility of punishment largely eliminates the downward trend in contributions over time. Fehr and Fischbacher (2004a), showed that the possibility of punishment enhances the contributions between participants.

The reinforcement effect. The second element in the effect of punishment or reward is the reinforcement effect. This is the effect of the actual punishment or reward on subsequent behavior. Bernhard et al., (2006a, b) showed that participants' behavior changes significantly as soon as a third party intervenes. Charness et al., (2008) showed that third party punishment and reward had strong impacts. Fehr and Fischbacher (2004a) found that transferred amounts by participants adapt to a fair equity level when punished by third party participants. Punishment expectations were shown to rise when punishment occurred (Fehr and Fischbacher, 2005). A similar pattern has been shown with rewards. Sutter et al., (2009) found that rewards have positively affected cooperation rates.

Third party responsiveness to proposer and responder actions. Intervention by a third-party will occur when the outcome between proposer and responder is perceived as unfair. The literature shows individuals are willing to punish strangers for unfair actions, even when these actions do not directly affect them (Carpenter et al., 2004; Fehr and Fischbacher, 2004a, b). Not only have third parties been shown to intervene when unfair actions occur, their punishments also are in line with an infraction's severity (Fehr and Fischbacher, 2004a, 2005; Charness et al., 2008; Ottone, 2005). Similar patterns have been documented concerning rewards. Sutter et al., (2009) found that reward activity is lowest in the very first period, when cooperation rates are by far highest and greater than 50%. As cooperation rates drop, third party intervention rises, in an attempt to increase cooperation rates.

3. Hypotheses

The monitoring hypothesis. As described in the introduction, our first and main conjecture is that monitoring should have a critical role to play, - a role so important that it might account for much of the documented effect of third party punishment. In short, the monitoring hypothesis is that the monitoring condition would have a large impact on proposers and responders' behavior. As noted in the above paragraph, Sutter et al., (2009) suggest that third party monitoring has a large effect-about 50% increase in cooperation-in a prisoner's dilemma game. But this monitoring effect can trace its roots back to the early experimental research about the effect of double-blind vs. single-blind procedures. Hoffman et al., (1996), for instance, showed that instituting a double-blind procedure in a dictator game, in which the experimenter cannot observe individual participants' contributions, drastically increases the incidence of selfish behavior. Thus, while it is well understood in the literature that a substantial monitoring effect exists, it is rarely discussed as a mechanism and is not generally separated from the punishment effect in mechanisms that institute punishment.

Please cite this article as: M. Fiedler, E. Haruvy, The effect of third party intervention in the trust game, Journal of Behavioral and Experimental Economics (2016), http://dx.doi.org/10.1016/j.socec.2016.10.003