What is Trustworthiness and What Drives It?*

James C. Cox[a], Rudolf Kerschbamer[b] and Daniel Neururer[c]

Forthcoming in *Games and Economic Behavior*

Abstract

This paper experimentally isolates the impact of various combinations of the following motives on trustworthiness: (i) unconditional other-regarding preferences; (ii) vulnerability-responsiveness; (iii) deal-responsiveness; and (iv) gift-responsiveness. Our results indicate that – besides unconditional other-regarding preferences like altruism and inequality aversion – vulnerability-responsiveness is the most important driver for trustworthiness. Prompted by our experimental findings we provide behavioral definitions of trust, trustworthiness, and trust-responsiveness based on revealed willingness to accept vulnerability and the response to it. An important difference from existing definitions is that ours allow trust to be present without generosity.

JEL classification: C70, C91, D63, D64

Keywords: trustworthiness, trust, trust game, investment game, vulnerability-responsiveness, deal-responsiveness, gift-responsiveness, generosity, reciprocity

---

* Financial support from the vice-rectorship of research of the University of Innsbruck through grant number 2013/1/VWL21, from the Austrian Science Fund (FWF) through grant numbers P20796 and P22669, and from the Austrian National Bank (OeNB Jubiläumsfonds) through grant number 13602 is gratefully acknowledged.

[a] Experimental Economics Center & Department of Economics, Georgia State University: jccox@gsu.edu

[b] Department of Economics, University of Innsbruck: Rudolf.Kerschbamer@uibk.ac.at

[c] Department of Economics, University of Innsbruck: Daniel.Neururer@uibk.ac.at

1 Introduction

The economics literature is full of statements stressing the importance of trust and trustworthiness for all kinds of human interactions. Fehr (2009, p.235), for instance, writes that "…trust plays a role in almost all human relationships. It permeates friendship relations, family relations, and economic relations. People rely on the support of their friends, children trust their parents, and sellers trust their buyers to pay the bill." Some empirical studies go even further by suggesting a (causal) relationship between people's perception of others' trustworthiness at the country level and important macro variables such as GDP growth (Knack and Keefer 1997), inflation (La Porta et al. 1997), or trade volume (Guiso et al. 2009).

Despite their importance for all kinds of economic relationships, there is no consensus in the literature as to what defines trust and trustworthiness and what drives them – see Coleman (1990), Bacharach et al. (2007) and Fehr (2009) for discussions. One contribution of the present paper is the provision of formal definitions of trust and trustworthiness based on observable variables. The second contribution is the experimental isolation of the relative importance of the following drivers for trustworthiness in a specific class of games: unconditional other-regarding preferences, vulnerability-responsiveness, deal-responsiveness and gift-responsiveness. As we will argue below, our two contributions are intimately related to each other.

The non-economic literature contains many verbal definitions of trust. In most of them "vulnerability" plays a central role. Mayer et al. (1995, p. 712), for instance, define trust as "… the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party." Similar definitions can be found in McAllister (1995), Rousseau et al. (1998), Ferrin and Dirks (2003), Ho and Weigelt (2005), and Colquitt et al. (2007), among others. The second term – trustworthiness – is frequently used, but without explicit definition.

The economic literature typically defines trust and trustworthiness in terms of behavior in specific games. The most familiar example is the investment game introduced by Berg et al. (1995).[1] In this two-stage game there are two players – a first mover (FM, he) and a second mover (SM, she). The players start with identical initial endowments – of e, say. In the first stage, the FM decides on the amount $t \in [0, e]$ he wants to transfer to the SM, knowing that if he transfers t the SM will receive kt, with $k > 1$ (typically, k is 2, 3, 4, or 5). In the second stage, the SM sees what the FM has done and then decides on the amount $r \in [0, kt]$ that she wants to return to the FM. After this move the game ends with material payoffs of $f = e - t + r$ for the FM and $s = e + kt - r$ for the SM.[2]

Different definitions of trust and trustworthiness have been used within the context of the investment game. Berg et al. (1995), Burks et al. (2003) and Ben-Ner and Putterman (2009) state that a FM trusts the SM if he sends more than the minimum possible amount. Cox (2004) states that (the outside observer can conclude that) the FM trusts the SM if he sends an amount that makes himself vulnerable to loss of utility, as evidenced by sending more in the investment game than in the *first-mover* dictator control game with the same feasible set. Some authors state that the SM is said to be trustworthy if she returns more than the minimum possible amount (Burks et al. 2003 and Ben-Ner and Putterman 2009 use this weak definition), or an amount that exceeds the amount sent (Schotter and Sopher 2006 and Chaudhuri and Gangadharan 2007 employ this more demanding definition).

One aim of the present paper is to propose definitions of trust and trustworthiness that are applicable not only to a specific game but rather to a non-trivial class of games – the class of two-player two-stage games with observable actions. The form of the definitions is inspired by the formal definitions of generosity and reciprocity introduced by Cox et al. (2008) and their content is motivated by the results of experiments designed to yield insights into the main drivers for trustworthiness.

The experimental identification of drivers of trustworthiness is the second aim of the paper. The investment game (IG) and *second-mover* dictator control game (DG) in Cox

[1] Indeed, this game is often called “the trust game” – by Croson and Buchan (1999) and Karlan (2005), for instance.

[2] Berg et al. (1995) implement a discrete version of this game with $e = 10$ and $k = 3$ and find in their ‘no history’ treatment that 28 out of 32 FMs send more than the minimum possible amount to SMs and that out of the 28 SMs who received more than the minimum possible amount 11 shared the increase in total surplus by returning more than the FM sent (but strictly less than they received). Qualitatively similar patterns are reported by Johnson and Mislin (2011) in their meta-study examining 162 replications of the Berg et al. (1995) investment game.

(2004) are relevant for this goal. The distribution of endowments over dictator-recipient pairs in the DG is chosen in such a way that it matches exactly the empirical distribution of material payoffs over SM-FM pairs after FMs' choices in the IG. In other words, the population of dictators in the DG treatment faces exactly the same distribution over opportunity sets as the population of SMs in the IG treatment, but in each pair the FM intentionally caused the choice set in the IG, while the experimenter predetermined the choice set in the DG. For a given choice set of the SM any difference between transfers in the DG treatment and back-transfers in the IG treatment indicates that unconditional other-regarding concerns (such as altruism, inequality aversion, maximin etc.) alone are not sufficient to explain back-transfers in the IG. Cox (2004) finds that back-transfers in the IG are approximately one-third higher than transfers in the DG and concludes that back-transfers in the IG are in part motivated by *conditional* other-regarding concerns.[3]

The Cox (2004) design does not allow – and is not intended to allow – for discrimination between different potential explanations for the conditional part of the other-regarding preferences of the SM. Our approach to digging deeper into drivers of trust and trustworthiness begins by identifying several possible motivators that co-vary in the investment game. First, because the SM can keep any proportion of the tripled amount transferred by the FM, the FM makes himself vulnerable to loss of monetary payoff by making a positive transfer. A SM may be motivated to respond to the FM's willingness to be vulnerable by making a positive back transfer. We call this motivation – reacting to the vulnerability of the trustor by adopting actions that do not hurt the trustor – *vulnerability-responsiveness*. Secondly, the tripling of amounts transferred makes possible mutually profitable "deals" for splitting the profit resulting from positive transfers. A SM may be motivated to respond by making a positive back transfer. We term this motivation – reacting to actions that allow for a mutual improvement by adopting behavior that implies a mutual improvement – *deal-responsiveness*. Thirdly, by transferring any positive amount the FM makes a gift of the tripled transfer should the SM choose to accept it. A SM may be motivated to respond to this gift by making a positive back transfer. We term this motivation – reacting

[3] Earlier contributions that disentangle unconditional from conditional other-regarding preferences are Charness (1996; published as Charness 2004) and Offerman (2002).

to choices that allow the trustee to obtain an improvement by adopting actions that benefit the trustor – *gift-responsiveness*.

