THE CENTRE FOR MARKET AND PUBLIC ORGANISATION

Centre for Market and Public Organisation
Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol BS8 1TX

Tel: (0117) 33 10799
Fax: (0117) 33 10705
E-mail: cmpo-office@bristol.ac.uk

Efficiency in the Trust Game: an Experimental Study of Preplay Contracting

Juergen Bracht and Nick Feltovich

August 2006

Working Paper No. 06/154

ISSN 1473-625X

CMPO Working Paper Series No. 06/154

Efficiency in the Trust Game: an Experimental Study of Preplay Contracting

Juergen Bracht[1]
and
Nick Feltovich[2]

[1]*Department of Economics, University of Aberdeen Business School*
[2] *Department of Economics, University of Houston*

August 2006

Abstract
We use a human-subjects experiment to test the effects of a simple mechanism designed to increase cooperation and efficiency in the *trust game*. In the equilibrium of the standard trust game, the investor does not invest, foreseeing that the allocator would have kept all of the returns from investment. Our mechanism adds a preplay *escrow* stage, in which the allocator places an amount (possibly zero) into escrow, to be forfeited if he keeps the proceeds of investment for himself. In the experiment, we vary the amounts that can be put into escrow. Our baseline treatment has no escrow. In a second treatment, only low escrow choices are possible, so the equilibrium is unaffected. In our third treatment, there is an escrow amount high enough that, in equilibrium, investment and sharing of the proceeds will occur. Two additional treatments mirror our second and third, except that in these, the escrow amount is randomly chosen and imposed on the allocator.
We find that the high escrow amount, when chosen, does lead to the predicted efficient result. Contrary to the equilibrium prediction, we also find substantial investment in both the baseline and "low-escrow" treatments, leading to markedly higher efficiency than predicted (albeit well below that when the high amount is chosen). Over time, however, cooperation and efficiency after low or zero escrow amounts decline. We find only weak evidence for "crowding-out", which predicts that given a low or zero amount placed into escrow in non-baseline treatments, investment and efficiency would actually be lower than in the baseline. We also find that initially, investment is more likely after allocators place the maximum possible amount into escrow – as if this action by allocators is being (mistakenly) read by investors as a signal that allocators plan to share. All of these results are seen when escrow choices are imposed as well as when they are voluntary.

Keywords: experiment, trust game, incentives, signal, crowding out

JEL Classification: C72, D82, A13

Acknowledgements
[2]Corresponding Author. Financial support from the University of Aberdeen, the University of Houston, and the British Academy is gratefully acknowledged. We thank Dieter Balkenborg, Tilman Börgers, Jim Engle-Warnick, David DeMeza, Steffen Huck, Oliver Kirchkamp, Nat Wilcox and Rick Wilson for helpful comments and suggestions. Any remaining errors are a result of bad incentives.

Address for Correspondence
CMPO, Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol
BS8 1TX
Juergen.bracht@abdn.ac.uk *or* nfelt@mail.uh.edu
www.bris.ac.uk/Depts/CMPO/

CMPO is jointly funded by the Leverhulme Trust and the ESRC

1 Introduction

The economics and game theory literatures teem with examples of group decision situations where self–interested behavior by individuals leads to outcomes that are inefficient from the perspective of the group. The prisoners dilemma (Flood (1952)), the tragedy of the commons (Hardin (1968)), and the market for lemons (Akerlof (1970)) are models of three such situations; these three are so well–known as to have crossed over into non–academic discourse. Many other such situations exist. Because these situations are so common, there has been some effort to theoretically study mechanisms aimed at improving efficiency in these situations. There have been relatively few empirical tests of such mechanisms, however.

The goal of this paper is to empirically examine the effects of a mechanism designed to improve efficiency in one particular situation. The situation we use is the *trust game* (Berg, Dickhaut, and McCabe (1995)), a simple collective–action game played between two players. One player—who will be referred to as the *investor*—has the choice of either investing or not investing in a project, which is administered by the other player—who will be referred to as the *allocator*. With certainty, the investment is successful, in the sense that the amount invested multiplies in value.[1] However, the allocator controls the proceeds of investment: he can either keep them for himself or share them with the investor.

