What is Trustworthiness and What Drives It

James Cox
Georgia State University

Rudolf Kerschbamer
University of Innsbruck

Daniel Neururer
University of Innsbruck

What is Trustworthiness and What Drives It?*

James C. Cox[†], Rudolf Kerschbamer[‡] and Daniel Neururer[§]

September 9, 2014

Abstract

This paper reports the results of experiments designed to isolate the impact of various combinations of the following motives on trustworthiness: (i) unconditional other-regarding preferences – like altruism, inequality aversion, quasi-maximin, etc.; (ii) deal-responsiveness – reacting to actions that allow for a mutual improvement by adopting behavior that implies a mutual improvement; (iii) gift-responsiveness – reacting to choices that allow the trustee to obtain an improvement by adopting actions that benefit the trustor; and (iv) vulnerability-responsiveness – reacting to the vulnerability of the trustor by adopting actions that do not hurt the trustor. Our results indicate 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 empirical findings we provide formal definitions of trust and trustworthiness based on revealed willingness to accept vulnerability and the response to it.

JEL classification: C70, C91, D63, D64

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

[†]Andrew Young School of Policy Studies, Georgia State University: <jccox@gsu.edu>
[‡]Department of Economics, University of Innsbruck: <Rudolf.Kerschbamer@uibk.ac.at>
[§]Department of Economics, University of Innsbruck: <Daniel.Neururer@uibk.ac.at>

1 Introduction

The economic 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 goal of the present paper is to provide formal definitions of trust and trustworthiness based on observable variables. The second goal is the identification of factors that drive trustworthiness in a specific class of games. As we will argue below, those two goals are intimately related to each other.

Regarding definitions, there exist many verbal ones for trust in the economic and the non-economic literature. 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 without explicitly defining it. One of the few exceptions to this is Glaeser et al. (2000) who state (on p. 3 of the 1999 working paper version of the article) that "trustworthiness is a behavior that increases the return to people who trust you."

The experimental economic literature typically defines trust and trustworthiness in terms of behavior in specific games. The most familiar example is the investment game 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.

Within the investment game a FM is said to trust if he sends more than the minimum possible amount and a SM is said to be trustworthy if she re-

---

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

turns 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).[2] However, it is not clear how these behavioral definitions of trust and trustworthiness extend to more general games where players have richer action spaces or where material payoffs have a different structure. 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 identification of drivers of trustworthiness is the other aim of this paper. It is important for our first aim because it has an immediate impact on the answer to the question what constitutes sensible definitions of trust and trustworthiness and what distinguishes those concepts from other related concepts. As an illustration, consider again the investment game. In this game a trust act by the FM (i.e. a positive transfer) is per design a generous act according to established definitions (as the one by Cox et al. 2008, for instance). If the SM interprets it that way and if the main motivation for her back-transfer is to repay the generosity of the FM by her own generosity, then there is not much room left for trust and trustworthiness as interesting concepts on their own. They are then simply context-dependent special cases of the more general concepts 'generosity' (or 'kindness') and 'retaliation' (or 'reciprocal response').

Regarding the identification of drivers for trustworthiness the pioneering work is by Cox (2004) who proposes a triadic (i.e., a three-games) design to discriminate between different motives for transfers (from the FM to the SM) and back-transfers (from the SM to the FM) in the investment game. Relevant for our goal of disentangling different drivers for trustworthiness are two of his three treatments, one is the standard investment game (IG), the other is a specially designed dictator game (DG). The crux of the Cox (2004) design is that 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

---

[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.

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. However, 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. The present paper digs deeper by isolating the impact of various combinations of the following motives: (i) unconditional other-regarding preferences – like altruism, inequality aversion, quasi-maximin, etc.; (ii) deal-responsiveness – reacting to actions that allow for a mutual improvement by adopting behavior that implies a mutual improvement; (iii) gift-responsiveness – reacting to choices that allow the trustee to obtain an improvement by adopting actions that benefit the trustor; and (iv) vulnerability-responsiveness – reacting to the vulnerability of the trustor by adopting actions that do not hurt the trustor.

For this purpose 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.

The five treatments of our experimental design correspond to the five two-player games shown in Figure 1. All of the treatments have the same non-trivial feasible set for the SM. In four of the treatments the non-trivial feasible set for the SM is reached only if the FM chooses option (b). The game-form to those treatments is displayed in Panel A of Figure 1. These four treatments 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 treatments are displayed in Table 1. The other treatment – 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 treatments –implemented in a between-subjects design – are:

- Treatment **IG** is a **standard investment game** à la Berg et al. (1995). The main difference to the 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 – she can return any amount between 0 and $4e$. Panel A of Figure 1 displays the resulting material game using the parameters and the discrete grid implemented in the experiment. Important features of the IG are that – in comparison to option (a) – option (b) (i) allows for a mutual improvement [if the SM returns more than e but less than $3e$ then both players are strictly better off]; (ii) involves a gift to the SM [the maximal material payoff she can realize with option (b) exceeds the payoff she receives in (a)]; and (iii) makes the FM vulnerable [if the SM returns less than e then the FM is worse off in material terms, compared to option (a)].

- Treatment **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.

- Treatment **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 an unilateral improvement. Still, the choice of (b) makes the FM vulnerable.

- Treatment **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 an unilateral improvement for the SM.

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

<Insert Figure 1 and Table 1 about here.>

The main results of our experiments involving 390 subjects are as follows: Consistent with the findings by Cox (2004) we observe strictly positive transfers from SMs to FMs in treatment CC and significantly higher transfers from SMs to FMs in IG than in CC.[3] Thus, unconditional other-regarding preferences seem to play a role for the behavior of SMs in the investment game, but such preferences alone are insufficient to explain behavior. Turning to the other treatments our most

---

[3] The two Cox treatments discussed above 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.

important finding is 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 empirical findings and inspired by the formal framework introduced by Cox et al. (2008) we 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.