To isolate the relative importance of each of these conditional motivations for SM back transfers we use an experimental design that comprises five treatments implemented between subjects. Treatment IG is a binary investment game in which the FM decides whether or not to send his endowment to the SM. If the FM does not send his endowment, the interaction ends and both players get as payoffs their equal endowments of 15. If the FM sends his endowment it is tripled and the SM then decides how to divide her enhanced endowment of 60 between herself and the FM. Arguably all the above motives affect SM behavior in this treatment.

We then investigate SM behavior in four other treatments. In all these treatments the SM faces exactly the same choice set as in IG. The treatments differ in what happens in case the FM decides not to send the endowment to the SM, and whether the FM can make such a decision at all. In the conditional-controlled treatment CC the FM makes no sending decision. Arguably, only unconditional other-regarding preferences can affect SM behavior in this treatment.[4] In the remaining three treatments the FM makes a sending decision. In the vulnerability-controlled treatment VC sending the endowment to the SM does not make the FM vulnerable to the behavior of the SM. This eliminates vulnerability-responsiveness as possible motivation for back-transfers. In the deal-controlled treatment DC sending the endowment does not increase efficiency. This eliminates deal-responsiveness as possible motivation. In the gift-controlled treatment GC sending the endowment does not allow the SM to get a higher payoff. This eliminates gift-responsiveness as possible motivation.

We then compare the return behavior of SMs across the 5 treatments. We find that average returns are around 12 in the treatments IG, DC, and GC, with no statistically significant differences across these three treatments. By contrast, average returns are only around 6 in the treatments VC and CC, and again, there is no statistically significant difference across these two treatments. Finally, the differences between the high return treatments (IG, DC, GC) and the low return treatments (VC, CC) are statistically significant.

[4] The treatments IG and DG in Cox (2004) correspond to our treatments IG and CC, the most important difference being that FMs in Cox's IG have a richer choice set – they are asked to transfer to the paired SM none, some or all of their endowment, while they face a binary choice in our IG.

We therefore conclude that the conditional part of the back-transfer in the IG is largely driven by vulnerability-responsiveness.

Prompted by our empirical findings we then provide formal definitions of trust-related concepts. For this purpose we first define a partial ordering over opportunity sets in the own-money, other's-money space – the "More Trusting Than" relation. A key ingredient of this definition is the vulnerability of the FM. Based on this definition we then provide behavioral definitions of trust act, trustworthy act and trust-responsive motivation. An important difference to existing definitions is that ours allow trust to be present without generosity. We consider this as important because in many interesting real world examples where we would consider it as natural to speak of "trusting behavior" there is no generosity involved in trust acts.

The rest of the paper is organized as follows. Section 2 describes the five mini-games used for discrimination between motives and provides definitions of vulnerability-responsiveness, deal responsiveness, and gift-responsiveness. The experimental design and the procedure are introduced in Section 3. Section 4 formulates the hypotheses. Section 5 presents and discusses the experimental results. Section 6 provides the behavioral definitions of trust, trustworthiness and trust-responsiveness and Section 7 concludes.

2 Identification Games, Revealed Intentions, and Revealed Preferences

In order to discriminate among the mentioned motives for conditional back transfers in the IG, we extend the Cox (2004) design to five diagnostic two-player games. In each of the games a FM chooses an opportunity set for a SM from a given collection of opportunity sets which is common information. Then the SM chooses an allocation (implying a material payoff for each of the two players) from the opportunity set chosen by the FM. In line with the formal framework introduced by Cox et al. (2008) our main hypothesis is that the SM's interpretation of a given choice by the FM depends not only on the actual choice made by the FM but also on the alternative choice(s) that would have been available to the FM. By systematically varying the alternative choice(s) available to the FM we address the question of how the SM's decision within a given opportunity set is affected by various combinations of revealed intentions behind the FM's choice.

2.1 Identification Games

The five treatments of our experimental design correspond to the five two-player games shown in Figure 1. All of the games have the same non-trivial feasible set for the SM. In four of the games the feasible set for the SM is reached only if the FM chooses option (b). The structure of those games is displayed in Panel A of Figure 1. These four games differ only with respect to the allocation (contained in a singleton opportunity set) that is implemented if the FM chooses option (a). The payoffs for option (a) for the different games are displayed in Table 1. The other game – shown in Panel B of Figure 1 – has the same option (b) feasible set for the SM but no option (a) for the FM. The five games are:

- **IG** is a binary **investment game**. The main difference from the original Berg et al. (1995) design is that the collection of opportunity sets the FM is asked to choose from consists of only two elements: The FM can either (a) not invest, resulting in a singleton opportunity set for the SM – the allocation $(f,s)=(e,e)$; or (b) invest the whole endowment, resulting in a non-trivial opportunity set for the SM – for which she can return any amount between 0 and $4e$.[5] Panel A of Figure 1 displays the resulting material game using the parameters and the discrete grid implemented in the experiment.[6] Note the following three important features of the choice of option (b) by the FM in the IG: (i) it makes the FM vulnerable – if the SM returns less than e then the FM is worse off in material terms, compared to option (a); (ii) it allows for a mutual improvement – if the SM returns more than e but less than $3e$ then both players are strictly better off; and (iii) it involves a gift to the SM – the maximal material payoff she can realize with option (b) exceeds the payoff she receives in (a).

< Insert Figure 1 and Table 1 about here >

- **CC** is the second-mover dictator control or **conditional-controlled investment game:** It differs from IG only in the fact that option (a) is not a feasible choice for the FM; that is, the collection of opportunity sets for the FM consists of a single element in this treatment,

[5] That is, in contrast to the standard investment game we allow the SM to transfer not only the amount received but also the initial endowment. We do not consider this detail as important, though.

[6] That is, in the experiment e =15 and the SM's choice set is restricted to elements in {0, 5, 10, ... , 50, 55, 60}.

the option (b). So, no motives of the FM are revealed in this treatment because no choice is made by the FM.

- **VC** is the **vulnerability-controlled investment game:** It differs from IG only in the payoffs for option (a), which are such that the choice of option (b) does not make the FM vulnerable in comparison to option (a). Still, option (b) allows for a mutual improvement and for a unilateral improvement for the SM.

- **DC** is the **deal-controlled investment game:** It differs from IG only in the payoffs for option (a), which are such that option (b) does not create the possibility of a mutual improvement in comparison to option (a). Still it keeps the features that – in comparison to option (a) – option (b) allows the SM to obtain an improvement and makes the FM vulnerable.

- **GC** is the **gift-controlled investment game:** It differs from IG only in the payoffs for option (a), which are such that the choice of option (b) does neither create the possibility of a mutual improvement nor allow the SM to obtain a unilateral improvement. Still, the choice of (b) makes the FM vulnerable.