The trust game is often used as metaphors for more complicated social situations. The amount of investment observed in this game is a measure of the amount of *trust* investors have in allocators. The portion of investment proceeds given back to investors measure the amount of *trustworthiness* of allocators. Under this interpretation, the prediction of game theory is dismal indeed: the unique subgame perfect equilibrium of this game has the investor refusing to invest, forseeing that the allocator would keep the entire proceeds of any investment.[2] This equilibrium is inefficient; total payoffs are higher if the investor invests, and indeed it is possible for the allocator to split the proceeds of investment in such a way as to make both players strictly better off than in equilibrium. However, in this simple game, there are no binding contracts, nor any other way for the allocator to credibly commit to share the proceeds rather than keep them all.

The mechanism we consider is relatively simple. We add a pre–play stage to the trust game, in which the allocator has the opportunity to place an amount of money into an escrow account, to be returned to him if he shares the proceeds from investment, but forfeited if he keeps them for himself. (The escrow amount is also returned to the allocator if the investor does not invest.) If the allocator places a large enough amount into escrow, he will have an incentive to share instead of keeping, as the loss of the escrow amount outweighs the gain from keeping. In this case, the mechanism is predicted to achieve an efficient (and equitable) outcome.

In order to examine the effects of this mechanism, we design and run a human–subjects experiment that looks at two versions of the escrow game, differing sharply in subgame perfect equilibrium predictions.

[1] In our treatment, the investment quadruples. Other versions have the investment doubling or trebling.

[2] Throughout this paper, unless otherwise stated, we will use terminology such as "theoretical prediction", "equilibrium prediction", or "prediction of game theory" to mean the combination of appropriate equilibrium concepts (usually subgame perfect equilibrium) and the assumption that players' preferences concern only their own monetary payoffs. We acknowledge that this is an abuse of terminology, as game theory itself makes no assumptions about what form preferences may take, and if players' preferences concern non–monetary aspects, the true equilibrium predictions may be different.

In one, it is possible for the allocator to choose a large enough escrow amount to achieve efficiency. In a second version, escrow is possible, but the amount is not large enough to be a credible commitment by the allocator, so investment should not occur. We compare the results of these games to those of three other games: a Control treatment in which escrow is not an option (that is, a basic trust game); and two "forced escrow" games, in which the escrow decision is not made by the allocator, but rather imposed on him by the experimenter.

Our primary source of hypotheses for the effects of our mechanism is standard game theory. Its predictions are simple. When the large amount is put into escrow, efficiency is high, as the investor invests. Also, the allocator splits the proceeds of investment in this case. When either the small amount or nothing at all is put into escrow, the result is the same as in the basic trust game: the investor does not invest (and if she did, the allocator would keep the proceeds), so efficiency is low. These predictions are unaffected by whether escrow decisions are voluntary or forced, and also by which escrow choices are possible.

Many experimental researchers have found that behavioral theories (other–regarding preferences, bounded rationality, or a combination of the two) can characterize aspects of decisions that standard game theory cannot. So, in addition to standard game theory, we examine two behavioral sources of hypotheses. According to "crowding out" (Ostrom (2000)), individuals' intrinsic tendency toward cooperative behavior is damaged by mechanisms providing financial incentives for such behavior. As a result, a mechanism that provides weak financial incentives (too small to change the monetary best responses) would actually result in less cooperation, and thus lower efficiency, than if there had been no mechanism at all—in contrast to the equilibrium prediction of no difference. We also considered a "signaling" theory, according to which a choice by the allocator of the largest possible escrow amount can be taken as a signal that the allocator intends to split the proceeds of investment—even if this largest escrow amount is too small to change the allocator's monetary incentives after investment. If this is true, then behavior following a given escrow amount will depend to some extent on which other amounts were permitted; specifically, investment (if investors interpret this behavior as signals) and splitting (if allocators actually are signaling) will be higher when the escrow amount is the highest possible, and voluntary rather than forced, than when either of these is not true.

