Tjalling C. Koopmans Research Institute

Discussion Paper Series nr: 06-21

Crowding out in an indefinitely repeated Asymmetric Trust Game

Thomas Dirkmaat, Stephanie Rosenkranz, Vincent Buskens

Universiteit Utrecht

Utrecht School of Economics

Tjalling C. Koopmans Research Institute
Utrecht School of Economics
Utrecht University

Janskerkhof 12
3512 BL Utrecht
The Netherlands
telephone +31 30 253 9800
website www.koopmansinstitute.uu.nl

The Tjalling C. Koopmans Institute is the research institute and research school of Utrecht School of Economics. It was founded in 2003, and named after Professor Tjalling C. Koopmans, Dutch-born Nobel Prize laureate in economics of 1975.

In the discussion papers series the Koopmans Institute publishes results of ongoing research for early dissemination of research results, and to enhance discussion with colleagues.

Please send any comments and suggestions on the Koopmans institute, or this series to M.deSwart-Gijsbers@econ.uu.nl

ontwerp voorblad: WRIK Utrecht

How to reach the authors

Please direct all correspondence to the first author.

Thomas Dirkmaat
Utrecht University
Utrecht School of Economics
Janskerkhof 12
3512 BL Utrecht
The Netherlands
E-mail: t.dirkmaat@econ.uu.nl

Stephanie Rosenkranz
Utrecht University
Utrecht School of Economics
Janskerkhof 12
3512 BL Utrecht
The Netherlands
E-mail: s.rosenkranz@econ.uu.nl

Vincent Buskens
Utrecht University
Department of Sociology/ICS
Heidelberglaan 2
3584 CS Utrecht
The Netherlands
E-mail: v.buskens@fss.uu.nl

Utrecht School of Economics
Tjalling C. Koopmans Research Institute
Discussion Paper Series 06-21

Crowding out in an indefinitely repeated Asymmetric Trust Game

Thomas Dirkmaat[a]
Stephanie Rosenkranz[b]
Vincent Buskens[c]

[a]Utrecht School of Economics
Utrecht University

[b]Utrecht School of Economics
Utrecht University

[c]Department of Sociology/ICS
Utrecht University

December 2006

Abstract
In this paper we introduce an alternative version of the trust game by Dasgupta (1988) and Kreps (1990) that allows for asymmetric information. We use this version to study the effect of checking on the trustee's behaviour, checking is a control option the trustor can decide to use and that takes place after both trustor and trustee made their initial decisions. 'Checking' differs in this respect from the often in the literature found 'monitoring' that allows the trustor to control the trustee's behaviour before the trustee makes his decision. The game theoretical analysis suggests that checking increases co-operation. The experimental results show that this is only true for the selfish part of the trustee population. Honest trustee react negatively to checking, which is more in line with crowding out theory.

Keywords: Trust, Asymmetric Information, Experiment, Checking, Crowding Out

JEL classification: C71, C91, D82

Acknowledgements
Authors would like to thank the Tjalling C. Koopmans Institute for financial support.

1 Introduction

In the economic literature one can find many articles about contractual relations and the role control and incentive devices, like reward and punishment, can have on opportunistic behaviour (e.g. Alchian and Demsetz (1972) and Prendergast (1999)). The traditional economic view is very well described by Williamson (1993): "... because opportunistic agents will not self-enforce open-ended promises to behave responsible, efficient exchange will be realized only if dependencies are supported by credible commitments." According to this point of view it is in the principal's interest to use control or incentive devices to make sure that opportunistic agents will behave "responsible". Crowding out theory on the other hand claims that these control and incentive devices can have a hidden cost in the sense that they can reduce the agent's willingness to co-operate instead of increasing it (Lepper and Greene (1978)). When the principal decides not to use control and incentive devices he needs to trust the agent. The principal's trustfulness is based on his expectations about the agent's future behaviour. The principal (trustor) is fully aware of the fact that the agent (trustee) has an option to abuse his trust, nevertheless he expects that this will not happen. This uncertainty forms the basis of the concept of trust, without this it is impossible to speak of trust.

There are many empirical studies regarding the crowding out effect, for an overview see Frey and Jegen (2001). Most related to our study are Falk and Kosfeld (2004) and Ploner (2006), both study crowding out in trust settings. Falk and Kosfeld (2004) in their experiment gave the principal two options: Give the agent a contract that allows him to decide his own effort level, in other words the principal trusts the agent. Give the agent a contract with a fixed effort requirement, in other words the principal distrusts the agent. Standard economic theory predicts that people will only put in effort if they are rewarded or forced to do so. So a fixed effort requirement will lead to a higher effort level. If some people see themselves as trustworthy, the principal signals that he does not trust them when he offers them a contract with a fixed effort requirement. Because their trustworthiness is not acknowledge they will deliver the effort the contract demands, in other words trustworthiness is replaced by extrinsic control. Would their trustworthiness have been acknowledge crowding theory argues that they would have put in more effort, crowding in. As a result of their experiment Falk and Kosfeld (2004) found that agents who are free to choose an effort level, indeed choose a higher effort level (note that effort is costly for the agents!) than those who are faced with a minimum effort requirement by the principal. This finding is consistent with crowding out theory. In Ploner (2006) the principal can either decide to play the standard investment game or he can play a modified investment game with intention detection. In the standard game the principal can choose to invest a certain amount of money in the agent. His investment is multiplied with an given factor m. When it is the agent's turn he can decide to return a certain amount of money back to the principal. In the modified game the principal can buy a detection technology, which asks the agent to write down what he would return given all the possible investments the principal can make. When the principal makes an investment the agent is forced to return the corresponding amount of money. When crowding out theory is correct we should assume that agents will react negatively on the intention detection system, because it does not recognise their reciprocal nature. Ploner (2006) found that this is not true. Agents are willing to return more in the modified game than in case of the standard game. However it should be noted that this study has very few observations.

In this study we are interested in the effects a weaker form of control, which we will label checking, has on the trustee's behaviour in trust relations. We use the simplest game describing a trust relation, namely the trust game introduced by Dasgupta (1988), Camerer and Weigelt (1988) and Kreps (1990). While in Falk and Kosfeld (2004) the trustee and in Ploner (2006) both trustor and trustee can make a decision by choosing a value from a discrete range.[1] In this game both trustor and trustee have binary decisions. The trustor can place trust or not. When trust is placed the trustee can either choose honour or abuse. In most trust situations the trustor does know at the end of an interaction if his trust has been abused or not. In this study we let go of this basic feature by adding asymmetric information to the trust game. Green and Porter (1984) developed a similar model, in their case firms observe market prices, which imperfectly reflect the output levels of other firms. In our asymmetric trust game the trustor doesn't know at the end of an interaction if his trust has been abused or not. The main reason we add asymmetric information to the trust game is the fact that it allows us to test the effect of checking. Checking is an option given to the trustor. It allows him to check what the trustee did after both he and the trustee made a decision. Checking is a weak form of control that

[1] In the first study the trustee chooses an effort level in the second study the trustor can choose the height of the investment and the trustee the height of the return rate.