The rest of the paper is organized as follows. Section 2 presents definitions for unconditional other-regarding preferences, deal-responsiveness, gift-responsiveness and vulnerability-responsiveness and describes the five mini-games used for discrimination between motives. 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 formal definitions of trust, trustworthiness and trust-responsiveness based on observable variables and Section 7 concludes.

2 Definitions of Motives and Identification Games

2.1 Definitions

Motivated by the formal framework introduced by Cox et al. (2008) we present in this subsection – for the class of two-player two-stage games of complete and perfect information – definitions of unconditional other-regarding preferences; revealed willingness to create the possibility of a mutual improvement and deal-responsiveness; revealed willingness to allow the DM to obtain an improvement and gift-responsiveness; and revealed willingness to accept vulnerability and vulnerability-responsiveness. For this purpose 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

- 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, $s_O^* = \sup\{s \geq 0 : \exists f \geq 0 \text{ such that } (f, s) \in O\}$;

5

- 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, $f_O^{eg} = \sup\{f : (f, s_O^*) \in O\}$; and

- c_O^* be the maximal feasible total cake size in the opportunity set O; that is, $c_O^* = \sup\{f + s : (f, s) \in O\}$.

Based on these definitions we now define three binary relations between opportunity sets – the "More Efficient Than" (MET) relation, the "More Beneficial Than' (MBT) relation and the "More Vulnerable Than" (MVT) relation.

Definition 1 *Consider two opportunity sets J and K in* **O**. *We say that*

- *opportunity set J allows for a more efficient choice than (is "**More Efficient Than**", MET) opportunity set K if $c_J^* > c_K^*$ and $\exists (f', s') \in J$ such that $f' > f_K^{eg}$ and $s' > s_K^*$;*

- *opportunity set J is more beneficial for the SM than (is "**More Beneficial Than**", MBT) opportunity set K if $s_J^* > s_K^*$;*

- *opportunity set J entails more vulnerability for the FM than (is "**More Vulnerable Than**'", MVT) opportunity set K if $f_J^{eg} < f_K^{eg}$.*

Each of those relations defines a partial ordering over opportunity sets. **MBT** and **MVT** yield only partial (and not complete) orderings because they are defined as strict inequalities. For **MET** an additional source for incompleteness comes from the fact that the relation imposes two requirements. The first condition says that for opportunity set J to be more efficient than some other opportunity set K, the maximal feasible cake size in set J must be larger than the maximal feasible cake size in set K. The second condition adds the requirement that the set J must contain 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.[4] The **MBT** relation defines opportunity set J as more beneficial for the SM than opportunity set K if the maximum material income the SM can get in set J exceeds the maximum material income she can realize in set K.[5] The **MVT** relation imposes the condition that for opportunity set J to entail more vulnerability for the FM than opportunity set K, 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.

Definition 2 *Let the FM choose the actual opportunity set for the SM from the collection of opportunity sets* **O**. *If the opportunity sets J and K are both available (that is, if $J, K \in$ **O**) and*

---

[4] Of course, one might think of alternative requirements here; for instance, the two inequalities could be replaced by the milder condition $c_J^* > f_K^{eg} + s_K^*$. There are pros and cons for each of these definitions. Since for the present context all plausible efficiency notions imply the same ordering over opportunity sets we do not discuss this issue further here.

[5] Our **MBT** definition resembles the More Generous Than (MGT) relation by Cox et al. (2008). We discuss the difference between the two in Subsection 5.2.

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

- J is **MBT** K then we say that the FM's choice of J reveals the **willingness to allow the SM to obtain an improvement**;

- J is **MVT** K then we say that the FM's choice of J reveals the **willingness to accept vulnerability**.

Borrowing again from Cox et al. (2008) we next define a partial ordering on preferences of the SM over income allocations – the "More Altruistic Than" (MAT) relation. For this purpose suppose that the SM has well-behaved preferences P over feasible income allocations and that those preferences can be represented by some smooth function $u_P(.)$ that assigns real numbers to any $o = (f, s)$ in a given domain $D \subset R_+^2$. Let $WTP_P(f, s)$ be the amount of own income a SM with preferences P is willing to give up in order to increase the FM's income by one unit at $o = (f, s)$. That is, $WTP_P(f, s) = (\frac{\partial u_P(f,s)}{\partial f} / \frac{\partial u_P(f,s)}{\partial s})$. Within this framework Cox et al. (2008) now formalize (in their Definition 1) the idea that one preference ordering is more altruistic than another. We restate their definition here as:

Definition 3 *For a given domain $D \subset R_+^2$ preference relation P on D is more altruistic than (**"More Altruistic Than", MAT**) preference relation P' on D if $WTP_P(f, s) \geq WTP_{P'}(f, s)$ for all $(f, s) \in D$.*

Based on those definitions we now define:

Definition 4 *Let the FM choose the actual opportunity set for the SM from the collection of opportunity sets **O**. Suppose the FM chooses $J \in$ **O** and suppose that this choice elicits preferences P_J in the SM.*

- *If J is the only element in **O** (so that the FM has no real choice to make) and if $WTP_{P_J}(f, s) \neq 0$ for some (f, s) then we say that the SM has **unconditional other-regarding preferences**.*

- *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 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 we say that the SM's preferences exhibit **gift-responsiveness**.*

- *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 we say that the SM's preferences exhibit **vulnerability-responsiveness**.*

2.2 Identification Games

Our experimental design involves five treatments that correspond to parameterized and discrete versions of the following five games:

IG is a **standard investment game** à la Berg et al. (1995). The main difference to the usual 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 – an allocation where both players earn their endowment e; or *(b) invest the whole endowment*, resulting in a non-trivial opportunity set for the SM – she receives the tripled transfer and has to decide which part of the augmented endowment of $4e$ she wants to return.[6] Panel A of Figure 1 and Table 1 display the material game using the parameters and the discrete grid as implemented in the experiment.