We next present an informal development of a theory of FM revealed intentions and SM revealed preferences for the class of two-player, two-stage games of complete and perfect information. The formal development of the theory is contained in Appendix A.

2.2 First Mover's Revealed Intentions

We begin by defining three partial orderings of FM opportunity sets. Opportunity set J is said to entail more vulnerability for the FM than opportunity set K (written as J **MVT** K) if the maximum income the FM can get in set J, given that the SM acts selfishly in own-money terms, falls short of his maximal income in set K under the same condition. Opportunity set J is said to be more efficient than opportunity set K (written as J **MET** K) if: (a) the maximal feasible sum of payoffs to the two players in set J is larger than the maximal feasible sum in set K; and (b) the set J contains an allocation that is better for both players (in own-money terms) than the allocation that is realized in set K when the SM acts selfishly in own-money terms. Opportunity set J is said to be more beneficial for the SM (written as J **MBT** K) if the maximum material income the SM can get in set J exceeds the maximum material income she can realize in set K. These partial orderings of opportunity sets are useful for analyzing the relationship between revealed intentions of FMs and revealed preferences of SMs.

First consider a FM's revealed intentions. Let the FM choose the actual opportunity set J for the SM from a known collection of available opportunity sets that also contains the set K:

(a) If J **MVT** K then we say that the FM's choice of J reveals the **willingness to accept vulnerability**;

(b) If J **MET** K then we say that the FM's choice of J reveals the **willingness to create the possibility for a mutual improvement**;

(c) If J **MBT** K then we say that the FM's choice of J reveals the **willingness to allow the SM to obtain an improvement**.

2.3 Second Mover's Revealed Preferences

Cox et al. (2008) define a partial ordering on preferences over income allocations – the "More Altruistic Than" (MAT) relation. We here apply that relation to SM's revealed preferences.

Let the FM choose the actual opportunity set *J* for the SM from a known collection of opportunity sets:

- If *J* is the only available opportunity set (so that the FM has no real choice to make) and if the SM is not globally indifferent to amounts of the FM's payoff then we say that the SM has **unconditional other-regarding preferences**.
- If the choice of *J* reveals the willingness to accept vulnerability and if this fact elicits more altruistic preferences in the SM compared to a situation where only *J* is available then we say that the SM's preferences exhibit **vulnerability-responsiveness**.
- If the choice of *J* reveals the willingness to create the possibility for a mutual improvement and if this fact elicits more altruistic preferences in the SM compared

to a situation where only *J* is available then we say that the SM's preferences exhibit **deal-responsiveness**.

- If the choice of *J* reveals the willingness to allow the SM to obtain an improvement and if this fact elicits more altruistic preferences in the SM compared to a situation where only *J* is available then we say that the SM's preferences exhibit **gift-responsiveness**.

3 Experimental Treatments and Procedures

For the experimental isolation of the relative importance of the four SM motivations defined in the previous section for behavior in the investment game we use an experimental design that comprises five treatments implemented between subjects. The experiment was conducted with paper-and-pen (and several other design features reported below were applied) to convince subjects that neither other subjects nor the experimenters could identify the person who has made any particular decision. This was done in an attempt to minimize the impact of experimenter demand and audience effects.[7]

3.1 Treatments

Our experimental design involves five treatments that correspond to discrete and parameterized versions of the five games described in the previous section, with the parameter e set to 15 experimental currency units (ECUs) – see Figure 1 and Table 1 for details.[8] In treatments IG, GC, VC, and DC the FM has to decide between option (a) and option (b) whereas in treatment CC only option (b) is available. In all treatments option (b) gives the SM the task of deciding whether she wants to transfer some, all, or none of 60 ECUs – in steps of five – to the FM, and keep the remainder. That is, in all treatments the task of the SM is to choose any number in {0, 5, 10, ... , 50, 55, 60}.

[7] Our experimental design is inspired by the (almost) double blind procedures employed by Hoffman et al. (1994), Cox (2004) and Cox and Sadiraj (2012). See List (2007) for a discussion on experimenter demand effects and Hoffmann et al. (1994), Andreoni and Petrie (2004), Cox and Deck (2005), and Andreoni and Bernheim (2009) for experimental evidence indicating that audience effects can have a significant impact on subjects' behavior.

[8] The instructions stated that the exchange rate was 5 ECUs (called Taler in the experiment) to 1 Euro.

3.2 Procedures

Ten experimental sessions were conducted at the University of Innsbruck from November 2011 to February 2013. Forty subjects who had not participated in similar experiments before were invited to each of the ten sessions using the ORSEE recruiting system (Greiner 2004). Since not all subjects showed up in time, 390 (instead of the invited 400) subjects from various academic backgrounds participated in total, and each subject participated in one treatment only.

After arrival, subjects assembled in one of the two laboratories and individually drew a sealed envelope containing a card with a number, henceforth called the private code. Then instructions were distributed and read aloud. Instructions informed subjects (i) that there are two roles in the experiment, the role of a "Group A member" and the role of a "Group B member"; (ii) that there is exactly the same number of Group A members and Group B members in the experiment and that roles are assigned randomly; (iii) that each Group A member is matched with exactly one Group B member and vice versa, and that at no point in time will a participant discover the identity of the person she or he is matched with; (iv) that Group A members are called to make a single decision that affects not only their own earnings from the experiment but also the earnings of the Group B member they are matched with (the instructions for treatment CC differ slightly in this and the next point – see Appendix E for details); (v) that Group B members have also a single decision to make and that the fact whether the decision of a Group B member is payoff-relevant or not depends on the choice of the corresponding Group A member;[9] and (vi) that cash payments could be collected a few days after the experiment from one of the secretaries who handles also the cash payments for other experiments (to ensure that the amount a subject earns cannot be linked to her or his decisions).

After reading the instructions and answering questions from the subjects in private, subjects were asked to open in private their envelopes with the private codes. Subjects whose private codes ended with an even number were assigned to Group A and asked to stay in the same room. Subjects with odd codes assumed the role of a Group B member and were escorted to the adjacent laboratory.

---

[9] That is, we employed the strategy method – i.e., Group A and Group B members made their decisions simultaneously.

In both rooms subjects were seated at widely separated tables with sliding walls. Group A and Group B members were each given a decision sheet and an empty envelope and they were asked to fill out the decision sheet in private.[10] After the subjects in both rooms made their decisions, they wrote their private code on the decision sheet and put the decision sheet into the unmarked envelope. Envelopes were then collected with a letterbox by an experimenter. Before leaving the room subjects were asked to fill out a questionnaire. Anonymous cash payments started a few days later – giving experimenters the opportunity to manually match Group A with Group B members in the meantime. Participants presented the cards with their private codes to an administrative staff person who did not know who did what for which purpose, nor how cash payments were generated, and they got their earnings in exchange. The fact that cash payments would be made that way was clearly indicated in the subject instructions. On average, subjects earned approximately 12.6 Euros (including a show up fee of 5 Euros and a flat fee of 2.5 Euros for filling out the questionnaire).[11]

<Insert Table 2 about here.>

4 Hypotheses

Appendix B presents a detailed discussion of our identification strategy and the hypotheses that can be tested with our experimental design. Here, we present a non-technical version of the main hypotheses and a short discussion on how we will test them.