takes place after the trustee made his decision and eliminates the information asymmetry. In the studies mentioned above the form of control is much stronger and control takes place before the trustee makes his decision. In Falk and Kosfeld (2004) the trustor can determine the trustee's choice set directly. Either the trustee has complete freedom in choosing his effort level, or a minimum level is required. In Ploner (2006) the trustor can detect the trustee's intention before he needs to make a decision, when this detection technology is used the trustee must respond according to his revealed intentions. In order for checking to have an effect on the behaviour of the trustee we need to repeat the game. In this paper we focus on long-run trust relations. Long-run relations can be split-up in two categories: relations with a predetermined end, like a temporary contract and relations with an uncertain end. In the latter case we know with certainty that it will end, we just don't know when it will end exactly. This difference is important and is reflected in our daily behaviour.[2] In this study we will limited ourselves to long-run relations with an uncertain end. To test if this checking option leads to more or less co-operation on the side of the trustee we conduct an experiment. Our experiment consists of two main treatments: the first is the indefinitely repeated asymmetric trust game without checking and the second is the same game only this time with checking. Intuitively one might expect to see more co-operation in case of checking, because uncertainty is reduced by an increase in information. This is also formally confirmed by our theoretical analysis. Selfish trustees should honour trust in the checking treatment compared to the no-checking treatment where they should not. Trustees might still honour trust in the no-checking treatment due to social preferences. Our predictions show that two groups of trustees with social preferences can be distinguished: honest trustees who will always honour trust regardless of the treatment they are in and moderate reciprocates who will co-operate in the no-checking treatment when the parameter of their social preferences meets a critical threshold value.[3] Crowding theory suggest that both honest and moderate reciprocating trustees might be crowded out by checking. Summarising the theoretical predictions: checking should

[2]We all know we will die one day. We also know what the average life expectancy is. Finally, we are aware of the fact that there is a very small probability we get involved in a fatal accident. Assume you get killed at age 40 due to an accident. You did not know the accident would happen, so you just lived your life like you were used to, until the moment of the accident. Now assume you would know from the beginning of your life that you would die in an accident at age 40. At certain moments in your live you probably would have made different choices.

[3]Because selfish trustees play honour in the checking treatment the moderate reciprocating trustees also co-operates in this treatment. For more details see section 3.

give selfish trustees the incentive to honour trust, while trustworthy trustees should be crowded out and honour less trust. The results from a pilot study give an indication in the same direction, see section 4.

This paper has the following outline. In the remainder of this section we will present our idea of checking in more detail. In section 2 we present the experimental design, followed by the theoretical predictions in section 3. In section 4 the empirical results are presented. The paper ends with a discussion in section 5.

1.1 Trust under asymmetric information

In this paper we will study how trustees react to being checked by the trustor in situations of asymmetric information. To get a better feeling for the topic we give the following example. We take a long run trust relationship as a starting point. Usually in a trust relationship both trustee and trustor know the outcome of an interaction. Think for example of an investor, the trustor, and an investment bank, the trustee. The investor does not know much about financial markets, but wants to have a high return on his invested money. So he asked the investment bank to invest his money for him. At the end of each period the trustor will see the return on his money. The size of the return depends on the effort the investment bank puts in selecting the best stock and bonds portfolio for the investor. When the investment bank is trustworthy and puts in high effort, we expect a good return. In case the investment bank abuses the trust of the investor by putting in low effort, the returns on investment will be low. In this example both trustor and trustee know if trust was placed and, when it was placed, if it was honoured or abused.

However in the short run this does not need to be true for the trustor. Although the trustor will usually find out what happened in the end, in the short run the action of the trustee might be unknown to the trustor. Assume that the returns on investment do not only depend on the effort level of the investment bank, but also on external shocks. Although the investment bank chose the portfolio carefully, market conditions can lead to low returns. Likewise when the investment bank puts in low effort while choosing a portfolio, for instance by selecting stocks at random, the results can still be good due to sheer luck. When external factors like luck and more importantly market conditions can influence the outcome of a sequence of actions in such a way that this outcome could have been the result from either action of the trustee, the trustor is no longer capable to directly infer the trustee's behaviour from this outcome. So in the short run it might be impossible for the investor to judge the trustworthiness of the investment bank, due to asymmetry in information. In the long run we can assume that, for example, the investor should be able to compare the returns the investment bank achieved on his money with macro economic indicators and thus get an impression if the investment bank is trustworthy or not.

It should be clear that the investor would be interested in having a way to find out if the investment bank is trustworthy or not. The investor might be able to check the trustworthiness of the investment bank when he employs a financial expert for a second opinion. We assume that the investment bank will find out that the investor hired a financial expert, because the financial world is a tight community. It should be noted that the effect of checking can only been seen in case of a long-run relationship. When the investor checks at time T the investment bank can react to this at time T+1 when the investor played trust again.

1.2 Checking

We have chosen to call this weak form of control 'checking' for two reasons: first, to make a distinction with the often in the literature found monitoring. Monitoring refers to ex ante controlling activities. The trustor gives two different signals, trusting and monitoring, to a trustee before the trustee can make a decision to either honour or abuse trust. Checking is ex post: the trustor first places trust, than the trustee can make his decision and finally the trustor has an option to check what the trustee did. Second, there is a more linguistic argument. A synonym for checking would be verifying. However, verifying has a more formal denotation, for example, when a scientist verifies a certain phenomenon. Given the informal status of the option to check in our context, we call this behaviour checking.
What makes checking a weak form of control? Checking, although in most situations costly for the trustor, does not reward nor punish the trustee in a direct material sense nor can the trustor force the trustee to take some action. In the example of the investor and the investment bank, checking is costly: a financial expert is hired. Checking means however that the investor is not allowed to promise bonuses to the investment bank for achieving good results. Nor is the investor allowed to demand a minimal return on

investment.
The trustor acquires additional information when he performs a check, but the check itself it not restricting the trustee's behaviour. Checking is of course not entirely without consequence, in the sense that the trustor will use the information he got for future decisions. When he finds out that the trustee has abused him, it is likely that he will punish by not placing trust for some periods, or when he is less forgiving, by terminating the trust relationship altogether. When he discovers that the trustee was trustworthy, the trustor should be willing to continue his trust relationship. In case crowding out plays a role the trustee will stop honouring trust when he sees that the trustor did not really trusted him.

2 Experimental design

2.1 The trust game

The starting point of our experiment is the trust game introduced by Dasgupta (1988), Camerer and Weigelt (1988) and Kreps (1990). Their trust game is the simplest game describing the essence of a trust relationship. It is a sequential game where the trustor moves first and can decide to place either trust or not. Then it is the trustee's turn who can decide to either abuse or honour trust. The trust game depicted in figure 1 makes clear that the relationship is beneficial to the trustor in case his trust is honoured, while if his trust is abused he is worse off because $S_1 < C_1$. For the trustee the gain from abusing (A_2) is greater than the gain from co-operation (C_2). This gives him an incentive to abuse. The trustor will anticipate the trustee's abuse by playing no trust, because $N_1 > S_1$. These three conditions are essential to speak of a trust game. When the trustee has no incentive to abuse or when the trustor would gain more from being abused than from choosing not to trust, this game does not reflect a trust situation. Note although the trust game describes a trust situation, the pure strategy Nash equilibrium is not to trust.