DC is the **deal-controlled investment game**: It differs from IG only in the payoff of the SM for option (a). This payoff is increased from e to $3e$, keeping all the other details of the game constant. This design feature implies that the choice of option (b) does not create the possibility of a mutual improvement in material terms – the available cake size is $4e$ if the FM decides for (a) and it is $4e$ if the FM decides for (b).

GC is the **generosity-controlled investment game**: It differs from IG only in the payoff of the SM for option (a). This payoff is increased from e to $4e$, keeping all the other details of the game constant. This design feature implies that option (b) does not allow the SM to obtain an improvement in comparison to option (a) – the highest material payoff the SM can realize under the opportunity set (b) is $4e$ which is exactly the amount she earns when the FM decides for (a).

VC is the **vulnerability-controlled investment game**: It differs from IG only in the payoff of the FM for option (a). This payoff is decreased from e to 0, keeping all the other details of the game constant. This design feature implies that the choice of option (b) by the FM does not make him more vulnerable than the choice of option (a) – the lowest material payoff he receives under the opportunity set (b) is 0 which is exactly the amount he earns when he decides for (a).

CC is the **conditional-controlled investment game**: It differs from IG only in the fact that option (a) is not a feasible choice for the FM. Thus, CC – shown in Panel B of Figure 1 – is a dictator game where the SM faces exactly the same choice set as in the other four games when the FM decides for option (b).

---

[6]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.

3   Experimental Treatments and Procedures

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]

Experimental 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, DC, GC and VC the FM has to decide between option (a) and option (b), 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 $\{0, 5, 10, ... , 50, 55, 60\}$.

Experimental 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 session 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/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 the Appendix 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

---

[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), and Andreoni and Bernheim (2009) for experimental evidence indicating that audience effects might have a significant impact on subjects' behavior in dictator-game like situations.

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

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

In both rooms subjects were seated at widely separated tables with sliding walls. Group A and Group B members were handed out 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 card with their private code 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 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

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 $\{0, 5, 10, \ldots, 50, 55, 60\}$, 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 subjects are exclusively interested in their own material income we should observe $r = 0$ across all subjects in all treatments. This will be our null hypothesis. The alternative is that the choices of SMs are shaped by some kind of other-regarding preferences. In that case r might differ across subjects

---

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

[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 and 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.

and treatments. Let r_G denote the average transfers from SMs to FMs in game $G \in \{IG, DC, GC, VC, CC\}$.

a) Identification of **Unconditional Other-Regarding Preferences** (H^U): 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 deal-responsiveness, gift-responsiveness, or vulnerability-responsiveness in this game and any transfer is arguably due to unconditional other-regarding preferences. The null hypothesis here is that unconditional other-regarding preferences are unimportant ($H_0^U : r_{CC} = 0$), with the alternative that they lead to positive transfers ($H_A^U : r_{CC} > 0$).

b) Identification of the **Joint Impact of Deal-Responsiveness, Gift-Responsiveness and Vulnerability-Responsiveness** ($H^{D,G,V}$): In IG, the choice of option (b) by the FM reveals willingness to create the possibility for a mutual improvement (or 'deal', D), willingness to allow the SM to obtain an improvement (or 'gift', G), and willingness to accept vulnerability – see Table 2 for details. Thus, in IG SMs can be motivated to return positive amounts by unconditional other-regarding preferences, by deal-responsiveness, by gift-responsiveness, and by vulnerability-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 deal-responsiveness, gift-responsiveness and vulnerability-responsiveness) are based on differences between IG and CC in the amounts of money returned by SMs to FMs. The null hypothesis here is that there is no joint impact of deal-responsiveness, gift-responsiveness and vulnerability-responsiveness ($H_0^{D,G,V} : r_{CC} \geq r_{IG}$), with the alternative that there is one ($H_A^{D,G,V} : r_{CC} < r_{IG}$).

c) Identification of the **Joint Impact of Gift-Responsiveness and Vulnerability-Responsiveness in the Absence of Deal-Responsiveness** ($H^{G,V(abs.D)}$): 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. Thus, in DC SMs can be motivated to return positive amounts by unconditional other-regarding preferences, by gift-responsiveness, and by vulnerability-responsiveness. By contrast, in CC 'SMs' are arguably only motivated by unconditional other-regarding preferences. We therefore address the question of whether there is a joint impact of gift-responsiveness and vulnerability-responsiveness (in the absence of D) by comparing the amounts of money transferred by SMs to FMs across DC and CC. The null hypothesis here is that there is no joint impact of gift-responsiveness and vulnerability-responsiveness in the absence of D ($H_0^{G,V(abs.D)} : r_{CC} \geq r_{DC}$), with the alternative that there is one ($H_A^{G,V(abs.D)} : r_{CC} < r_{DC}$).

d) Identification of the **Impact of Vulnerability-Responsiveness in the**

Absence of Deal-Responsiveness and Gift-Responsiveness ($H^{V(abs.D,G)}$):
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. We therefore address the question of whether there is an impact of vulnerability-responsiveness in the absence of deal-responsiveness and gift-responsiveness by comparing the amounts of money returned by SMs to FMs across GC and CC. The null hypothesis here is that there is no impact of vulnerability-responsiveness in the absence of D and G ($H_0^{V(abs.D,G)} : r_{CC} \geq r_{GC}$), with the alternative that there is one ($H_A^{V(abs.D,G)} : r_{CC} < r_{GC}$).