Our results are largely in line with standard game theory. When the large amount is placed by allocators into escrow—irrespective of whether it was chosen or forced—high levels of efficiency result, as investors generally invest in this case, correctly anticipating that allocators will split the proceeds with them. Since not only investors, but also allocators, earn high payoffs compared to the no–investment outcome, it is not surprising that when allocators do have the option of the large escrow amount, they nearly always choose it. On the other hand, when escrow is not possible at all, or when only a low escrow amount is possible, allocators are much less likely to split the proceeds of investment, and the frequency of their doing so declines over time. Perhaps in response, investment in these cases also declines over time, from initial levels comparable to the high–escrow case to final levels much lower, in some cases even zero. We do find higher levels of investment and splitting following the low escrow amount than following a zero escrow amount, which is inconsistent with the theory, though for allocators, these differences die out over time.

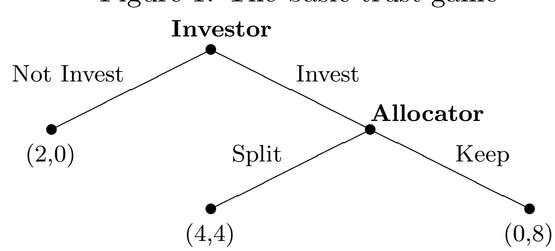
On the other hand, our results show little support for the behavioral theories; at best, they describe some aspects of early–round decisions, which go away as the experiment progresses. By and large, crowding out does not occur. If we consider a weakened version of the crowding–out hypothesis, restricting ourselves to behavior following only the zero escrow amount (not the low amount), then investment and splitting are indeed less frequent in the escrow treatments than in the Control treatment; however, this effect is compensated for by the higher levels of investment and splitting following the low escrow amount, making the net effect insignificant most of the time. As for our signaling hypotheses, we do find that in early rounds, investment following an allocator's choice of a low escrow amount is substantially higher when that was the highest amount possible than when a higher amount was possible, and by the same token, investment following nothing put into escrow was initially higher when escrow was not possible than when it was. However, we do not find the same differences in allocators' subsequent decisions, suggesting that the investors' interpretation of the escrow decision as a signal is mistaken. Investors seem to eventually figure this out, so that the effect dies out over time.

The rest of the paper proceeds as follows. In Section 2, we discuss the basic trust game, the mechanism, and the associated predictions from standard game theory and behavior game theory. In Section 3, we describe the procedures used in the experiment. In Section 4, we list the experimental results and compare these results to our hypotheses. Finally, in Section 5, we summarize our main results and discuss their implications.

2 Theories and implications

The basic game is shown in Figure 1. The investor has two units of money that she can either invest or not invest. (Investing a partial amount is not possible.) If she does not invest, the game ends, she keeps her money, and the allocator gets nothing. If she invests, her money quadruples in amount and becomes property of the allocator, who may either split it evenly with the investor or keep it all for himself. Under the assumption that both players' payoffs are identical to their monetary earnings, this game has a unique subgame perfect equilibrium in which the investor does not invest because she correctly foresees that, if she does, the allocator returns nothing to her.[3]

Figure 1: The basic trust game



Many researchers have studied experimentally the trust game and related games (such as the labor–market games with incomplete contracts examined by Fehr, Kirchsteiger, and Riedl (1993)). The main

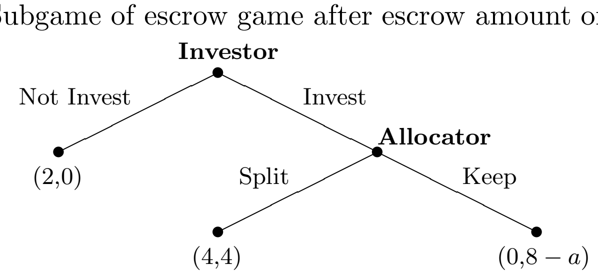
[3]This game, as well as the game described in the next section, also has non–subgame–perfect Nash equilibria. These equilibria are all equivalent along the path of play: in all of them, the investor does not invest, and the allocator keeps with high enough probability that it does not pay the investor to invest.