2.2 The asymmetric trust game

In the original trust game the payoffs are common knowledge: the trustor will always know what the trustee has done. In order to introduce a checking

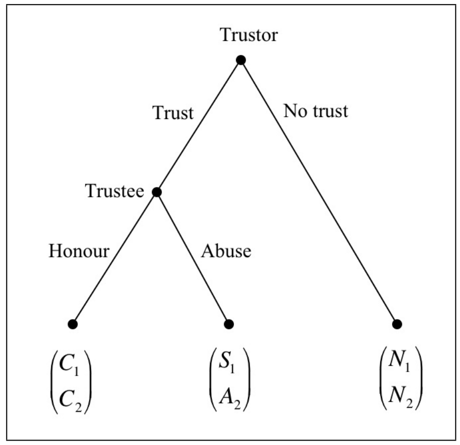


Figure 1: The trust game. C_1 and C_2, S_1 and A_2, N_1 and N_2 are all material payoffs combinations the trustor and trustee can realize when playing the game.

option the trustor should not immediately be able to see this in our set-up. That is why we introduce an asymmetric version of the trust game. In this version, the trustor's payoffs when playing trust are expected values based on two realisable earnings with two different probability distributions. Which distribution is used depends on the trustee's decision. The two realisable earnings are two payoffs one is a high payoff (H) and the other a low payoff (L). Let p_a be the probability linked with L_a, the low payoff in case of abuse. Thus $(1-p_a)$ is the probability of H_a, the high payoff in case of abuse. Let p_h be the probability of L_h, the low payoff in case of honour, so the probability of H_h, the high payoff in case of honour, is $(1-p_h)$. When the trustor should not be able to distinguish between his trust being abused or honoured, then it needs to be the case that $L_h = L_a$ and $H_h = H_a$. In other words, the trustor only observes H and L. The trustee has the same information he has in the standard trust game. For a graphical representation of the asymmetric trust game see figure 2.

It should be noted that in order to speak of a trust game p_a and p_h cannot take on every possible value between 0 and 1. To maintain strategic equivalence to the original trust game we need in expected terms $E(C_1) > E(S_1)$. This implies:

$$p_h L + (1 - p_h)H > p_a L + (1 - p_a)H \tag{1}$$

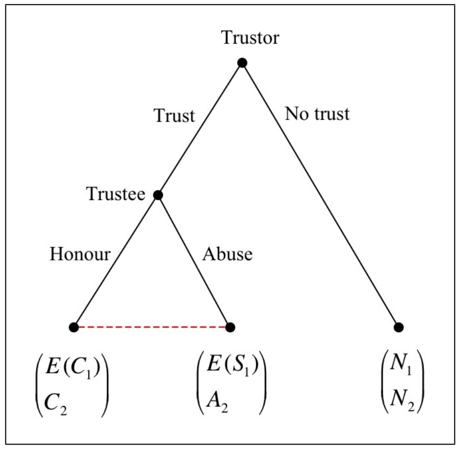


Figure 2: The asymmetric trust game. The dotted line reflects that the trustor does not know in which note he is.

Which simplifies to:

$$p_h < p_a \tag{2}$$

2.3 The experiment

In the experiment we use the following values $p_a = \frac{3}{4}$, $p_h = \frac{1}{4}$, $H = 24$ and $L = 4$. In case of abuse the trustor's payoff is given by the following probability distribution:

$S_1 : \frac{3}{4}4 + \frac{1}{4}24$, with an expected value of 9

and in case of honour it is:

$C_1 : \frac{1}{4}4 + \frac{3}{4}24$, with an expected value of 19

When the trustor earns 4 or 24, it is not clear to him whether his trust has been abused or honoured that period. C_1 has an expected value of 19 and S_1 of 9. Payoffs N_1 and N_2 have both a certain value of 12 and C_2 and A_2 have a certain value of 16 and 24 respectively. In appendices B and C a graphical representation of the asymmetric trust game with and without checking option are presented. In our opinion an indefinitely repeated trust game is good resemblance of real world trust relations, with uncertainty about the moment it ends, where checking can play a role, think for example of investor investment bank relations, employer-employee relations or long-term supply contracts. In the experiment we simulate this by letting the subjects play a game with a continuation rule. All subjects know that after every round the game continues to the next round with a probability δ. In the experiment we use $\delta = 0.925$. The expected number of rounds equals 13,33.[4]

In the asymmetric trust game the realised earnings do not give the trustor information about the behaviour of the trustee. This game is called the no-checking treatment (NCT). In the checking treatment (CT) the subjects will play the same game, but this time the trustor has an option to check on the trustee's decision however. When the trustor checks he will still earn either 4 or 24 but this time it will also be revealed whether the trustee honoured or abused his trust. If the trustor decided to check, the trustee is immediately notified of this. We made checking costless so we can focus on the effect of checking on the trustee's behaviour. The cost of checking has an effect on the behaviour of the trustor. The instructions for both treatments are enclosed in the appendices B and C.

In section 3 it will become clear that we distinguish different types of trustees based on differences in preferences. Two disentangle the selfish from the honest trustee in the data we start each session with a type treatment (TT). In this treatment the subjects play an isolated encounter trust game, with the same payoffs as in the asymmetric trust game. However in this case the trustor's payoffs do not depend on a probability distribution. The trustor either earns 16 in case his trust is being honoured or 8 in case of abuse. The game is played in the following way: all subjects are asked to make the decision of the trustor. Next they are asked to make the decision of the trustee, under the assumption that the trustor has chosen trust. After they made both decisions they are randomly given a role and they are matched with someone with a different role. Their earnings are calculated in such a way that the decision they made in the role assigned to them is linked to the decision made by their partner in his role. In this way the equilibrium strategy remains the same, but we can observe for all participants if they would honour trust. When we would play a normal isolated encounter game, with matching before the game starts, we would not be able to observe the reaction of trustees who were not trusted. The instructions for the type treatment can be found in appendix A.

In the experiment we use the 'within subjects' design. The comparison of

[4]The expected number of rounds is equal to $\frac{1}{1-\delta}$.

	Treatment number
Session type A	(TT) Type game (NCT) Trust game without checking option (CT) Trust game with checking option Questionnaire
Session type B	(TT) Type game (CT) Trust game with checking option (NCT) Trust game without checking option Questionnaire

Table 1: Ordering of the treatments.

both treatments allows us to find out whether there is any effect of checking. To compensate for the ordering effect we need to run two different sessions. The first session will start with the NCT and finish with the CT. The second session is ordered vice versa, see table 1. Both sessions end with a questionnaire.

2.4 Carrying out the experiment

The experiment is computer based, using z-Tree by Fischbacher (1999) both for programming the treatments and to run the experiment. Subjects are students from the subject pool of ELSE.[5] Twenty subjects were invited for each session, the maximum number of subjects the computer laboratory of Utrecht University can occupy. When subjects enter the laboratory, numbers are handed out at random. The number a subject receives is linked to a computer. By handing out the numbers randomly we want to avoid that people, who might know each other, will sit closely together. Deviders separate the computers so that subjects cannot give signals to each other during the experiment. This, combined with the random numbers, should limit communication as best as possible. The two roles, trustor and trustee, are randomly assigned to the subjects. In each session we have ten trustors and ten trustees. During each session the subjects keep the same role. The computer matches a trustor and a trustee before a treatment starts. During a treatment, trustor and trustee stay together throughout the 20 periods. The experiment starts with handing out the instructions for the first treat-

[5]Experimental Laboratory for Sociology and Economics of Utrecht University.

ment. The subjects get enough time to read the instructions. The instructor announces the beginning of the first treatment. After the first treatment, the instructions for the second treatment are handed out. For the second treatment, the computer matches new combinations of trustors and trustees. We ensured the subject that they will not play two treatments with the same person. After all subjects have read them the instructor announces that the second treatment will start. Each session ends with a questionnaire. In the meantime, the instructor(s) count(s) out the money earned by the subjects. After all subjects have answered the questionnaire, the subjects receive their earnings. We pay subjects according to the strategies they play in the games. Each number depicted in our trust game reflects points, which are translated into Eurocents afterwards. Where 1 point corresponds to 1.75 eurocent. On average subjects earn €10 an hour. A session takes about one hour and 15 minutes. To avoid home-grown priors we used abstract labels of the players and the strategies in both treatments (see appendices A and B).