e) Identification of the **Incremental Impact of Vulnerability-Responsiveness in the Presence of Deal-Responsiveness and Gift-Responsiveness** ($H^{V(pres.D,G)}$): 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) to IG (see Figure 1 and Tables 1 and 2), conclusions about whether there is an incremental impact of vulnerability-responsiveness in the presence of D and G are based on differences between IG and VC in the amounts of money returned by SMs to FMs. The null hypothesis here is that there is no incremental impact of vulnerability-responsiveness in the presence of D and G ($H_0^{V(pres.D,G)} : r_{VC} \geq r_{IG}$), with the alternative that there is one ($H_A^{V(pres.D,G)} : r_{VC} < r_{IG}$).

f) Identification of the **Joint Impact of Deal-Responsiveness and Gift-Responsiveness in the Absence of Vulnerability-Responsiveness** ($H^{D,G(abs.V)}$): 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 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 V by comparing the amounts of money returned by SMs to FMs between VC and CC. The null hypothesis here is that there is no joint impact of deal-responsiveness and gift-responsiveness in the absence of V ($H_0^{D,G(abs.V)} : r_{CC} \geq r_{VC}$), with the alternative that there is one ($H_A^{D,G(abs.V)} : r_{CC} < r_{VC}$).

g) Identification of the **Joint Impact of Deal-Responsiveness and Gift-Responsiveness in the Presence of Vulnerability-Responsiveness** ($H^{D,G(pres.V)}$): 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 those latter properties are the only differences (regarding revealed intentions) to IG, conclusions about whether there is a joint impact of deal-responsiveness and gift-responsiveness in the presence of V are based on differences between IG and GC in the amounts of money returned by SMs to FMs. The null hypothesis here is that there is no joint impact of deal-responsiveness and gift-responsiveness in the presence of V ($H_0^{D,G(pres.V)} : r_{GC} \geq r_{IG}$), with the alternative that there is one ($H_A^{D,G(pres.V)} : r_{GC} < r_{IG}$).

h) Identification of the **Incremental Impact of Deal-Responsiveness in the Presence of Gift-Responsiveness and Vulnerability-Responsiveness** ($H^{D(pres.G,V)}$): 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, conclusions about whether there is an incremental impact of deal-responsiveness in the presence of G and V are based on differences between IG and DC in the amounts of money returned by SMs to FMs. The null hypothesis here is that there is no incremental impact of deal-responsiveness in the presence of G and V ($H_0^{D(pres.G,V)} : r_{DC} \geq r_{IG}$), with the alternative that there is one ($H_A^{D(pres.G,V)} : r_{DC} < r_{IG}$).

i) Identification of the **Incremental Impact of Gift-Responsiveness in the Presence of Vulnerability-Responsiveness but Absence of Deal-Responsiveness** ($H^{G(pres.V,abs.D)}$): 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) to DC, conclusions about whether there is an incremental impact of gift-responsiveness in the presence of V and absence of D are based on differences between GC and DC in the amounts of money returned by SMs to FMs. The null hypothesis here is that there is no incremental impact of gift-responsiveness in the presence of V and absence of D ($H_0^{G(pres.V,abs.D)} : r_{GC} \geq r_{DC}$), with the alternative that there is one ($H_A^{G(pres.V,abs.D)} : r_{GC} < r_{DC}$).

<Insert Table 3 about here.>

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

13

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 our null 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.[12] A SM returns never more than 45 ECUs to the paired FM in any of the treatments. 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 4, 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).

<Insert Figure 2 and Table 4 about here.>

Turning to the hypotheses formulated in Section 4 statistical tests yield the following results:

a) H^U suggests testing for the **relevance of unconditional other-regarding preferences** for the transfers from SMs to FMs by analyzing transfers in CC. In the diagonal of Table 4 we see that the mean amount transferred by 'SMs' 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 other-regarding preferences established in previous work (by Cox 2004, for instance).

b) $H^{D,G,V}$ suggests testing for the **joint impact of conditional other-regarding concerns** (i.e., the joint impact of deal-responsiveness, gift-responsiveness and vulnerability-responsiveness) by comparing back-transfers in IG with transfers in CC. In the diagonal of Table 4 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.[13]

---

[12] Here note that a transfer of 30 ECUs implies an equal split of the 60 ECUs between the two players.

[13] Taken together the results for H^U and $H^{D,G,V}$ confirm the findings by Cox (2004) that both unconditional and conditional other-regarding preferences play a role for the behavior of SMs in the investment game.

c) $H^{G,V(abs.D)}$ suggests testing for the **joint impact of gift-responsiveness and vulnerability-responsiveness in the absence of D** by comparing back-transfers in DC with transfers in CC. The diagonal of Table 4 reports that the mean amount transferred by SMs is 12.69 ECUs in DC, but only 6.18 ECUs in CC, suggesting that the mean amount transferred is significantly greater in DC than in CC. Significance is confirmed by the two-tailed t-test ($p = 0.021$) and the MWU-test ($p = 0.031$). Hence, both tests support the hypothesis that there is a joint impact of gift-responsiveness and vulnerability-responsiveness, even in absence of D.

d) $H^{V(abs.D,G)}$ suggests testing for the impact of **vulnerability-responsiveness in the absence of D and G** by comparing back-transfers in GC with transfers in CC. The diagonal of Table 4 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 of D and G.

e) $H^{V(pres.D,G)}$ suggests testing for the **incremental impact of vulnerability-responsiveness in the presence of D and G** by comparing back-transfers between IG and VC. The diagonal of Table 4 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 there is indeed an incremental impact of vulnerability-responsiveness in the presence of D and G.

f) $H^{D,G(abs.V)}$ suggests testing for the **joint impact of deal-responsiveness and gift-responsiveness in the absence of V** by comparing back-transfers in VC with transfers in CC. The diagonal of Table 4 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 V.

g) $H^{D,G(pres.V)}$ suggests testing for the **joint impact of deal-responsiveness and gift-responsiveness in the presence of V** by comparing back-transfers in IG and in GC. The diagonal of Table 4 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 in the presence of V.

h) $H^{D(pres.G,V)}$ suggests testing for the **incremental impact of deal-responsiveness in the presence of G and V** by comparing back-transfers in IG and in DC. The diagonal of Table 4 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 presence of G and V.

i) $H^{D(abs.G,V)}$ suggests testing for the **incremental impact of gift-responsiveness in the presence of V** (but absence of D) by comparing back-transfers in DC and in GC. The diagonal of Table 4 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 there is an incremental impact of gift-responsiveness in the presence of V.