results of these experiments have been as follows: investors often invest with nonnegligible positive frequency; allocators return a positive amount with nonnegligible frequency (though many return nothing, as theory predicts); when multiple investment and return amounts are possible, returns by allocators tend to increase with the amount invested; and amounts returned average about the same as, or somewhat below, amounts invested (so that investors typically do not earn a profit by investing). These qualitative results have been replicated in different countries (e.g., Barr (2003) in Zimbabwe); using abstract language or various types of context; with binary choices (that is, invest either all or nothing) or nearly–continuous choices; with the investment amount being multiplied by varying amounts; and with many different stake sizes, including very large stakes (e.g., Johansson–Stenman, Mahmud, and Martinsson (2004), who used a stake size of two weeks' income) and the case of hypothetical payments (Holm and Nystedt (2004)). Researchers have also examined relationships between investor and/or allocator behavior and many possible correlates, including measures of risk attitudes from surveys or lottery–choice problems (Eckel and Wilson (2004a), Bohnet and Zeckhauser (2004)), measures of trust attitudes from surveys or other games (Glaeser et al. (2000)), measures of social distance (Glaeser et al. (2000)), demographic features of subjects or their opponents (Fershtman and Gneezy (2001), Eckel and Wilson (2002, 2004b), Croson and Buchan (1999)), opponents' facial expressions (Scharlemann et al. (2001), Eckel and Wilson (2004b)), political ideology and political–party identification (Anderson, Mellor, and Milyo (2004)), finite versus indefinite repetition of the game (Engle–Warnick and Slonim (2004)), experience playing the game (Engle–Warnick and Slonim (2004)), whether the game is in normal– or extensive–form representation (Deck (2001)), and elicitation and/or transmission of beliefs about opponent's play (Guerra and Zizzo (2004)).

2.1 The escrow mechanism

Because the subgame perfect equilibrium of the basic trust game is not efficient, it is of interest to examine small modifications of the game that may lead to increased efficiency. We consider the following variation, which we call an "escrow game". Before the investor's decision, the allocator has the opportunity to place a nonnegative amount $a \in A$ in "escrow" (where the nonempty set A of allowable escrow amounts is our treatment variable). This amount is forfeited if the investor invests and the allocator keeps the returns; it is returned to the allocator otherwise. Figure 2 shows the subgame that results *after* the allocator's

Figure 2: Subgame of escrow game after escrow amount of a is chosen



decision a is made. The nature of the subgame perfect equilibrium of this subgame depends on the size of a. If $a < 4$, the escrow amount is small enough that the allocator would still keep the returns from any investment (forfeiting the escrow amount) rather than splitting them, so that the investor will not invest.

In this case, subgame perfect equilibrium predicts no behavioral difference between this subgame and the original game shown in Figure 1. If $a > 4$, on the other hand, the escrow amount is large enough that the allocator would rather split the returns from investment, so that the investor will invest. In this case, the subgame perfect equilibrium of this subgame is efficient.

The subgame perfect equilibrium of the entire game (the allocator choosing some $a \in A$, then play of the resulting subgame shown in Figure 2) depends on the elements contained in A. If A contains only amounts less than 4, then regardless of the allocator's choice, the investor will not invest (since the allocator would keep any returns from investment). Since any escrow choice by the allocator leads to the same payoffs, subgame perfect equilibrium doesn't predict which escrow choice he will make. If A contains any amount(s) greater than 4, however, the allocator will choose some such amount, the investor will invest, and the allocator will split. Since any escrow choice above 4 gives the allocator the same payoff, subgame perfect equilibrium again does not predict which such choice he will make.