3 Behavioural predictions

When using experiments to test economic theories one runs quickly into the problem of distinguishing utility from material payoffs. In theory payoffs can represent utility. In the laboratory the payoffs can only represent material wealth. Recall that the points subjects earn are translated into money. This means that subjects might experience the game differently. Some care only about the material wealth, others might derive additional utility from their strategy. In the literature social preference functions try to explain the difference between theoretical results and experimental findings. In this paper we use a simple model introduced in Levine (1998) that allows for altruism and reciprocity, where v_i represents the adjusted utility of person i:

$$v_i = u_i + \frac{a_i + \lambda a_j}{1 + \lambda} u_j \tag{3}$$

According to this model a person receives a direct utility, in our terms material wealth, of u_i. Every person also has an altruism coefficient $-1 < a_i < 1$, which makes him care about the utility of another person. But because $-1 < a_i < 1$ he never cares more about that other person than about himself. When $a_i > 0$ a person is altruistic. If $a_i < 0$ the person is called spiteful. And finally when $a_i = 0$ the person is selfish. A person can have a higher regard for altruistic opponents than for spiteful ones, this is reflected by λ. Altruism and spitefulness on the side of your opponent is expressed by a_j. When $\lambda = 0$ we are dealing with a person driven only by altruism. When $\lambda > 0$ the person is also reciprocal in the sense that he cares about the other person given that the other person cares positively or negatively about him. In the remainder of this paper we will denote $\frac{a_i+\lambda a_j}{1+\lambda}$ by ρ. ρ is an indicator for social preferences, reflecting how much one player cares about the well-being and the kindness of the other player. Note that ρ is increasing in a_i and a_j at a constant rate. The effect of λ depends on the relative values a_i and a_j. When $a_j > a_i$, the trustee cares more about the trustor then the other way around, ρ is in creasing in λ. When $a_j < a_i$, the trustee cares less about the trustor then the other way around, ρ is in decreasing in λ.

In the checking treatment the trustor has an option to check on the behaviour of the trustee. This option to check is costless. When he checks every round the information structure will be the same as in a game of symmetric information. Although his payoffs in case of trust still depend on a probability distribution, the trustor can turn the asymmetric trust game into a normal trust game in terms of strategies. This is why we will start our analysis with describing co-operation strategies in the trust game with checking.

3.1 Indefinitely repeated asymmetric trust game with checking

TG(∞, δ) denotes the indefinitely repeated version of the one-stage trust game TG, where δ represent the probability that the game will continue in the next period.[6] For each period t, the outcomes of period t-1 are observed before stage t begins. Recall from section 2 that $E(C_1)$ and C_2 are the material payoffs the trustor and respectively the trustee receive from co-operation. N_1 and N_2 are the material payoffs the trustor and respectively the trustee receive from the trustor's decision not to place trust. The trustor will receive $E(S_1)$ and the trustee A_2 when the trustee abuses the

[6] In this study we ignore temporal preference. We assume subjects will not be sensitive to time in an experiment that has a duration of less then an hour. We think it is very unlikely that subjects will appreciate earnings now much higher compared to earnings 1 minute from now. Given the short time span of the experiment we think this assumption can be justified.

trust of the trustor. The utility both trustor and trustee derive from the different strategy combinations can be found in table 2.

	Trustor's payoffs	Trustee's payoff
No Trust	$N_1 + \rho N_2$	$N_2 + \rho N_1$
Trust/Honour	$E(C_1) + \rho C_2$	$C_2 + \rho E(C_1)$
Trust/abuse	$E(S_1) + \rho A_2$	$A_2 + \rho E(S1)$

Table 2. Trustor's and trustee's utility per strategy combination.

It should be noted that while allowing for social preferences several cases can occur. For example it is possible that $E(S_1) + \rho A_2 > N_1 + \rho N_2$ for some values of ρ. Or to put it differently, when $\rho > \frac{N_1 - E(S_1)}{A_2 - N_2}$ the trustor's altruistic preferences make him prefer being abused above placing no trust. In this case it becomes difficult to speak of a trust game or relationship. When $\rho > \frac{A_2 - C_2}{E(C_1) - E(S_1)}$ the trustee prefers honouring trust above abusing trust, due to his altruistic and reciprocal preferences.[7] Because ρ is not common knowledge it is still possible to speak of a trust game in this case. Both trustor and trustee can have their own values of ρ.

In total four equilibria will be discussed, which can also be found in table 3. We refer to the table by writing a roman number between brackets behind an equilibrium. When the trustor's $\rho > \frac{N_1 - E(S_1)}{A_2 - N_2}$, he prefers abuse above choosing no trust. In an isolated encounter trust game, depending on the trustee's ρ, the equilibrium is either (Trust, Honour) (I) in case of $\rho > \frac{A_2 - C_2}{E(C_1) - E(S_1)}$ or (Trust, Abuse) (II) in case of $\rho < \frac{A_2 - C_2}{E(C_1) - E(S_1)}$. When the trust game is indefinitely repeated these equilibria remain the same. It is not possible for the trustor to increase co-operation in case of the (Trust, Abuse) equilibrium because he cannot credibly punish the trustee. As mentioned above these equilibria are an exception, because they are based on extreme altruism on the side of the trustor. We don't expect them to have serious predictive power in our experiment.

In case the trustor's $\rho < \frac{N_1 - E(S_1)}{A_2 - N_2}$, the trustor can play the following trigger strategy: The trustor will start playing trust and will continue doing so as long as the trustee plays honour. As soon as the trustee plays abuse the trustor will play no trust in the next and all coming rounds.

[7] Because the trustor moves first it is impossible to have any reciprocal feelings towards the trustee.

Given the trustor's strategy the trustee will honour trust as long as:

$$\sum_{t=1}^{\infty}(C_2 + \rho E(C_1))\delta^{t-1} > A_2 + \rho E(S_1) + \sum_{t=2}^{\infty}(N_2 + \rho N_1)\delta^{t-1} \tag{4}$$

or when:

$$\delta > \delta^* \equiv \frac{(A_2 + \rho E(S_1)) - (C_2 + \rho E(C_1))}{(A_2 + \rho E(S_1)) - (N_2 + \rho N_1)} \tag{5}$$

This trigger strategy constitutes a subgame perfect Nash Equilibrium. Depending on δ and ρ three cases can occur. When $\rho = 0$, we call the trustee selfish, and when $\rho > \frac{A_2 - C_2}{E(C_1) - E(S_1)}$, we call him honest. A third group of trustees, that we would like to call moderate reciprocates, is defined by the following condition: $\frac{A_2 - C_2}{E(C_1) - E(S_1)} > \rho > 0$. First when (5) is true a selfish trustee will prefer honouring trust above abuse. This is the standard equilibrium condition that allows for co-operation in indefinitely repeated games (III). Second if $\delta < \delta^*$ the selfish trustee will prefer to play abuse whenever he can. The trustor now prefers to play no trust. (IV). For an honest trustee $\delta > 0$, in other words he will always co-operate for every probability that the game will continue to the next round.[8] Moderate reciprocating trustees will also co-operate when selfish trustees co-operate, for every $\rho \in (0, 1)$. When for selfish trustees co-operation is not an equilibrium strategy, it might still be for moderate reciprocating trustees depending on ρ. Although for moderate reciprocates it is not optimal to co-operate in an isolated encounter trust game, they do co-operate in an indefinitely repeated game, given $\delta > \delta^*$.Solving (5) for ρ will yield us the critical ρ moderate reciprocates need to honour placed trust:

$$\rho > \rho^* \equiv \frac{A_2 - C_2 + \delta(N_2 - A_2)}{E(C_1) - E(S_1) + \delta(E(S_1) - N_1)} \tag{6}$$

To show this, we can also look at the partial derivative of (5), $\frac{\partial\delta}{\partial\rho}$, which is a decreasing function, in other words the right hand side of (5) will decrease when ρ gets larger, making co-operation more likely.