Taken together the tests reported here support the following summary statement:

Result 1: *Unconditional other-regarding preferences seem to play a role for the behavior of SMs in the investment game, but such preferences alone are insufficient to explain behavior. Vulnerability-responsiveness seems to be an important driver for SM behavior independently of whether vulnerability of the FM comes together with a gift to the SM or not – or allows for a Pareto improvement or not. Deal-responsiveness and gift-responsiveness seem to be behaviorally relevant only in settings where allowing for a Pareto improvement or for an improvement for the SM entails vulnerability of the FM. Indeed, the binary comparisons between treatments reveal significant differences in the amounts transferred from SMs to FMs in the comparisons IG vs. VC, DC vs. CC and GC vs. CC, but no significant differences in the comparisons IG vs. DC, IG vs. GC and DC vs. GC, as well as for VC vs. CC.*

5.2 Discussion

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 (**"More Generous Than", MGT**) opportunity set K if two conditions are fulfilled: (i) $s_J^* \geq s_K^*$; and (ii) $s_J^* - s_K^* \geq f_J^* - f_K^*$. Part (i) of this relation corresponds almost exactly to our MBT condition, the only difference being that the latter is defined as a strict inequality while part (i) of the MGT relation entails a weak inequality. Part (ii) of MGT adds 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. So, our finding that the behavior of SMs in IG is qualitatively similar to that in DC and GC, but significantly different to that in VC and CC does not contradict the Cox et al. (2008) reciprocity axiom – the axiom rather correctly predicts the difference in SM behavior between IG and CC, but is agnostic about the other comparisons. One conclusion regarding reciprocity one can draw from our data is that a reciprocity notion based exclusively on the gift part of the MGT definition misses the point – that is, something like part (ii) of the MGT definition seems to be an important ingredient to a sensible definition of revealed generosity. A second conclusion worth stressing is 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.

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 D and G has the same effect as adding vulnerability-responsiveness in the absence of D and G. 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. Engler et al. (2014) 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.

6   Definitions of Trust Act, Trustworthy Act and Trust-Responsive Motivation

Prompted by the experimental results reported in Section 5 we now present formal definitions of trust act, trustworthy act and trust-responsive motivation. Our definitions are more specific than the verbal definitions typically employed in the non-economic literature and at the same time more general than the game-specific definitions typically used in the (experimental) economic literature. Based on the notation introduced in Section 2 we first 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 5 *Consider two opportunity sets J and K in* **O**. *We say that opportunity set $J \subset R_+^2$ is more trusting than (***"More Trusting Than", MTT***)*

17

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 definition of a trust act based exclusively on observable variables asking for the prospect of an improvement in utility terms is not feasible. Definition 5 therefore insists on the prospect of a material improvement.

While the two inequalities in Definition 5 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 5 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 definition 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 a definition of a trust act.

Definition 6 *Let the FM choose the actual opportunity set for the SM from the collection of opportunity sets **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. Data provide support for the empirical significance of both motivations of FM befavior 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 **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 **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 7 *Suppose the FM chooses the opportunity set $J \in \boldsymbol{O}$ from the collection of opportunity sets **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 8 *Let the FM choose the actual opportunity set for the SM from the collection of opportunity sets **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: (i) unconditional other-regarding preferences – like altruism, inequality aversion, quasi-maximin, etc.; (ii) deal-responsiveness – reacting to actions that allow for a mutual improvement by adopting behavior that implies a mutual improvement; (iii) gift-responsiveness – reacting to choices that allow the trustee to

obtain an improvement by adopting actions that benefit the trustor; and (iv) vulnerability-responsiveness – reacting to the vulnerability of the trustor by adopting actions that do not hurt the trustor. 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 to 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

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

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

[3] Arrow, K. J., 1974. "The Limits of Organization", *Norton.*

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

[5] Baier, A., 1986. "Trust and Antitrust", *Ethics*, 96, 231-260.

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

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

[8] Brandts, J. and G. Charness, 2011. "The strategy versus the direct-response method: a first survey of experimental comparisons", *Experimental Economics*, 14, 375–398.

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

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

[11] Coleman, J., 1990. "Foundations of Social Theory", *The Belknap Press of Harvard University Press.*

[12] Colquitt, J. A., S. A. Brent and J. A. LePine, 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.

[13] Cox, J. C., 2004. "How to Identify Trust and Reciprocity", *Games and Economic Behavior*, 46, 260-281.

[14] Cox, J. C. and C. A. Deck, 2005. "On the Nature of Reciprocal Motives", *Economic Inquiry*, 43, 623-635.

[15] Cox, J. C., D. Friedman and V. Sadiraj, 2008. "Revealed Altruism", *Econometrica*, 76, 31-69.

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

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

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

[19] Engler, Y., R. Kerschbamer and L. Page, 2014. "Intention-Based Social Preferences: A New Elicitation Method", Mimeo, School of Economics, Queensland University of Technology, Brisbane, Australia.

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

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

[22] Glaeser, E. L., D. Laibson, J. A. Scheinkman and C. L. Soutter, 2000. "Measuring Trust", *The Quarterly Journal of Economics*, 115, 811-846.