In all treatments subjects in the role of SMs face exactly the same choice set, they can transfer to an anonymously paired person in the FM role any number in {0, 5, 10, …, 50, 55, 60}, and keep the remainder of the 60 ECUs. Let r stand for the chosen transfer from a SM to the paired FM. Under the assumption that all subjects are exclusively interested in their own material income we should observe $r = 0$ across all subjects in all treatments. By contrast, when the choices of SMs are shaped by some kind of other-regarding preferences

[10] In CC, Group A members have no decision to make. In this treatment we asked them to answer a question which did not affect the payments of Group A members or Group B members.

[11] Since subjects in our experiments receive the variable part of their earnings from a single game and since we expected the payoffs from that game to differ substantially across subjects (the lowest possible payoff is 0 ECUs and the highest payoff is 60 ECUs for players in both roles in all our treatments) we decided for a relatively high fixed income to ensure that lowest earnings remained in a reasonable range.

then r might differ across subjects and treatments. Let r_G denote the average transfer from the SM to the FM in game $G \in \{IG, DC, GC, VC, CC\}$.

Based on previous findings – by Offerman (2002), Charness (2004) and Cox (2004), for instance – we would expect that both unconditional and conditional other-regarding preferences play a role for SM behavior in the IG. We state this as:

Hypothesis 1 (The Impact of Unconditional Other-Regarding Preferences and the Joint Impact of Conditional Other-Regarding Preferences): *Both unconditional and conditional other-regarding preferences play a role for the behavior of second movers in the investment game.*

In line with the identification strategies employed by Charness (2004) and Cox (2004) our statistical tests for this hypothesis are based on the following considerations: In CC, FMs do not have a decision to make. Thus, nothing is revealed in this game by the choice of the FM. So, arguably, SMs cannot be motivated by vulnerability-responsiveness, deal-responsiveness, or gift-responsiveness in this game and any transfer is arguably due to unconditional other-regarding preferences. Following this line of reasoning we interpret r_{CC} as the average transfer resulting from unconditional other-regarding preferences in a situation where one person can transfer to an anonymously paired second person any number in {0, 5, 10, …, 50, 55, 60}. From observing that r_{CC} is significantly larger than zero we would then conclude that unconditional other-regarding preferences are important for SM behavior, while the absence of a difference between r_{CC} and zero would lead us to conclude that unconditional other-regarding preferences are unimportant.

In IG, the choice of option (b) by the FM reveals willingness to accept vulnerability (V), willingness to create the possibility for a mutual improvement (or "deal", D), and willingness to allow the SM to obtain an improvement (or 'gift', G) – see Table 2 for details. Thus, in IG SMs can be motivated to return positive amounts by unconditional other-regarding preferences, by vulnerability-responsiveness, by deal-responsiveness, and by gift-responsiveness. Since SMs who are asked to make a choice in IG face exactly the same opportunity set as 'SMs' in CC, but – in contrast to CC – know that the FM has chosen the opportunity set, conclusions about whether back-transfers in IG are motivated by conditional other-regarding preferences (i.e., jointly by vulnerability-responsiveness, deal-responsiveness and gift-responsiveness) are based on differences between IG and CC in the amounts of money returned by SMs to FMs. From observing that r_{IG} is significantly larger than r_{CC} we would conclude that there is a joint impact of conditional other-regarding preferences, while the absence of a significant difference would lead us to conclude that there is no such impact.

Our main empirical contribution is the isolation of the relative importance of different conditional other-regarding preferences for SM behavior. An important hypothesis in this regard is:

Hypothesis 2 (Impact of Vulnerability-Responsiveness): *Vulnerability-responsiveness is an important driver for second-mover behavior in the investment game and beyond.*

Our main empirical test for the relevance of vulnerability-responsiveness for SM behavior in the IG is based on the comparison of return transfers between IG and VC. This comparison addresses the question how SM behavior changes if – starting from the IG – we turn off vulnerability-responsiveness as a possible driver for SM behavior. To see this, note that in VC the choice of option (b) by the FM – while creating the possibility for a mutual improvement and allowing the SM to obtain an improvement – does not make the FM vulnerable. Since this latter property is the only difference (regarding revealed intentions) from IG (see Figure 1 and Tables 1 and 2), the observation that r_{IG} is significantly higher than r_{VC} would lead us to the conclusion that vulnerability-responsiveness is an important driver of SM behavior in the IG.

Comparison of IG with VC identifies the incremental impact of vulnerability responsiveness in the presence of deal-responsiveness and gift-responsiveness. From observing that r_{IG} is significantly larger than r_{VC} we would therefore conclude that vulnerability-responsiveness has an impact on SM behavior in the presence of deal-responsiveness and gift-responsiveness. A second comparison that is relevant for the question whether vulnerability-responsiveness is an important driver for SM behavior in two-stage games is the comparison between GC and CC. This comparison identifies the incremental impact of vulnerability-responsiveness in the absence of deal-responsiveness and gift-responsiveness: In GC, the choice of option (b) by the FM – while making the FM vulnerable – does neither create the possibility for a mutual improvement nor allow the SM to obtain an unilateral improvement. Thus, SMs in GC can be motivated to return positive amounts by unconditional other-regarding preferences and by vulnerability-responsiveness while “SMs” in CC are arguably only motivated by unconditional other-regarding preferences. From observing that r_{GC} is significantly larger than r_{CC} we would therefore conclude that vulnerability-responsiveness has an impact on SM behavior in the absence of deal-responsiveness and gift-responsiveness.

Vulnerability-responsiveness is unlikely to be the only conditional other-regarding preference shaping SM behavior. We therefore hypothesize:

Hypothesis 3 (Joint Impact of Deal-Responsiveness and Gift-Responsiveness): *Deal-responsiveness and gift-responsiveness together are important drivers for second-mover behavior in the investment game and beyond.*

Our main test for the joint impact of deal-responsiveness and gift-responsiveness for SM behavior in the IG is based on the comparison of back transfers between IG and GC. This comparison addresses the question how SM behavior changes if – starting from the IG – we turn off deal-responsiveness and gift-responsiveness as possible drivers for SM behavior. This follows from the observation that in GC, the choice of option (b) by the FM – while making the FM vulnerable – does neither create the possibility for a mutual improvement nor allow the SM to obtain an improvement. Since the joint impact of deal-responsiveness and gift-responsiveness is the only difference (regarding revealed intentions) to IG, the observation that r_{IG} is significantly higher than r_{GC} would lead us to conclude that deal-responsiveness and gift-responsiveness together are important drivers for SM behavior in the IG.

The comparison IG vs. GC identifies the joint impact of deal-responsiveness and gift-responsiveness in the presence of vulnerability-responsiveness. A second comparison that is relevant for the question whether deal-responsiveness and gift-responsiveness together are important drivers for SM behavior in two-stage games is the comparison between VC and CC. This comparison identifies the joint impact of the two motives in the absence of vulnerability-responsiveness: In VC, the choice of option (b) by the FM – while creating the possibility for a mutual improvement and allowing the SM to obtain an improvement – does not make the FM vulnerable. Thus, in VC the SMs can be motivated to return positive amounts by unconditional other-regarding preferences, by deal-responsiveness and by gift-responsiveness, while "SMs" in CC are arguably only motivated by unconditional other-regarding preferences. We therefore address the question of whether there is a joint impact of deal-responsiveness and gift-responsiveness in the absence of vulnerability-responsiveness by comparing r_{VC} with r_{CC}.