2.2 Experimental design and hypotheses

In our experiment, we consider five treatments, differing in A—the set of possible a—as well as how the escrow decision is made. In our control treatment, no escrow is possible, so that $A = \{0\}$. In our "Escrow03" treatment, we set $A = \{0, 3\}$, and in our "Escrow036" treatment, $A = \{0, 3, 6\}$. In addition, there are two "forced escrow" treatments, where a third party (the computer program) determines the escrow amount, rather than the allocator making the choice. These treatments parallel our Escrow03 and Escrow036 treatments; in our "Forced03" treatment, we again have $A = \{0, 3\}$, and in our "Forced036" treatment, $A = \{0, 3, 6\}$. (We will sometimes refer to the Escrow03 and Escrow036 treatments as our "voluntary escrow" treatments in contrast.) A summary of the treatments and corresponding subgame perfect equilibrium predictions is shown in Table 1; also shown is a measure of efficiency, which we define as the sum of investor and allocator payoffs, normalized so that 0 and 1 represent the minimum and maximum efficiencies. Notice that from a game–theoretic standpoint, the only determinant of investment and splitting is the amount placed into escrow; whether the escrow choice was voluntary or forced does not matter, nor does the existence of larger or smaller alternative escrow choices.

The pre–play escrow stage is an example of a mechanism designed to improve efficiency. As discussed above, according to subgame perfect equilibrium, this mechanism should work if and only if an escrow amount larger than 4 is chosen. This is possible in the Escrow036 and Forced036 treatments, but not in the Control treatment (where there is no escrow at all), nor in the Escrow03 and Forced03 treatments (where the escrow amounts are too small). This implies that when the escrow amount is 6, investors will choose Invest and allocators will choose Split, whereas when the escrow amount is 0 or 3, they will not. This implies the following hypotheses.

Hypothesis 1 *The frequency of Invest will be higher following an escrow amount of 6 than following an escrow amount of 0 or 3.*

Hypothesis 2 *The frequency of Split will be higher following an escrow amount of 6 than following an escrow amount of 0 or 3.*

Table 1: Treatments and game–theoretic predictions

Treatment	Escrow amount	Probability chosen	Conditional Prob(Invest)	Conditional Prob(Split)	Efficiency
Control	0	1	0	0	0
Escrow03	0	*	0	0	0
	3	*	0	0	0
	0	0	0	0	0
Escrow036	3	0	0	0	0
	6	1	1	1	1
Forced03	0	—	0	0	0
	3	—	0	0	0
	0	—	0	0	0
Forced036	3	—	0	0	0
	6	—	1	1	1

*: *Either escrow amount can be chosen in subgame perfect equilibrium.*

Hypothesis 3 *The frequency of Invest will be the same following an escrow amount of 0 as following an escrow amount of 3.*

Hypothesis 4 *The frequency of Split will be the same following an escrow amount of 0 as following an escrow amount of 3.*

While the game–theoretic prediction of the impact of the escrow mechanism is clear, there is a good amount of evidence from the experimental economics literature suggesting that the actual impact may be different. Ostrom (2000) summarizes a large body of research on collective–action problems (including the trust game), and finds several empirical regularities. Two of these regularities are (1) in situations like our basic trust game, levels of cooperative behavior are substantially higher than would be predicted by game theory, but (2) when rules are added to the game in an attempt to motivate cooperative behavior, people act approximately as game theory predicts. Together, these results imply that "externally imposed rules tend to 'crowd out' endogenous cooperative behavior" (p. 147).[4]

This "crowding out" hypothesis has implications for the games used in our experiment. Our Control, Escrow03 (and Forced03), and Escrow036 treatments correspond, respectively, to the three cases that can occur: (1) no externally–imposed rules, (2) weak externally–imposed rules, and (3) strong externally–imposed rules. In our Control treatment—where no external rules are imposed—levels of investment and splitting (cooperative play) ought to be substantially higher than game theory predicts. When rules are

[4]Lazzarini, Miller, and Zenger (2004) discuss some of the more recent research on crowding out, including experiments in which crowding out did not occur. They also present the results of their own experiment, in which crowding out does not occur. They conclude that under some circumstances, formal mechanisms can actually be *complements* to informal social norms, rather than *substitutes*, as crowding out implies.