[23] Glaeser, E. L., D. Laibson, J. A. Scheinkman and C. L. Soutter, 1999. "What Is Social Capital? The Determinants of Trust and Trustworthiness", *NBER working paper No. 7216.*

[24] 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.

[25] Guiso, L., P. Sapienza and L. Zingales, 2008. "Trusting the Stock Market", *The Journal of Finance*, 63, 2557-2600.

[26] Ho, T. H. and K. Weigelt, 2005. "Trust Building among Strangers", *Management Science*, 51, 519-530.

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

[28] Johnson, N. D. and A. Mislin, 2011. "Trust games: A meta-analysis", *Journal of Economic Psychology,* 32, 865-889.

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

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

[31] LaPorta, R., F. Lopez-de-Silane, A. Shleifer and R. Vishny, 1997. "Trust in Large Organizations", *American Economic Review*, 87, 333–338.

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

[33] Lorenz, E. H., 1988. "Neither friends nor strangers: informal networks of subcontracting in French industry", *Trust. Making and breaking cooperative relations*, 194-209.

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

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

[36] Noorderhaven, N. G., 1992. "Trust and inter-firm relations", *Paper for the 1992 EAEPE Conference.*

[37] Rousseau, D., S. Sitkin, R. Burt and C. Camerer, 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.

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

[39] Sabel, C. F., 1990. "Studied trust: building new forms of cooperation in a volatile economy", F. Pyke, W. Sengenberger (Eds.), *Industrial Districts and Local Economic Regeneration, International Institute for Labor Studies*, 215-250.

Tables and Figures

Table 1: Treatments and Number of Observations

treatment	choice of *option (a)* implements	# of observations (pairs)
IG	$(f,s) = (15,15)$	40
DC	$(f,s) = (15,45)$	39
GC	$(f,s) = (15,60)$	39
VC	$(f,s) = (0,15)$	39
CC	n.a.*	38
overall		195

* n.a. stands for 'not available': there is no *option (a)* in CC

Table 2: Revealed Intentions in the Five Games

game	choice of *option (b)* reveals
IG	D, G, V
DC	G, V
GC	V
VC	D, G
CC	Nothing

In the Table the letters D, G and V stand for:

D: 'willingness to create the possibility for a mutual improvement';

G: 'willingness to allow the SM to obtain an improvement';

V: 'willingness to accept vulnerability'.

Table 3: Hypotheses Overview

| Abbreviation	Tests for	Data Source
H^{U}	unconditional other-regarding preferences	CC
$H^{D,G,V}$	joint impact of D, G and V	IG vs. CC
$H^{G,V(abs.D)}$	joint impact of G and V in the absence of D	DC vs. CC
$H^{V(abs.D,G)}$	impact of V in the absence of D and G	GC vs. CC
$H^{V(pres.D,G)}$	incremental impact of V in the presence of D and G	IG vs. VC
$H^{D,G(abs.V)}$	joint impact of D and G in the absence of V	VC vs. CC
$H^{D,G(pres.V)}$	joint impact of D and G in the presence of V	IG vs. GC
$H^{D(pres.G,V)}$	incremental impact of D in the presence of G and V	IG vs. DC
$H^{G(pres.V,abs.D)}$	incremental impact of G in the presence of V but absence of D	DC vs. GC

Table 4: Parametric and Nonparametric Tests of SMs Back-Transfers

| Treatment	IG	DC	GC	VC	CC
IG	12.13 [2.30] {40}				
DC	(0.18) [0.19]	12.69 [2.16] {39}			
GC	(0.217) [0.19]	(0.041) [0.022]	12.82 [2.23] {39}		
VC	(1.81*) [1.96**]	(2.074**) [2.20**]	(2.081**) [2.08**]	6.67 [1.95] {39}	
CC	(2.06**) [1.83*]	(2.37**) [2.16**]	(2.37**) [2.03**]	(0.187) [0.29]	6.18 [1.69] {38}

Diagonal: means in first line, standard errors in brackets and number of observations in braces.

Below the diagonal: two-tailed t-test statistics in parentheses and Mann-Whitney statistics in brackets.

* and ** indicate significance at the 10% and 5% level, respectively.

25
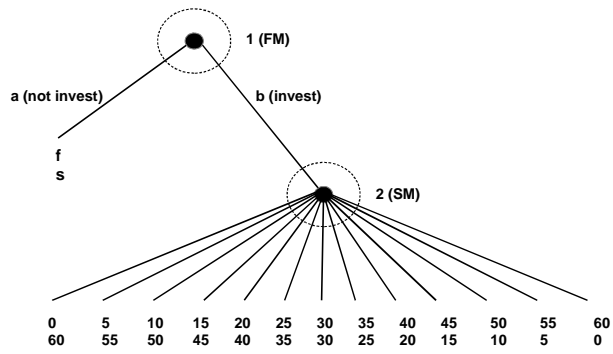
Panel A: Game Tree for Treatments IG, DC, GC and VC (with (f,s) as specified in Table 1)
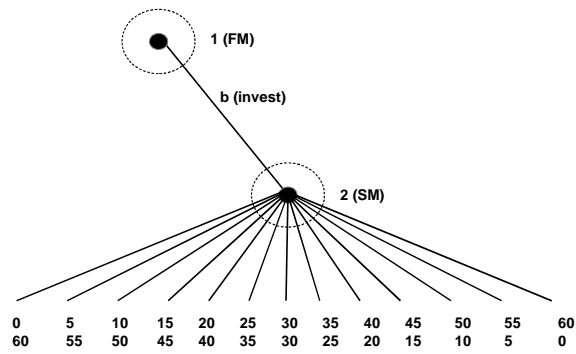
Panel B: Game Tree for Treatment CC