If we find support for Hypothesis 3 we still do not know whether deal-responsiveness, gift-responsiveness, or any combination of these motives is relevant for the finding. We therefore also test the impact of each of the two motives in isolation:

Hypothesis 4 (Impact of Deal-Responsiveness): *Deal-responsiveness is an important driver for second-mover behavior in the investment game.*

Hypothesis 4 is tested by comparing SM behavior between IG and DC. In DC, the choice of option (b) by the FM – while allowing the SM to obtain an improvement and making the FM vulnerable – does not create the possibility for a mutual improvement. Since this latter property is the only difference (regarding revealed intentions) to IG, from observing that r_{IG} is significantly higher than r_{DC} we would conclude deal-responsiveness is an important driver of SM behavior in the IG.

Hypothesis 5 (Impact of Gift-Responsiveness): *Gift-responsiveness is an important driver for second-mover behavior in the investment game.*

Hypothesis 5 is tested by comparing SM behavior between DC and GC.[12] In GC, the choice of option (b) by the FM – while making the FM vulnerable – does neither create the possibility for a mutual improvement nor allow the SM to obtain an improvement. Since the improvement for the SM is the only difference (regarding revealed intentions) between DC and GC, from observing that r_{DC} is significantly higher than r_{GC} we would conclude that gift-responsiveness has an incremental impact in the presence of vulnerability-

[12] While the main tests for hypotheses 1-4 all start form the IG and turn off one or more motives that are potentially relevant for SM behavior, this is not feasible here because by turning off the gift channel (which implies an improvement for the SM) we necessarily also turn off the joint improvement channel.

responsiveness (and absence of deal-responsiveness) and is therefore likely to be an important driver of SM behavior in the IG.

5. Experimental Results and Discussion

5.1 Experimental Results

Figure 2 shows the relative frequencies of transfers from SMs to FMs in the five treatments and reveals some notable differences between treatments. While about half of the SMs transfer zero in treatments IG, DC and GC (the exact fractions are 50.00% in IG, 46.15% in DC, and 48.72% in GC), more than 70% do so in VC and more than 65% do so in CC (the exact fractions are 71.79% for VC and 65.79% for CC). The behavior of those subjects is consistent with the hypothesis that the behavior of SMs is exclusively driven by material self-interest, whereas the behavior of the rest of the subjects is not in line with that hypothesis. At the other extreme, about 30% of SMs in treatments IG, DC and GC transfer at least 30 ECUs to the paired FM (the exact fractions are 27.50% in IG, 30.77% in DC and 33.33% in GC), while the respective fractions are 15.38% in VC and only 10.53% in CC.[13]

The pronounced differences in the tails of the distribution of transfers from SMs to FMs between treatments IG, DC and GC on the one hand, and treatments VC and CC on the other hand, also manifest themselves in the mean transfers from SMs to FMs. Those transfers are reported in the diagonal of Table 3, while the cells below the diagonal record the results of statistical tests of the pairwise comparisons between treatments – the results of the t-test are recorded in parentheses, while the results of the Mann-Whitney U-test (referred to as MWU-test below) are recorded in brackets. As can be seen in the diagonal the mean amount transferred by SMs is above 12 ECUs in treatments IG, DC and GC (specifically, it is 12.13 ECUs in IG, 12.69 ECUs in DC, and 12.82 ECUs in GC), but below 7 ECUs in treatments VC and CC (specifically, 6.67 ECUs in VC and 6.18 ECUs in CC). For our first result we use the data from treatment CC and compare it to that in treatment IG. This comparison yields:

[13] Here note that a transfer of 30 ECUs implies an equal split of the 60 ECUs between the two players. Also note that no SM returned more than 45 ECUs to a FM in any of the treatments.

Result 1 (The Impact of Unconditional Other-Regarding Preferences and the Joint Impact of Conditional Other-Regarding Preferences): *In line with Hypothesis 1 unconditional other-regarding preferences play a role for the behavior of second movers in the investment game, but such preferences alone are insufficient to explain behavior.*

The discussion in Section 4 suggests testing for the relevance of unconditional other-regarding preference**s** for behavior of SMs in the IG by analyzing transfers in CC. In the diagonal of Table 3 we see that the mean amount transferred by "SMs" in CC is 6.18 ECUs, which is significantly greater than zero according to the one-tailed t-test (p=0.000). This confirms the behavioral relevance of unconditional (i.e. altruistic) other-regarding preferences established in previous work (by Cox 2004, for instance).

<Insert Figure 2 and Table 3 about here>

To assess the joint impact of conditional other-regarding preferences (i.e., the joint impact of vulnerability-responsiveness, deal-responsiveness and gift-responsiveness) we compare the back-transfers in IG with transfers in CC. In the diagonal of Table 3 we see that the mean amount transferred by SMs is 12.13 ECUs in IG, but only 6.18 ECUs in CC, suggesting that the back-transfers in IG are significantly higher than the transfers in the CC. Significance is confirmed by the two-tailed t-test (p=0.042) and the MWU-test (p=0.067). Thus, both tests support the hypothesis that there is a joint impact of conditional other-regarding concerns. This confirms the behavioral relevance of conditional other-regarding preferences established in previous work (by Charness 1996, Offerman 2002, or Cox 2004, for instance).

Result 2 (Impact of Vulnerability-Responsiveness): *In line with Hypothesis 2, vulnerability-responsiveness is an important driver for second-mover behavior independently of whether vulnerability of the first mover comes together with a gift to the second mover or not – or allows for a Pareto improvement or not.*

The discussion in Section 4 suggests testing for the behavioral relevance of vulnerability-responsiveness in the IG by comparing back-transfers between IG and VC. The diagonal of Table 3 reveals that the mean transfer is 12.13 in IG, but only 6.67 in VC.

Applying our tests, we find that the transfers are significantly higher in IG than in VC according to the two-tailed t-test (p=0.075) and the MWU-test (p=0.050). Hence, both tests support the conclusion that vulnerability-responsiveness is an important driver for SM behavior in the IG.

While the comparison IG vs. VC assesses the impact of vulnerability-responsiveness in the presence of deal-responsiveness and gift-responsiveness, the comparison GC vs. CC assesses the impact of vulnerability-responsiveness in the absence of these two motives. The diagonal of Table 3 reports that the mean amount transferred by SMs is 12.82 ECUs in GC, but only 6.18 ECUs in CC, suggesting that the mean amount transferred is significantly greater in GC than in CC. This is confirmed by the two-tailed t-test (p=0.021) and the MWU-test (p=0.043). Thus, both tests support the hypothesis that there is an impact of vulnerability-responsiveness even in the absence deal-responsiveness and gift-responsiveness.

Result 3 (Joint Impact of Deal-Responsiveness and Gift-Responsiveness): *In sharp contrast to Hypothesis 3 there seems to be neither a joint impact of deal-responsiveness and gift-responsiveness on second-mover behavior in the presence of vulnerability-responsiveness, nor a joint impact of these motives in the absence of vulnerability-responsiveness.*

To test for the joint impact of deal-responsiveness and gift-responsiveness in the IG we compare the back-transfers in IG and in GC. The diagonal of Table 3 reports that the mean amount transferred by SMs is 12.13 ECUs in IG and 12.82 ECUs in GC. Not surprisingly, there is no significant difference according to the two-tailed t-test (p=0.829) and the MWU-test (p=0.853). Hence, the tests do not support the conclusion that deal-responsiveness and gift-responsiveness have a joint impact on SM behavior in the IG.