imposed, levels of investment and splitting ought to be similar to the game–theoretic prediction, but the prediction itself will depend on the strength of the rules. In our Escrow036 treatment, the rules are strong enough to make cooperative behavior rational (in the sense of maximizing monetary payoffs), so there should be high levels of investment and splitting. In our Escrow03 and Forced03 treatments, where rules are in place, but they are not strong enough to make cooperative behavior rational, levels of investment and splitting should be as game theory predicts. Noting that the game–theoretic prediction for the Escrow 03 and Forced03 treatments is the same as that for the Control treatment, and that as mentioned above, actual levels for the latter should be higher than the game–theoretic prediction, the implication is that levels of investment and splitting should actually be even less in the Escrow03 and Forced03 treatments than in the Control treatment. In the Forced036 treatment, externally–imposed rules are either strong or weak, depending on whether the escrow amount imposed is 6 or less than 6. In either case, levels of investment and splitting should be as game theory predicts: high (as in the Escrow036 treatment) when the escrow amount is 6, and low (as in the Escrow03 and Forced03 treatments) when the escrow amount is either 0 or 3.

Thus, the "crowding out" hypothesis implies that Hypotheses 1 and 2 above should still hold, but Hypotheses 3 and 4 should be replaced by:

Hypothesis 5 *The frequency of Invest will be higher in the Control treatment than in each of the other treatments following an escrow amount of 0 or 3.*

Hypothesis 6 *The frequency of Split will be higher in the Control treatment than in each of the other treatments following an escrow amount of 0 or 3.*

An alternative "signaling" theory makes almost the opposite prediction. According to this theory, allocators who intend to Split will signal their cooperative intention by placing the maximum possible amount into escrow, thus making it more costly to Keep later (if the investor invests). In the Escrow036 treatment, of course, this is no different from what equilibrium predicts. If allocators are signaling with their escrow choices, however, then even in the Escrow03 treatment, an escrow choice of 3 should increase cooperative behavior: investors will anticipate that allocators intend to choose Split, so they will choose Invest.

This reasoning implies that other things equal, cooperative behavior should be more likely when the escrow amount chosen by the allocator was the largest escrow amount possible—and of course, that this amount was actually chosen by the allocator, not imposed externally. This leads to the following hypotheses (in addition to Hypotheses 1 and 2 above):

Hypothesis 7 *In the Escrow03 treatment, the frequency of Invest will be higher following an escrow choice of 3 than following an escrow choice of 0.*

Hypothesis 8 *In the Escrow03 treatment, the frequency of Split will be higher following an escrow choice of 3 than following an escrow choice of 0.*

Hypothesis 9 *Following an escrow choice of 3, the frequency of Invest will be higher in the Escrow03 treatment than in the Escrow036, Forced03, or Forced036 treatments.*

Hypothesis 10 *Following an escrow choice of 3, the frequency of Split will be higher in the Escrow03 treatment than in the Escrow036, Forced03, or Forced036 treatments.*

Several other researchers have looked experimentally at mechanisms for improving outcomes in games where the equilibrium is inefficient. Many of these mechanisms have been designed specifically for collective–action problems. Andreoni and Varian (1999) examine the ability of a "compensation mechanism" (analyzed by Varian (1994)) to facilitate cooperation. In their experiment, subjects play 15 rounds of an asymmetric prisoners' dilemma, then 25 rounds of a modified version of the game in which, prior to play, each player chooses how much to offer her opponent in exchange for cooperating, and then each player is told what she has been offered to cooperate. This mechanism changes the equilibrium prediction to one in which both players cooperate (along with suitable compensations in the pre–play stage). However, Andreoni and Varian's results were mixed. They did find, as predicted, that cooperative behavior was higher under the mechanism than without it. However, the increase in cooperative behavior was much smaller than predicted: cooperation was higher than predicted without the mechanism, but lower than predicted with the mechanism. It is not clear why cooperation under the mechanism was so low; even following choices large enough to induce cooperation in equilibrium, the frequency of cooperation was only about 70%. It is worth keeping their results in mind, as their mechanism is similar in nature to the escrow mechanism we use (the main difference is that their mechanism involves players making decisions that change each other's incentives, while ours involves a player changing his own incentives).