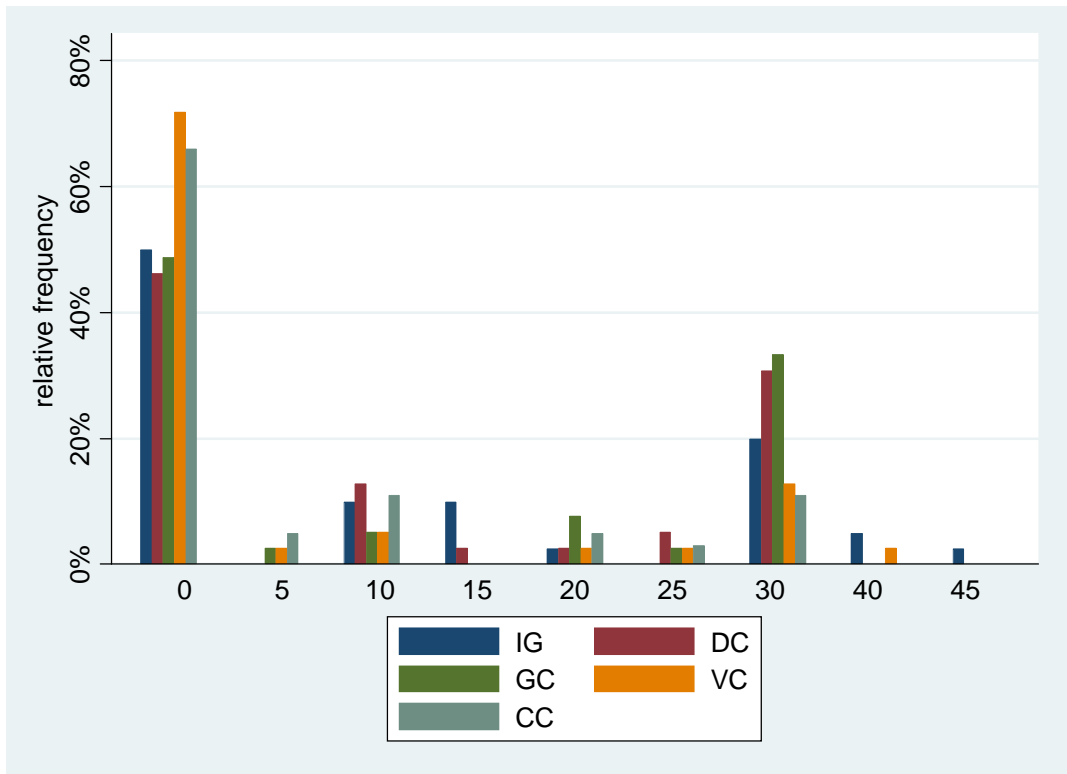Figure 1: Game Trees for the Five Treatments

26

Figure 2: Rel. Frequencies of Transfers from SMs to FMs in the Five
Treatments

Appendix

Experimental Instructions

In the following we provide an English translation of our originally German instructions. We provide the instructions and the decision sheets for treatments IG and CC here. The instructions and decision sheets for the treatments DC, GC and VC are identical to those for treatment IG (except for the payoffs when Group A member chooses left). German instructions and questionnaire are available on request.

[General Instructions, at start of session]

Welcome to an experiment on decision making. We thank you for your participation!

The experiment consists of two parts: a decision making part and a questionnaire at the end of the experiment. A research foundation has provided the funds for conducting the experiment. You can earn a considerable amount of money by participating. The text below will tell you how the amount you earn will be determined.

No Talking Allowed
Please do not talk to any other participant until the experiment is over. If there is anything that you don't understand, please raise your hand. An experimenter will approach you and clarify your questions in private. In about ten minutes this document will also be read aloud (by an experimenter).

Anonymity
You will never be asked to reveal your identity to anyone during the experiment. Neither the experimenters nor the other subjects will be able to link you to any of your decisions. In order to keep your decisions private, please do not reveal your choices to any other participant. The following means help to guarantee anonymity:

Non-Computerized Experiment and Private Code
The task you have to complete during the experiment is conducted in private on a printed form; that is, the experiment is not computerized. You have drawn a small sealed envelope from a box upon entering the room. PLEASE DO NOT OPEN YOUR ENVELOPE BEFORE THE EXPERIMENT STARTS. Your envelope contains your participation number. We will refer to it as "your private code" in the sequel. Your private code is the only identification used during the experiment and you will also need it to collect your cash payments.

Cash Payments

During the experiment we shall not speak of Euros but rather of Talers. At the end of the experiment the total amount of Talers you have earned will be converted to Euros at the following rate:

$$5 \text{ Taler} = 1 \text{ Euro}$$

Cash payments can be collected from Monday onwards in room w3.29 on the third floor (West) of this building. You will present your private code to an admin staff person (Mrs. xy) and you will receive your cash payment in exchange. The admin staff person will not know who has done what and why, nor how payments were generated. No experimenter will be present in the room when you collect your money. Also, the private codes of this experiment will be mixed up with the codes of other experiments. This will again help to guarantee that the amount you earn cannot be linked to your decisions. Mrs. xy is available from Monday to Friday between 9 a.m. and 11 a.m., as well as between 2 p.m. and 3 p.m. in room w3.29 on the third floor (West) of this building. Please collect your earnings within a weak (you find those details also on the card displaying your private code).

Two Groups

Before the experiment starts, the participants in this room will be randomly divided into two groups of equal size. The groups are called Group A and Group B. Members of Group A will be seated in this room, members of Group B will be seated in the adjacent room.

Role Assignment and Start of the Experiment

After the instructions at hand and the detailed instructions have been read aloud and all questions have been answered you (and all other participants in this room) will be asked to open the sealed envelope you drew from the box when entering this room. The envelope contains a card with your private code. The code ends with a number. If this number is even, you are a member of Group A, if it is odd, you are a member of Group B. Members of Group A are asked to take a seat at one of chairs in this room. Members of Group B will be escorted to the adjacent room and asked to take a seat there. Then the decision sheets will be handed out and the experiment starts.