While the comparison of IG with GC assesses the joint impact of deal-responsiveness and gift-responsiveness in the presence of vulnerability-responsiveness, the comparison of VC with CC assesses the joint impact of these two motives in the absence of vulnerability-responsiveness. The diagonal of Table 3 reports that the mean amount transferred by SMs is 6.67 ECUs in VC and 6.18 ECUs in CC, suggesting no significant difference. This is confirmed by the two-tailed t-test (p=0.852) and the MWU-test (p=0.770). Hence, both tests confirm that there is no joint impact of deal-responsiveness and gift-responsiveness in absence of vulnerability-responsiveness. Result 3 already suggests the following two results:

Result 4 (Impact of Deal-Responsiveness): *Deal-responsiveness seems not to be an important driver for second-mover behavior in the investment game.*

Hypothesis 4 is tested by comparing SM behavior between IG and DC. The diagonal of Table 3 reports that the mean amount returned by SMs is 12.13 ECUs in IG and 12.69 ECUs in DC, suggesting that there is no significant difference in back-transfers between these two treatments. This is confirmed by the two-tailed t-test (p=0.858) and the MWU-test (p=0.846). Hence, the tests do not support the hypothesis that deal-responsiveness is an important driver for the behavior of SMs in the IG.

Result 5 (Impact of Gift-Responsiveness): *Gift-responsiveness seems not to be an important driver for second-mover behavior in the investment game.*

Hypothesis 5 is tested by comparing SM behavior between DC and GC. The diagonal of Table 3 reports that the mean amount returned by SMs was 12.69 ECUs in DC and 12.82 ECUs in GC, suggesting that there is no significant difference in back-transfers between these two treatments. This is confirmed by the two-tailed t-test (p=0.967) and the MWU-test (p=0.983). Hence, the tests do not support the hypothesis that gift-responsiveness is an important driver for the behavior of SMs in the IG.

5.2 Discussion

Taken together our experimental results suggest that – apart from unconditional other-regarding preferences – SMs in the IG react mainly to the vulnerability of the FM, not to the fact that the FM transfer allows for a mutual improvement or implies a gift to the SM. One might suspect that this finding is at least in part driven by our efficiency notion. According to that notion opportunity set J is more efficient than opportunity set K (written as J **MET** K) if: (a) the maximal feasible sum of payoffs to the two players in set J is larger than the maximal feasible sum in set K; and (b) the set J contains an allocation that is better for both players (in own-money terms) than the allocation that is realized in set K when the SM acts selfishly in own-money terms. Given this definition opportunity sets (a) and (b) cannot be ranked with respect to the efficiency criterion in treatment GC although sending the endowment reduces the size of the cake to be distributed.

There are sensible alternatives to our efficiency definition – for instance, requirement (b) in MET could simply be eliminated yielding an efficiency notion exclusively based on the (maximal feasible) cake size. How would our results change if we replaced MET by such an efficiency notion? The answer is that our main results are robust against such a change in the definition of efficiency. To see this, note that if cake-size concerns were important for back transfers, then we should see differences in the amounts of money returned by SMs to FMs between VC and CC, between IG and DC, and between DC and GC (see Appendix D for details). In the diagonal of Table 3 we see that the mean amount transferred by the SM is 6.67 ECUs in VC and 6.18 ECUs in CC, 12.13 ECUs in IG and 12.69 ECUs in DC, and 12.69 ECUs in DC and 12.82 ECUs in GC. None of these differences is significant at any reasonable level. We therefore conclude that applying a cake-size notion of efficiency does not change any of our findings: we still do not find any evidence in our data that efficiency concerns or gift-considerations are important for SM behavior; and we still find that vulnerability-responsiveness is an important driver for SM behavior.

An open question is whether adding a given motive (or a given combination of motives) has the same effect independently of the presence or absence of other motives. While our identification approach does not assume this kind of additivity of motives, our results seem to suggest that motives are indeed additive. For instance, SM behavior in VC is qualitatively similar to SM behavior in CC, and SM behavior in IG is qualitatively similar to that in GC, suggesting that adding vulnerability-responsiveness in the presence of deal-responsiveness and gift-responsiveness has the same effect as adding vulnerability-responsiveness in the absence of these motives. We would not put that much emphasis on this finding, though – our experiments have not been designed to test the additivity issue and our results in that regard are at best preliminary.[14]

[14] Engler et al. (2016) investigate this issue more systematically in experiments aimed at estimating the utility function of SMs in two-player two-stage games of perfect information using econometric techniques. Preliminary results suggest that an additive utility function does quite a good job in fitting the data regarding SM behavior.

Our results might seem to support the conclusion that positive reciprocity – as formally defined by Cox et al. (2008) – is not an important driver for SM behavior in the investment game and for related games. This conclusion would be premature, however. To see this recall that Cox et al. (2008) formulate positive reciprocity (in a two-player two-stage game of perfect information) as the assertion that a more generous choice by the FM elicits more altruistic preferences in the SM. According to their generosity definition opportunity set *J* is more generous than opportunity set *K* (written as *J* **MGT** *K*) if (a) the maximum feasible income the SM can get in set *J* (weakly) exceeds the maximum feasible income she can realize in set *K*; and (b) the difference in the maximum feasible income of the SM between sets *J* and *K* (weakly) exceeds the difference in the maximum feasible income of the FM across these sets. That is, MGT adds to our MBT the requirement that generosity is revealed by the FM's choice only if – with his choice – he does not increase his own potential income more than the SM's potential income.

According to MGT the choice of opportunity set (b) reveals generosity in IG, but opportunity sets (a) and (b) are not MGT-ranked in games DC, GC and VC. Thus, the reciprocity axiom by Cox et al. (2008) only makes a prediction for the comparison IG vs. CC, but remains agnostic regarding the other comparisons we perform in Subsection 5.1. We therefore conclude that unconditional other-regarding preferences plus reciprocity à la Cox et al. (2008) alone are insufficient to explain SM behavior in our games. By contrast, unconditional other-regarding preferences plus vulnerability-responsiveness as defined here can potentially explain SM behavior in all the games we consider.

One might ask whether the concept of betrayal aversion (as defined by Bohnet and Zeckhauser 2004, and Bohnet et al. 2008) – or some variant of it – can help to explain our data. Loosely speaking betrayal aversion refers to the hypothesis that people are less willing to take risk if the source of the risk is the behavior of another person than in the alternative situation where the source of the risk is just nature – possibly, because a bad outcome caused by selfish human behavior produces a disutility that goes beyond the disutility of the bad material outcome itself. Bohnet and Zeckhauser (2004) isolate the impact of betrayal aversion on trusting behavior in the context of the IG and find that it is important for FM behavior in the IG (almost equally important as risk aversion).[15] Although betrayal aversion as defined by Bohnet and coauthors is a hypothesis about FM motivation and behavior, one might argue that it is closely related to vulnerability-responsiveness (a hypothesis about SM motivation): If a SM in the IG anticipates that the FM experiences an extra disutility if he receives a low back transfer in a situation where he has exposed himself to social risk (he has made himself "vulnerable") and if she feels guilty for this extra disutility, then she might be inclined to reward actions that entail vulnerability to avoid these negative feelings. This is kind of second-degree betrayal-aversion story (where a SM increases the back transfer to avoid feeling guilty for the negative feelings of the FM) and it might help to explain our finding that vulnerability-responsiveness is an important driver for SM behavior.