Using the payoffs from our experiment we can calculate the equilibrium value of δ. For selfish trustees, $\rho = 0$: $\delta > \frac{24-16}{24-12}$ or $\delta > \frac{2}{3} \approx 0.67$. With $\delta = 0.95$ in the experiment all selfish, moderate reciprocating as well as honest trustees should co-operate in case of constant checking.

[8] A honest trustee will also co-operate in an isolated encounter trust game, this means that in an indefinitely repeated trust game he will co-operate regardless of a trigger strategy.

In table 3 we summarise all possible combinations of δ and ρ for the CT.

	Trustee		
Trustor	Honest $\rho > \frac{A_2 - C_2}{E(C_1) - E(S_1)}$	Selfish $\rho = 0$	Moderate reciprocate $\frac{A_2 - C_2}{E(C_1) - E(S_1)} > \rho > 0$
Altruistic $\rho > \frac{N_1 - E(S_1)}{A_2 - N_2}$	I (Trust, Honour)	II (Trust, Abuse)	II (Trust, Abuse)
Normal $\rho < \frac{N_1 - E(S_1)}{A_2 - N_2}$	All values of δ III Trigger strategy	$\delta > \delta^*$ III Trigger strategy	$\delta > \delta^*$ III Trigger strategy
		$\delta < \delta^*$ IV (No Trust, Abuse)	$\delta < \delta^*$ IV (No Trust, Abuse)

Table 3. Equilibrium strategies depending on ρ and δ.

3.2 Indefinitely repeated asymmetric trust game without checking

Equilibria I and II still hold in the indefinitely repeated TG without checking. However the trigger strategy described above, is not any longer applicable. This strategy is based on punishment by the trustor as soon as he discovers abuse. In the asymmetric trust game it is not immediately clear when this is the case. Closely related to our game of asymmetric information is the model of Green and Porter (1984) and later applications such as Ivaldi et al. (2003). We will use the latter as the starting point of our analysis. The trustee cannot see abuse, but he does know that it is more likely to observe L in case of abuse. Therefore the most strict trigger strategy he can play is not accepting a single L. The trustee will co-operate when his expected utility of co-operation is larger then his expected utility of abuse. Where V_1 is the expected profit when the response to the trigger is co-operation:

$$V_1 = C_2 + \rho E(C_1) + (1 - p_h)\delta V_1 + p_h \sum_{t=2}^{\infty} \delta^{t-1}(N_2 + \rho N_1) \qquad (7)$$

and V_2 the expected profit when the response to the trigger is abuse:

$$V_2 = A_2 + \rho E(S_1) + (1 - p_a)\delta V_2 + p_a \sum_{t=2}^{\infty} \delta^{t-1}(N_2 + \rho N_1) \qquad (8)$$

In order to see co-operation we need $V_1 > V_2$, or:

$$\frac{(C_2 + \rho E(C_1))(\delta - 1) - \delta(N_2 + \rho N_1)p_h}{(\delta - 1)(\delta(p_h - 1) + 1)} >$$

$$\frac{(A_2 + \rho E(S_1))(\delta - 1) - \delta(N_2 + \rho N_1)p_a}{(\delta - 1)(\delta(p_a - 1) + 1)} \tag{9}$$

This leads to

$$\delta > \hat{\delta} \equiv \frac{(\widetilde{A_2} - \widetilde{C_2})}{(\widetilde{A_2} - \widetilde{C_2}) - P_h(\widetilde{A_2} - \widetilde{N_2}) + P_a(\widetilde{C_2} - \widetilde{N_2})} \tag{10}$$

Where for simplification, we denote $A_2 + \rho E(S_1)$ by $\widetilde{A_2}$, $N_2 + \rho N_1$ by $\widetilde{N_2}$ and $C_2 + \rho E(C_1)$ by $\widetilde{C_2}$.

Equilibrium III and (IV) have a new condition: when (10) is true a selfish trustee will honour. An honest trustee will always honour independent of the value of δ. If (10) is not satisfied the trustor will prefer to play no trust, above trust and then being abused.

Substituting the numbers of the experiment in equation (10) we find $\delta = 1$, for selfish trustees co-operation is no longer an attractive option, honest trustees will still co-operate. For moderate reciprocal trustees co-operation is possible when:

$$\rho > \hat{\rho} \equiv \frac{A_2(\delta(p_h - 1) + 1) - C_2(\delta(p_a - 1) + 1) + \delta N_2(p_a - p_h)}{E(C_1)(\delta(p_a - 1) + 1) - \delta(N_1(p_a - p_h) + E(S_1)(p_h - 1)) - E(S_1)} \tag{11}$$

or when $\rho > \frac{2}{21}$. Again $\frac{\partial \delta}{\partial \rho}$ is a decreasing function, thus the larger ρ the more likely co-operation is. Trustors will only play trust when they believe they are playing against a honest trustee or against a moderate reciprocal trustee with $\rho > \frac{2}{21}$.

In table 4 we summarise all possible combinations of δ and ρ for the NCT.

	Trustee		
Trustor	Honest $\rho > \frac{A_2 - C_2}{E(C_1) - E(S_1)}$	Selfish $\rho = 0$	Moderate reciprocate $\frac{A_2 - C_2}{E(C_1) - E(S_1)} > \rho > 0$
Altruistic $\rho > \frac{N_1 - E(S_1)}{A_2 - N_2}$	I (Trust, Honour)	II (Trust, Abuse)	II (Trust, Abuse)
Normal $\rho < \frac{N_1 - E(S_1)}{A_2 - N_2}$	All values of δ III Trigger strategy	$\delta > \hat{\delta}$ III Trigger strategy	$\delta > \hat{\delta}$ III Trigger strategy
		$\delta < \hat{\delta}$ IV (No Trust, Abuse)	$\delta < \hat{\delta}$ $\rho > \hat{\rho}$ III Trigger strategy $\rho > \hat{\rho}$ IV (No Trust, Abuse)

Table 4. Equilibrium strategies depending on ρ and δ.

In the literature it has been suggested that co-operation is also possible when players use a trigger strategy with limited punishment. In Ivaldi et al. (2003)

the trustor punishes the trustee only for several rounds upon observing L and he will return to co-operation afterwards until he observes the next L, then he will punish again and afterwards go back to co-operation, etc. So every time he is confronted with L he will punish for several rounds and return to co-operation afterwards. The trustee will only co-operate when $V_1 > V_2$, where D reflects the number of punishment rounds, where V_1 and V_2 are given as follows:

$$V_1 = C_2+\rho E(C_1)+(1-p_h)\delta V_1+p_h\left(\left(\sum_{t=2}^{D+1}(N_2+\rho N_1)\delta^{t-1}\right)+\delta^{D+1}V_1\right) \quad (12)$$

$$V_2 = A_2+\rho E(S_1)+p_a\left(\left(\sum_{t=2}^{D+1}(N_2+\rho N_1)\delta^{t-1}\right)+\delta^{D+1}V_2\right)+(1-p_a)\delta V_2 \quad (13)$$

Rearranging leads to:

$$\frac{(N_2+\rho N_1)p_h\delta^{d+1}+(C_2+\rho E(C_1))(\delta-1)-\delta(N_2+\rho N_1)p_h}{(1-\delta)(p_h\delta^{d+1}+\delta(1-p_h)-1)} >$$
$$\frac{(N_2+\rho N_1)p_a\delta^{d+1}+(A_2+\rho E(S_1))(\delta-1)-\delta(N_2+\rho N_1)p_a}{(1-\delta)(p_a\delta^{d+1}+\delta(1-p_a)-1)} \quad (14)$$

The more lenient the trustor becomes, the larger the right hand side (RHS) of (14) will become in comparison to the left hind side (LHS). This makes intuitively sense, because the trustee will be able to earn more from abusing payoffs, which are by definition larger then the payoffs from honouring. Given our payoffs, selfish trustees do not co-operate in case of infinite punishment. Neither will they co-operate in case of punishment for several rounds, because abuse only becomes more rewarding. Honest trustees always co-operate, thus also in this situation. The more interesting group are the moderate reciprocals. Here two effects are at play: the higher ρ the more willing the moderate reciprocals become to co-operate, because the LHS of (14) will become larger compared to the RHS. On the other hand the lower D, the larger the RHS will become. For higher values of ρ co-operation can be maintained with only 1 round of punishment. For lower values of ρ more rounds are required.