[Instructions for Treatment IG]

Matching

Each member of Group A is anonymously paired with a member of Group B. The matching is 1:1; that is, each member of Group A is exactly matched with one member of Group B and vice versa. You will **never learn the identity of the member of the other group you are paired with.** In the same way, the member of the other group you are paired with will not learn your identity. In the following we call the member of the other group you are matched with **the other person.**

Decision Tasks
Each member of Group A and each member of Group B will be asked to make a single decision. The decision of a Group A member always affects not only her or his own payment but also the payment of the other person. Whether the decision of a Group B member has an impact on the own payment of the Group B member and on the payment of the paired Group A member depends on the decision of the Group A member. We will explain this in detail:

Decision Task of Group A Members and Possible Payments
Each member of Group A will be asked to decide between the two alternatives LEFT and RIGHT. This is done by marking one of the two options on the top of the decision sheet (you find a copy of the decision sheet of a Group A member on the next page).

- If the Group A member chooses LEFT, the payments to both persons (Group A member and paired Group B member) **are not determined** by the decision of the paired Group B member.

- If the Group A member chooses RIGHT, the payments to both persons (Group A member and paired Group B member) **are determined** by the decision of the paired Group B member.

Decision Task of Group B Members and Possible Payments
Each member of Group B will be asked to decide between thirteen alternatives. This is done by marking one of the 13 rows on the bottom/right of the decision sheet (you find a copy of the decision sheet of a Group B member on the next page but one).

- If the paired Group A member has chosen LEFT, the decision of the Group B member is irrelevant for the payments. In this case the payments for the Group A member and for the Group B member are as shown on the left hand side of the decision sheet.

- If the paired Group A member has chosen RIGHT, the decision of the Group B member determines the payments for the Group A member and for the Group B member. In the column "Group A member receives" on the right hand side of the decision sheet you see how many Talers the member of Group A gets and in the column "Group B member receives" you see how many Talers the member of Group B gets if the respective row is chosen by the Group B member.

Pay attention to the fact that the members of Group A and the members of Group B make their decisions simultaneously. Therefore the member of group B has to make a decision without knowing the corresponding decision of the Group A member.

In the case of wrong decisions (the member of Group A marks both options or none of them or the member of Group B marks more than one row or none) you get no payment from this part of the experiment and the payment to the other person is determined randomly (by randomly implementing one of your possible choices and matching it with the choice of the other person).

After you have made your decision on the decision sheet, put the decision sheet into the envelope and wait until an experimenter will collect it. Pay attention that the envelope is not marked in any way.

Decision Sheet of a Member of Group A

If you are a member of Group A your decision sheet will look as shown below. Note that this page is NOT the decision sheet. The decision sheet will be handed out to you at the start of the experiment.

Decision Sheet - You Are a Member of Group A

Please enter your private code (you will get no payment out of the experiment in case the private code is missing or wrong):

....

Decision Sheet of a Member of Group B

If you are a member of Group B your decision sheet will look as shown below. Note that this page is NOT the decision sheet. The decision sheet will be handed out to you at the start of the experiment.

Decision Sheet – You Are a Member of Group B

Please enter your private code (you will get no payment out of the experiment in case the private code is missing or wrong):

....

[Instructions for Treatment CC]

Matching
Each member of Group A is anonymously paired with a member of Group B. The matching is 1:1; that is, each member of Group A is exactly matched with one member of Group B and vice versa. You will **never learn the identity of the member of the other group you are paired with.** In the same way, the member of the other group you are paired with will not learn your identity. In the following we call the member of the other group you are matched with **the other person.**

Decision Tasks
Only members of Group B will be asked to make a single decision. The decision of a Group B member always affects not only her or his own payment but also the payment of the paired Group A member. Group A members will be asked to answer a single question (you find a copy of the decision sheet of a Group A member on the next page) in the meantime and the answer to this question does not affect her or his own payment nor the payment of the paired Group B member. We will explain this in detail:

Decision Task of Group B Members and Possible Payments
Each member of Group B will be asked to decide between thirteen alternatives. This is done by marking one of the 13 rows on the bottom/right of the decision sheet (you find a copy of the decision sheet of a Group B member on the next page but one).

- The decision of the Group B member determines the payments for the Group A member and for the Group B member. In the column "Group A member receives" on the right hand side of the decision sheet you see how many Talers the member of Group A gets and in the column "Group B member receives" you see how many Talers the member of Group B gets if the respective row is chosen by the Group B member.

In the case of wrong decisions (the member of Group B marks more than one row or none) the Group B member gets no payment from this part of the experiment and the payment to the other person is determined randomly (by randomly implementing one of the possible choices and matching it with the other person).

After you have made your decision on the decision sheet respectively after you have answered the question on the question sheet, put the decision sheet (question sheet) into the envelope and wait until an experimenter will collect it. Pay attention that the envelope is not marked in any way.

Question Sheet of a Member of Group A

If you are a member of Group A your question sheet will look as shown below. Note that this page is NOT the question sheet. The question sheet will be handed out to you at the start of the experiment.

Question Sheet – You Are a Member of Group A

Please enter your private code (you will get no payment out of the experiment in case the private code is missing or wrong):

....

Question: "How many Talers do you expect from your paired Group B member?"

....

Decision Sheet of a Member of Group B

If you are a member of Group B your decision sheet will look as shown below. Note that this page is NOT the decision sheet. The decision sheet will be handed out to you at the start of the experiment.

Decision Sheet – You Are a Member of Group B

Please enter your private code (you will get no payment out of the experiment in case the private code is missing or wrong):

....