6 Identifying Trust Acts, Trustworthy Acts, and Trust-Responsive Motivation

Prompted by our empirical findings we now provide formal definitions of trust-related concepts. For this purpose we first define a partial ordering over opportunity sets in the own-money, other's-money space – the "More Trusting Than" relation. A key ingredient of this definition is the vulnerability of the FM. Based on this definition we then provide behavioral definitions of trust act, trustworthy act and trust-responsive motivation. An important difference from existing definitions is that ours allow trust to be present without generosity. We consider this as important because in many interesting real world examples where we would consider it as natural to speak of "trusting behavior" there is no generosity involved in trust acts. Also note that definitions of trust that require that a gift is involved in a trust act imply that trust-responsiveness is a special case of gift-responsiveness. By contrast, according to our definitions trust-responsiveness is a motivation in its own that can come – but does not need to come – together with gift-responsiveness. To present our formal definitions of trust-related concepts we first need to introduce some notation.

6.1 Notation

Let $o = (f, s)$ denote an income allocation that gives material payoff f to the FM and material payoff s to the SM and consider a two-stage game in which

[15] Bohnet et al. (2008) investigate the impact of betrayal aversion in six different countries (Brazil, Oman, China, Switzerland, Turkey, and the U.S.) and find evidence for it in all of them.

- the FM is asked to choose an opportunity set O consisting of income allocations $o = (f, s)$ out of some collection of opportunity sets $\boldsymbol{O}$ which is common knowledge;
- the SM observes the opportunity set O chosen by the FM, acquires preferences P_O and then chooses a payoff vector $(f, s) \in O$.

For a given opportunity set $O \in \boldsymbol{O}$ let

- f_O^* be the maximal feasible income of the FM in the opportunity set O; that is, $f_O^* = \sup\{f \geq 0 : \exists s \geq 0 \text{ such that } (f, s) \in O\}$;
- s_O^* be the maximal feasible income of the SM in the opportunity set O; that is, $s_O^* = \sup\{s \geq 0 : \exists f \geq 0 \text{ such that } (f, s) \in O\}$;
- f_O^{eg} be the maximal feasible income of the FM in the opportunity set O, given that the SM acts selfishly in own money terms; that is, $f_O^{eg} = \sup\{f : (f, s_O^*) \in O$;
- c_O^* be the maximal feasible total cake size in the opportunity set O; that is, $c_O^* = \sup\{f + s : (f, s) \in O\}$.

6.2 Revealed Trust

We next introduce a partial ordering over opportunity sets – the “More Trusting Than” (MTT) relation – and then use it in our definition of a trust act:

Definition 1 *Consider two opportunity sets J and K in the collection of opportunity sets $\boldsymbol{O}$. We say that opportunity set $J \subset R_+^2$ is more trusting than (**"More Trusting Than", MTT**) opportunity set $K \subset R_+^2$ iff $f_J^{eg} < f_K^{eg}$ and $f_J^* > f_K^{eg}$.*

The MTT relation imposes two requirements. The first condition says that for opportunity set J to be more trusting than some other opportunity set K, the maximum income the FM can get in set J given the SM acts selfishly in own-money terms falls short of his maximum income in set K under the same condition. This requirement corresponds to the "more vulnerable than" (MVT) relation introduced in Section 2. The MTT relation imposes in addition the requirement that set J must contain an allocation that is better for the FM (in own-money terms) than the maximum income he can get in set K given the SM acts selfishly (in own-money terms). Although the second requirement is fulfilled in the investment game and some variant of it is imposed in most existing definitions of trust acts, it is debatable whether it is really necessary for a sensible definition of that concept. Consider, for instance, a (male) decision maker who is asked by a (female) fellow to lend her some money. It is feasible that the decision maker agrees and as a consequence he makes himself vulnerable (i.e., the decision maker bears the risk that he will not get back the money) without having the prospect of an improvement in material terms. However, here the prospect of improvement is arguably present in utility terms – for instance, the decision maker might experience a kind of warm-glow utility from helping his friend. Since our aim is to provide a behavioral definition of a trust based exclusively on observable variables asking for the prospect of an improvement in utility terms is not feasible. Definition 1 therefore insists on the prospect of a material improvement.

While the two inequalities in Definition 1 seem necessary for characterizing trust acts, one might argue that they are not sufficient. Fehr (2009, p.238), for instance, writes: "An individual (let's call her the trustor or investor) trusts if she voluntarily places resources at the disposal of another party (the trustee) without any legal commitment from the latter. In addition, the act of trust is associated with an expectation that the act will pay off in terms of the investor's goals." This definition adds to the two conditions in Definition 1 the requirement that a trust act by the FM must allow for an improvement for the SM (because placing resources at the disposal of the SM allows the latter to improve in material terms according to our definitions in Section 2). While this requirement has appeal, adding it has the consequence that trust becomes a special case of generosity as defined by Cox et al. (2008) – in the sense, that we can have generosity without trust but not vice versa. Also, there are many interesting real world examples of behavior where we would consider it natural to speak of "trusting behavior" although there is no gift involved. Consider, for instance, a hungry escaped prisoner who rings at the door of a secluded house to ask for a piece of bread. The escaped prisoner trusts that the landlord will help him and will not call the police. If his trust is fulfilled he is better off, if it is violated he is worse off. However, there is no gift involved here. The landlord is neither better off if she fulfills nor if she violates the trust – arguably, she would have been better off if the escaped prisoner had trusted someone else! In sum, we think that there is room for a plausible definition of trusting behavior that does not involve generosity. We next present such a definition.

Definition 2 *Let the FM choose the actual opportunity set for the SM from the collection of opportunity sets $\boldsymbol{O}$. If the opportunity sets J and K are both available (that is, if $J, K \in O$) and J is MTT K then we say that the FM's choice of J is a* ***trust act*** *(or trusting behavior).*

This definition of a trust act is an identification of behavior not motivation for that behavior. As explained by Cox (2004), a trust act can be motivated by FM's altruistic preferences or by FM's expectation that SM will reciprocate trust acts. Existing evidence provides support for the empirical significance of both motivations of FM behavior in the investment game (Cox, 2004), the moonlighting game (Cox, Sadiraj and Sadiraj, 2008) and the trust game (Cox and Deck, 2005).

We now turn our attention to SM behavior. Given a collection of opportunity sets $\boldsymbol{O}$ and given that the FM's actual choice $J \in \boldsymbol{O}$ is a trust act, we define $\boldsymbol{O}^J \subseteq \boldsymbol{O}$ as that subset of the original collection of opportunity sets that contains only those elements of $\boldsymbol{O}$ that are **MTT**-ordered with respect to J. By the definition of a trust act $\boldsymbol{O}^J$ is non-empty and we denote the least trusting element in $\boldsymbol{O}^J$ by L^J. Using this notation we now define $f_J^{safe} = f_{L^J}^{eg} = \sup\{f : (f, s_O^*) \in L^J\}$ as "the safety payoff for the FM".

Definition 3 *Suppose the FM chooses the opportunity set $J \in \boldsymbol{O}$ from the collection of opportunity sets $\boldsymbol{O}$ and that the choice of J is a trust act. Then the behavior of the SM is a* ***trustworthy act*** *iff she chooses an element $(f^c, s^c) \in J$ such that $f^c \geq f^{safe}$.*

Having characterized trust act by FM and trustworthy act by SM we now define trust-responsive *motivation* by SM.