3.3 Crowding out and checking

As has been suggested in section 1 checking might have a crowding out effect. Before we will discuss the consequences of checking we first will discuss crowding out in more detail. Crowding theory started with intrinsic motivation, this is where we start as well, in order to continue to the crowding out effect checking might have on another behavioural motivation, namely trustworthiness. In our model trustworthiness is represented by ρ. We will claim that checking can have a crowding out effect on altruistic and reciprocal preferences that allow for trustworthy behaviour.

Titmuss (1970) was the first who mentioned the basic idea of crowding out. He argued that people's willingness to donate blood would be reduced when they would get paid for doing so. Crowding out, which is also known as the hidden cost of reward or control, is however mainly developed by cognitive social psychologist, see for example Deci summarised in Deci and Ryan (1985) and Lepper and Greene (1978). Frey (1997) makes a distinction between two psychological processes that can account for crowding out:

1. "When individuals perceive the external intervention to be 'controlling' in the sense of reducing the extent to which they can determine actions themselves, intrinsic motivation is substituted by extrinsic control.
2. An intervention from the outside undermines the actor's intrinsic motivation if it carries the notion that the actor's intrinsic motivation is not acknowledged."

In the economic literature, this materializes in two basic mechanisms that are responsible for crowding out intrinsic motivation: First, crowding out can be due to a perceived loss of autonomy by the agent. For instance, when the principal limits the choices he can make the agent's intrinsic motivation is replaced by extrinsic control, which can potentially lead to a lower effort level of the agent. See for example Falk and Kosfeld (2004). Second, the introduction of a (financial) reward or punishment system may decrease intrinsic incentives. The precise effect of the reward depends on the level of intrinsic motivation. For example, when people are asked to perform a certain task without getting paid for it, they might still perform it, because they are intrinsically motivated. Next, an extrinsic reward is introduced that crowds out the intrinsic motivation. If it is not high enough to compensate for the loss of intrinsic motivation, the effort level chosen by people might drop below the original level because people feel that their intrinsic motivation is not recognised.

Frey (1997) obviously relates each of his psychological processes, which account for crowding out, with one of the mechanisms of control and reward: a control system reduces the extent to which individuals can determine their actions themselves, whereas an extrinsic reward scheme deters intrinsic motivation by not acknowledging it. In our setting, checking is a weak form of control, and unlike strong control that can crowd out trustworthiness by replacing it with external force, we think it is possible that checking can crowd out trustworthiness by not acknowledging it. Checking can be perceived as a contradicting signal. The trustor indicates that he trusts the trustee by placing trust, but later on he signals distrust by checking the trustee. The trustee wants to respond to the trustor's kindness, placing trust. When the trustor checks the trustee's kind response, his trustworthiness (ρ), is not acknowledged. When checking crowds out the trustee's trustworthiness, ρ should become negative. The trustee's trustworthiness has changed to spitefulness. Like Rabin (1993) says "If somebody is being nice to you, fairness dictates that you be nice to him. If somebody is being mean to you, fairness allows- and vindictiveness dictates- that you be mean to him." The trustee feels better when he can punish the trustor for not recognising his trustworthiness. While a selfish trustee only cares about himself, a trustworthy trustee cares about the trustor and the intention behind his actions. Fair behaviour must be rewarded but unfair behaviour punished. When do to checking $\rho = 0$, the trustworthy trustee would have become selfish. Because it is optimal even for selfish trustees to honour trust in the checking treatment we should only observe less honour, and hence crowding out, when $\rho < 0$.

3.4 Hypotheses

We start by looking at the NCT, where we assume for the moment that both trustor and trustee only have selfish preferences. We come to the following two hypotheses:

Hypothesis 1a: *All trustors play no trust during the NCT.*

Out of equilibrium when a trustor does play trust.

Hypothesis 1b: *All trustees play abuse during the NCT.*

Now assume social preferences do play a role. In the NCT the predictions are that selfish trustees do not co-operate. Trustors should only co-operate when they believe they are playing against honest trustees or moderate reciprocates, with $\rho > \frac{2}{21}$. Trustors might prefer to play the trigger strategy which only punishes the trustee for several rounds, because this will lead to an higher income. Trustors who believe that the trustee's ρ is large will punish for a shorter number of rounds. When trustors believe trustees have social preferences hypothesis 1a should be rejected, when trustees actually display social preferences hypothesis 1b should also be rejected.

Now we take a look at the CT. Again we assume that both trustor and trustee only have selfish preferences. Given that checking is costless we can say the following for the CT:

Hypothesis 2: *All trustors play trust in the CT.*

Hypothesis 3: *All trustees play honour in the CT.*

Hypothesis 4: *All trustors always check in the CT.*

In the CT all trustors should always check and all trustees should co-operate. When we allow for social preferences trustees should still co-operate in the CT unless they are crowded out. In case hypothesis 3 is rejected it can be interested to make a distinction between the different types of trustees. When we assume honest and moderate reciprocal trustees do react negatively on checking, in other words their trustworthiness is crowded out, we can come to the following hypothesis:

Hypothesis 5: *Honest and moderate reciprocal trustees honour trust less in the CT then in the NCT.*

On the other hand checking leaves less room for trustees to abuse, because their behaviour is immediately detected. Checking thus allows for a more efficient punishment strategy, which can increase co-operation.[9]

Hypothesis 6: *Selfish trustees co-operate in the CT.*

[9] This is not crowding in. Checking does not acknowledge nor reward trustworthiness. Checking changes the game in such a way that for selfish trustees co-operation is the best response to the trigger strategy the trustor plays.

So selfish trustee who did not honour trust in the NCT should do so in the CT.

4 Results

As this is work in progress and the experiment mentioned in this paper is scheduled for March this year, we can at this moment not present an extensive data analysis. To get an impression of how checking might affect behaviour we briefly present some results based on a pilot experiment. It should be noted that the pilot experiment differs in some respects from the experiment described in section 2. This pilot experiment only consists of two treatments: the NCT and the CT. The game was finitely repeated for 20 rounds. In the experiment we use slightly different payoffs for both trustor and trustee: $A_2 = 32, C_2 = 16, N_2 = 10, E(C_1) = 16, E(S_1) = 8$ and $N_1 = 10$. The values of P_a and P_h are the same, but the values of H and L again differ slightly, namely 20 and 4 respectively. The trustor also needed to pay a small cost for checking of 1, so in the CT he earns 15 and 7 in case he checks. We used slightly different instructions in the pilot experiment than the ones that can be found in the appendices which correspond to the experiment mentioned in section 2 of this paper.

The pilot experiment consisted of 4 sessions, each session consisted of one treatment. There were two NCT with a total of 36 subjects, 20 and 16 respectively. In case of the CT the number of subjects was 30 in total, 16 and 14 respectively. The total number of subjects is 66, of whom 36 are male and 30 female. The average age of the subjects was 22.5 years and they all studied economics.

The estimations of our parameters are based on the linear probability model, where we correct for heteroskedasticity by using robust standard errors. At the moment we limit ourselves to a comparative analysis between the treatments. Hypothesis 1a states that all trustors should play no trust in the NCT and hypothesis 1b says that all trustee should play abuse. Both hypotheses should be rejected. 46% of the time trustors chose to play trust in the NCT and 48% of the time trustees chose to play honour, see table 5. In line with earlier experiments this result suggests that social preferences play a role.

Probability to play Trust	
constant	0.467***
CT	0.037
Probability to play Honour	
constant	0.476***
CT	0.087

Table 5. Differences for playing honour and trust between NCT and CT.

Now we will take a look at the behaviour in the CT. The second and third hypotheses should be rejected, the data suggest that trustors and trustees do not significantly behave differently in the CT compared to the NCT, see table 5. The trustor can only check when he places trust first. In the CT trustors played trust half of the time. A little bit more than half of the time they checked the trustee, see table 6. Hypothesis 4 should thus be rejected.

CT	Number of observations
Trustor plays no trust	149
Trustor plays trust but no check	72
Trustor plays trust and check	79
Total number of observations	300

Table 6. Total number of checks performed in the CT compared to possibility to check.

Given the fact that we did not run the type game in the pilot experiment we use gender in the remainder of this section as an indication for type. In the literature it is often argued that women are on average more trustworthy then men. The last results can be found in table 7. The data shows that men react positively to checking, they start to co-operate more. They co-operate 58% of the time in the CT compared to only 30% of the time in the NCT. The women, although they initially co-operate more then the men (+42% in the NCT), are negatively affected by checking and co-operate less in the CT. In the NCT woman honour trust 70% of the time compared to 55% in the CT. Due to this opposite effect of checking men and women are almost equally co-operative in the CT.

Probability to play Honour	
constant	0.298***
CT	0.281***
Female	0.419***
Female CT	−0.429***

Table 7. Differences in trustworthiness in NCT and CT after correction for gender.

5 Discussion

Given the fact that the data we present at the moment is produced by an experiment that does not match the theoretical framework completely it does not make sense to draw too strong conclusions. Although it seems that checking has no effect on the behaviour of the group of trustees as a whole, the data suggests that checking can increase or decrease trustworthiness when we take into account the difference between specific groups. We are optimistic to come to more elaborate conclusions when we have run the main experiment in April this 2007.

Appendices

A Introduction and type game

Before the experiment starts please read the following instructions.

Welcome to this research project! Thank you for participating.

In this experiment you will encounter three different game situations. For each game you will receive a new set of instructions. The instructions for the first game are on the back of this instruction form. Before the first game is explained we will give you some basic rules, which apply to all three game.

The basic rules:

1. **The experiment is conducted anonymously, participants will only be known as a number. Neither the researchers nor other participants will know what you decided.**
2. **It is not allowed to communicate with the other participants during the entire experiment.**
3. **In this experiment you will be asked to make decisions. In the instructions we will explain which options you will face. These options will appear on the screen of your computer. You will be asked to make a choice for one of the options.**
4. **If the instructions are unclear, your computer doesnt work properly or you want to ask a question raise your hand and one of the instructors will come to you.**
5. **During this experiment you will earn points. How much points you earn depend on the decisions you make. For every game the payoff structure will be explained. At the end of the experiment the total number of points you earned will be translated into a monetary reward. We use the following exchange rate: 1 point = 1.75 eurocent.**
6. **When the experiment is finished remain at your seat until you are given the signal that you can come forward to claim you reward. You can claim your reward by handing in the number given to you when entering the laboratory together with all three complete instructions sets.**
7. **Please turn off your mobile phone**

Game 1 (For a graphical representation see figure 1 on the next page.):

This game has two players: person A and person B.

Possible moves: Person A moves first and has two choices: 1) Left 2) Right

If person A decides to play Right person B can do nothing and the game ends.

If person A decides to play Left person B will have the following two choices: 1) Left 2) Right

After person B made his decision the game ends.

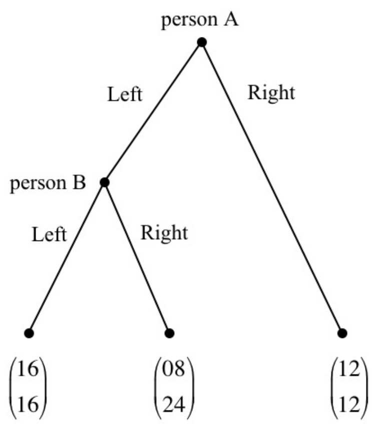
This game will be played in the following way: All participants present in the laboratory will first be given the choice of person A. Next all participants will be given the choice of person B under the assumption that person A did choose Left. In other words you are twice asked to make a decision between Left and Right once in the role of person A and once in the role of person B. After all participants made both decisions the computer will randomly assign to half of the participants the role of person A and to the other half the role of person B. Next the computer will randomly match each person A to a person B.

To calculate your payoffs the decision made by person A in the first stage of this game will be match with the decision made by person B in the second stage. In other words you will earn points depending on the role assigned to you and the other participant you are matched with and the decisions you and the other participant made in this role. For a detailed overview of the payoff structure see this table:

Assume you are assigned the role of person A	**Payoffs:**
You chose Right	You and person B both receive 12 points
You chose Left and person B chose also Left	You and person B both receive 16 points
You chose Left and person B chose Right	You receive 8 and person B receives 24 points
Assume you are assigned the role of person B	**Payoffs:**
Person A chose Right	You and person A both receive 12 points
Person A chose Left and you chose also Left	You and person A both receive 16 points
Person A chose Left and person B chose Right	You receive 24 and person A receives 8 points

Table 1. Payoff structure.

Figure 1. The game tree of game 1.



Payoffs:
person A's payoffs are in the upper row.
person B's payoffs are in the lower row.

B Trust game without checking

Before experiment continues please read the following instructions.

Game 2 (For a graphical representation see figure 2 on the next page.):

This game will be repeated, the number of repetitions is however unknown.

This game has two players: person A and person B.

Possible moves in one round: Person A moves first and has two choices: 1) Left 2) Right

If person A decides to play Right person B can do nothing and the round ends.

If person A decides to play Left person B will have the following two choices: 1) Left 2) Right

After person B made his or her decision the round ends. At the end of each round both players will receive the number of points they earned.

Although game 2 looks similar to game 1, there are some major differences. So read the remainder of these instructions carefully. Before game 2 starts the computer will randomly assign you the role of either person A or person B. Next the computer will randomly match each person A to a person B. Once matched you will play together for the entire duration of the game. After each round the game will continue with a probability of 0.95. In other words at the end of each round the chance that the game will stop is 5%. The stopping chance is constant over time; this means that regardless of how many rounds you have played the chance that there will be a next round is still 95%.

This game has the following payoff structure: If person A decides to choose Left Person B can earn either 16 points if he chooses Left or 24 points when he chooses Right.

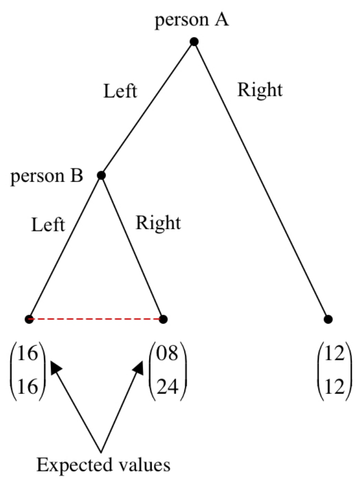
The payoffs of person A and person B are 12 points if person A chooses Right.