Definition 4 *Let the FM choose the actual opportunity set for the SM from the collection of opportunity sets $\boldsymbol{O}$. Suppose the FM chooses $J \in \boldsymbol{O}$. If the choice of J is a trust act and if this fact elicits more altruistic preferences in the SM compared to a situation where only J is available we say that the SM is motivated by* ***trust-responsiveness.***

7 Conclusion

This paper has used a five-games design to decompose the drivers for trustworthiness. Our diagnostic games have isolated the impact of various combinations of the following motives for back-transfers in the investment game. First, trustees have unconditional other-regarding preferences – like altruism, inequality aversion, quasi-maximin, etc. Second, trustees are motivated by vulnerability-responsiveness – they reward actions that potentially lead to a decrease in trustors' payoffs. Thirdly, trustees are motivated by deal-responsiveness – they reward actions that have the potential to increase the payoffs of both parties. And fourthly, trustees are motivated by gift-responsiveness – they reward actions that potentially increase their own payoffs. Our results have shown that – besides unconditional other-regarding preferences – vulnerability-responsiveness is an important determinant of trustworthiness even in cases where the vulnerability of the trustor does not come together with a gift to the trustee. Motivated by our experimental findings we have provided definitions of trust acts, trustworthy acts and trust-responsive motivation based on revealed willingness to accept vulnerability and the response to it. An important difference from well-established definitions is that ours allows trust to come without generosity. This seems more in line with the everyday usage of that term and it gives room for a role for trust and trust-responsiveness as interesting concepts on their own (independent of generosity).

References

Andreoni, J., Bernheim, D., 2009. Social Image and the 50–50 Norm: A Theoretical and Experimental Analysis of Audience Effects. Econometrica. 77, 1607-1636.

Andreoni, J., Petrie, R., 2004. Public Goods Experiments Without Confidentiality: A Glimpse into Fund-Raising. Journal of Public Economics. 88, 1605-1623.

Bacharach, M., Guerra, G, Zizzo, D. J., 2007. The Self-Fulfilling Property of Trust: An Experimental Study. Theory and Decision. 63, 349-388.

Ben-Ner, A., Putterman, L., 2009. Trust, communication and contracts: An experiment. Journal of Economic Behavior & Organization. 70, 106-121.

Berg, J., Dickhaut, J., McCabe, K. A., 1995. Trust, Reciprocity, and Social History. Games and Economic Behavior. 10, 122-42.

Bohnet, I., Zeckhauser, R., 2004. Trust, Risk and Betrayal. Journal of Economic Behavior and Organization. 55, 467-484.

Bohnet, I., Greig, F., Herrmann, B., Zeckhauser, R., 2008. Betrayal Aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States. American Economic Review. 98, 294-310.

Burks, S. V., Carpenter, J. P., Verhoogen, E., 2003. Playing both roles in the trust game. Journal of Economic Behavior & Organization. 51, 195-216.

Charness, G., 1996. Attribution and reciprocity in a simulated labor market: An experimental investigation. Economics Working Paper 283, Department of Economics and Business, Universitat Pompeu Fabra, revised Oct 1997.

Charness, G., 2004. Attribution and reciprocity in a simulated labor market: An experimental investigation. Journal of Labor Economics. 22, 665-688.

Chaudhuri, A., Gangadharan, L., 2007. An Experimental Analysis of Trust and Trustworthiness. Southern Economic Journal. 73, 959-985.

Coleman, J., 1990. Foundations of Social Theory. The Belknap Press of Harvard University Press.

Colquitt, J. A., Brent, S. A., LePine, J. A., 2007. Trust, Trustworthiness, and Trust Propensity: A Meta-Analytic Test of Their Unique Relationships with Risk Taking and Job Performance. Journal of Applied Psychology. 92, 909-927.

Cox, J. C., 2004. How to Identify Trust and Reciprocity. Games and Economic Behavior. 46, 260-281.

Cox, J. C., Deck, C. A., 2005. On the Nature of Reciprocal Motives. Economic Inquiry. 43, 623-635.

Cox, J. C., Friedman, D., Sadiraj, V., 2008. Revealed Altruism. Econometrica. 76, 31-69.

Cox, J. C., Sadiraj, V., 2012. Direct Tests of Individual Preferences for Efficiency and Equity. Economic Inquiry. 50, 920-931.

Cox, J. C., Sadiraj, K., Sadiraj, V., 2008. Implications of Trust, Fear, and Reciprocity for Modeling Economic Behavior. Experimental Economics. 11, 1-24.

Croson, R., Buchan, N., 1999. Gender and Culture: International Experimental Evidence from Trust Games. American Economic Review. 89, 386-391.

Engler, Y., Kerschbamer, R., Page, L., 2016. Why Did He Do That? Using Counterfactuals to Study the Effect of Intentions in Extensive Form Games. Mimeo. School of Economics, Queensland University of Technology, Brisbane, Australia.

Fehr, E., 2009. On the Economics and Biology of Trust. Journal of the European Economic Association. 7, 235-266.

Ferrin, D. L., Dirks, K. T., 2003. The Use of Rewards to Increase and Decrease Trust: Mediating Processes and Differential Effects. Organization Science. 14, 18-31.

Greiner, B., 2004. The Online Recruiting System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics. WP Series in Economics. 10, University of Cologne.

Guiso, L., Sapienza, P., Zingales, L., 2008. Trusting the Stock Market. The Journal of Finance. 63, 2557-2600.

Ho, T. H., Weigelt, K., 2005. Trust Building among Strangers. Management Science. 51, 519-530.

Hoffman, E., McCabe, K., Shachat, K., Smith, V., 1994. Preferences, Property Rights, and Anonymity in Bargaining Games. Games and Economic Behavior. 7, 346-380.

Johnson, N. D., Mislin, A., 2011. Trust games: A meta-analysis. Journal of Economic Psychology. 32, 865-889.

Karlan, D. S., 2005. Using Experimental Economics to Measure Social Capital and Predict Financial Decisions. American Economic Review. 95, 1688-1699.

Knack, S., Keefer, P., 1997. Does Social Capital Have an Economic Payoff? A Cross-Country Investigation. The Quarterly Journal of Economics. 112, 1251-1288.

LaPorta, R., Lopez-de-Silane, F., Shleifer, A., Vishny, R., 1997. Trust in Large Organizations. American Economic Review. 87, 333-338.

Levine, D. K., 1998. Modeling Altruism and Spitefulness in Experiments. Review of Economic Dynamics. 1, 593-622.

List, J., 2007. On the Interpretation of Giving in Dictator Games. Journal of Political Economy. 115, 482-493.

Mayer, R., Davis, J., Schoorman, F., 1995. An Integrative Model of Organizational Trust. *The* Academy of Management Review. 20, 709-734.

McAllister, D. J., 1995. Affect- and Cognition-Based Trust as Foundations for Interpersonal Cooperation in Organizations. The Academy of Management Journal. 38, 24-59.

Offerman, T., 2002. Hurting hurts more than helping helps. European Economic Review. 46, 1423-1437.

Rousseau, D., Sitkin, S., Burt, R., Camerer, C., 1998. Introduction to Special Topic Forum: Not so Different after All: A Cross-Discipline View of Trust. The Academy of Management Review. 23, 393-404.

Schotter, A., Sopher, B., 2006. Trust and trustworthiness in games: An experimental study of intergenerational advice. Experimental Economics. 9, 123-145.