Person As payoffs from choosing Left are expected values. This means that person A will earn these payoffs if the game is repeated several times. After person A and person B have made their decision, the computer randomly determines person As payoff to be either 4 or 24 points according to the probabilities given in table 2. Person A will earn an expected payoff of either 16 points when person B chose Left or 8 points when person B chose Right.

At the end of each round both players will know what they have earned. Person A however will not know if person B has chosen Left or Right, because his payoff does not reveal this. This is reflected by the dotted line in figure 2.

A final remark: We will be playing this game for the coming 30 minutes. Because the number of rounds is unknown, it can be possible that the game is played more then once. When this happens you will maintain your role, but you will be matched with another person of the opposing role. The game is played until it stops, with the following limitation that the instructors will stop it when 45 minutes have past since the beginning of this part of the experiment.

Figure 2: The game tree of game 2.



Payoffs:
person A's payoffs are in the upper row.
person B's payoffs are in the lower row.

Payoffs Game 2	Person A	Person B
Left, Left	$\frac{1}{4} \times 4 + \frac{3}{4} \times 24 = 16$ points	16 points
Left, Right	$\frac{3}{4} \times 4 + \frac{1}{4} \times 24 = 8$ points	24 points
Right	12 points	12 points

Table 2: The payoff structure of game 2.

C Trust game with checking

Before the experiment continues please read the following instructions.

Game 3 (For a graphical representation see figure 3 on the next page.):

This game will be repeated, the number of repetitions is however unknown.

This game has two players: person A and person B.

Possible moves in one round: Person A moves first and has two choices: 1) Left 2) Right

If person A decides to play Right person B can do nothing and the round ends.

If person A decides to play Left person B will have the following two choices: 1) Left 2) Right

After person B has chosen either Left or Right person A has an additional decision moment; he or she can decide to check on the behaviour of person B. Person A has two choices: 1) Check (Check on the behaviour of person B) 2) No check (Do not check on the behaviour of person B)

After the final move of person A the round ends. At the end of each round both players will receive the number of points they earned.

Again game 3 looks similar to game 2, but please take into account the minor differences. You maintain your role from game 2. Before game 3 starts the computer will randomly match you with another person who has been assigned an opposing role. Once matched you will play together for the entire duration of the game. After each round the game will continue with a probability of 0.95. In other words at the end of each round the chance that the game will stop is 5%. The stopping chance is constant over time; this means that regardless of how many rounds you have played the chance that there will be a next round is still 95%.

This game has the following payoff structure: If person A decides to choose Left Person B can earn either 16 points if he chooses Left or 24 points when he chooses Right.

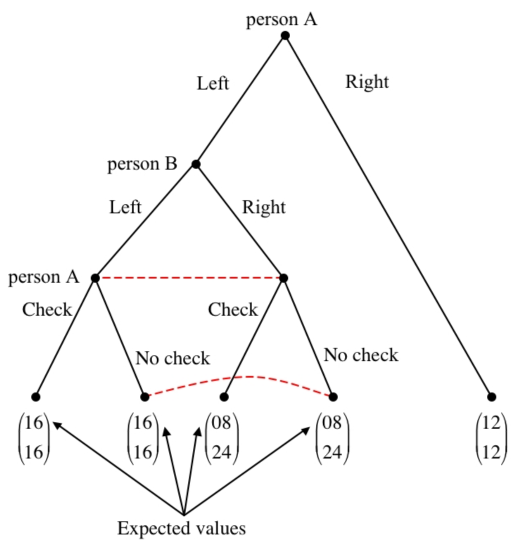
The payoffs of person A and person B are 12 points if person A chooses Right.

Person As payoffs from choosing Left are expected values. This means that person A will earn these payoffs if the game is repeated several times. After person A and person B have made their decision, the computer randomly determines person As payoff to be either 4 or 24 points according to the probabilities given in table 2. Person A will earn an expected payoff of either 16 points when person B chose Left or 8 points when person B chose Right.

At the end of each round both players will know how many points they earned. If person A did not check on the behaviour of person B he/she will not know if person B has chosen Left or Right, because his/her payoff does not reveal this. If person A decides to check on the behaviour of the person B he/she will also know what decision person B has made. This piece of information will appear in the lower right corner of your screen. If person A does not check he/she will only know his/her own payoff. The dotted lines in figure 3 depict the situations in game 3 when person A does not know what person B has chosen.

A final remark: We will be playing this game for the coming 30 minutes. Because the number of rounds is unknown, it can be possible that the game is played more then once. When this happens you will maintain your role, but you will be matched with another person of the opposing role. The game is played until it stops, with the following limitation that the instructors will stop it when 45 minutes have past since the beginning of this part of the experiment.

Figure 3: The game tree of game 3.



Payoffs:
person A's payoffs are in the upper row.
person B's payoffs are in the lower row.

Payoffs Game 3	Person A	Person B
Left, Left, Check or No Check	$\frac{1}{4} \times 4 + \frac{3}{4} \times 24 = 16$ points	16 points
Left, Right, Check or No Check	$\frac{3}{4} \times 4 + \frac{1}{4} \times 24 = 8$ points	24 points
Right	12 points	12 points

Table 3: The payoff structure of game 3.

References

Alchian, A. and Demsetz, H.: 1972, Production, information costs and economic organization., *The American Economic Review* **62**(5), 777–795.

Camerer, C. and Weigelt, K.: 1988, Experimental tests of a sequential equilibrium reputation model., *Econometrica* **56**(1), 1–36.

Dasgupta, P.: 1988, Trust as a commodity., *in* D. Gambetta (ed.), *Trust: Making and Breaking Cooperative Relations.*, Oxford, Blackwell, pp. 51–72.

Deci, E. and Ryan, R.: 1985, *Instrinsic motivation and self-determination in human behaviour.*, New York, Plenum Press.

Falk, A. and Kosfeld, M.: 2004, Distrust: The hidden cost of control, CEPR Discussion Paper No. 4512.

Fischbacher, U.: 1999, z-tree: A toolbox for readymade economic experiments., IEW Working paper No. 21, University of Zurich.

Frey, B.: 1997, A constitution for knaves crowds out civic virtues., *The Economic Journal* **107**, 1043–1053.

Frey, B. and Jegen, R.: 2001, Motivation crowding theory, *Journal of Economic Surveys* **15**(5), 589–611.

Green, E. and Porter, R.: 1984, Noncooperative collusion under imperfect price information, *Econometrica* **52**(1), 87–100.

Ivaldi, M., Jullien, B., Rey, P., Seabright, P. and Tirole, J.: 2003, The economics of tacit collusion, IDEI, Toulouse.

Kreps, D.: 1990, Corporate culture and economic theory., *in* J. E. Alt and K. Shepsle (eds), *Perspectives on Positive Political Economy*, Cambridge, Cambridge University Press, pp. 90–143.

Lepper, M. and Greene, D.: 1978, *The hidden costs of reward: new perspectives on psychology of human motivation.*, Hillsdale, New York, Erlbaum.

Levine, D.: 1998, Modeling altruism and spitefulness in experiments, *Review of Economic Dynamics* **1**, 593–622.

Ploner, M.: 2006, Trust and detection: An experimental investigation of motivational crowding out, CEEL Working Paper N. 02/2005.

Prendergast, C.: 1999, The provision of incentives in firms., *Journal of Economic Literature* **37**(1), 7–63.

Rabin, M.: 1993, Incorporating fairness into game theory and economics, *The American Economic Review* **83**, 1281–1302.

Titmuss, R.: 1970, *The Gift Relationship.*, London, Allen and Unwin.

Williamson, O.: 1993, Calculativeness, trust, and economic organization, *Journal of Law and Economics* **36**(1), 453–486.