Houser et al. (2004) designed a trust game experiment involving a mechanism that is similar in some ways to Andreoni's. Along with choosing an investment amount, investors choose a desired amount to be returned to them by allocators, and threaten punishment if the allocator returns less than that amount. (Thus, they also have players making decisions that change their counterparts' incentives, but in this case, they are punishments rather than rewards.) Their results are consistent with crowding out. When no sanctions were threatened, allocators typically returned a positive amount—though less than the investor requested. When sanctions were threatened, allocator behavior depended on the severity of the sanctions: strong sanctions led to allocators returning the amount requested (though not more), while weak sanctions often led to nothing at all being returned. Interestingly, their results were robust to whether the threat was made by the investor or randomly by the experiment computer program.

Falkinger et al. (2000) considered a "tax–subsidy mechanism" (proposed by Falkinger (1996)) in which a third party—such as a government—sets a tax/subsidy rate before the players choose their contributions toward a public good. After contributions are chosen, players are rewarded or fined according to their deviation from the mean contribution level; players making above–average contributions are rewarded proportionally to how far above average their contributions were, and those making below–average contributions are punished in a similar way.[5] Falkinger et al. (2000) designed an experiment in which subjects played public–good games under several parameterizations (they varied the number of players, the production function for the public good, players' incomes and marginal rates of substitution between the public and private goods, and the tax/subsidy rate). Depending on the treatment, subjects played either 10 or 20

[5]Falkinger (2004) extends his previous model by adding an earlier stage in which players invest in an enforcement technology, which determines the effective tax/subsidy rate for the second stage.

rounds of a basic public–good game with no mechanism followed by the same number of rounds with the mechanism, or either 10 or 20 rounds with the mechanism followed by the same number of rounds without it. They found that cooperative behavior—and therefore, provision of the public good—was substantially higher with the mechanism than without it, but again, the difference was less than the equilibrium prediction. Subjects contributed much more than predicted in the basic public–good game (that is, without the mechanism), but either less than predicted (when the predicted contribution level was the player's entire income, and thus the maximum of the strategy set) or roughly as much as predicted (when the predicted contribution level was less than the player's income, and thus in the interior of the strategy set) when the mechanism was present.

Bracht, Figuières, and Ratto (2004) extended the work of Andreoni and Varian (1999) and Falkinger et al. (2000), by more directly comparing the two mechanisms studied by them. The game they used was a two–player public–good game with utility linear in the public good and concave in the private good, so that both the Nash equilibrium and the joint–payoff–maximizing outcome were in the interior of the strategy set (both for the basic game and under each mechanism). Subjects in the experiment played 20 rounds of this basic public–good game, then 20 rounds of a game with one of the two mechanisms. Bracht, Figuières, and Ratto found that both mechanisms led to increased cooperative behavior, but the increases were smaller than predicted. When there was no mechanism, contributions were consistently above the equilibrium level (though they moved toward it over time). Under the tax–subsidy mechanism, similarly, contributions started above the equilibrium level but reached it by the end of the session. Under the compensation mechanism, however, contributions started at the equilibrium level but decreased over time until ending well below it (though still higher than without the mechanism).

The results of these four experiments were largely consistent with each other, and carry two implications for us. First, the performance of these mechanisms seems to be relatively robust to small changes in experimental parameters and procedures. Second, there is substantial crowding out: while cooperative behavior is well above equilibrium levels when no mechanism is in place, under any of these mechanisms, levels of cooperation are usually no higher, and indeed are often lower, than the equilibrium prediction.

3 Experimental procedures

Our design is made up of five treatments, as listed in Table 1 above. In all experimental sessions, subjects started by playing 5 rounds of the basic trust game—with no escrow possible. This was intended to familiarize subjects with the strategic situation and the computer interface. After the first 5 rounds were completed, subjects played 10 rounds of a game that depended on the treatment. In the Control treatment, these next 10 rounds were also of the basic trust game; in the remaining treatments, these 10 rounds were of the corresponding game (for example, the Escrow03 game in our Escrow03 treatment).[6] All subjects in an experimental session were playing the same game. Each session involved 20 subjects, with the exception of one Control session that had only 18 subjects and one Escrow036 session that had only 10 subjects.

[6]Note here the distinction we draw between game and treatment in our nomenclature. Each treatment begins with 5 rounds of the basic trust game, then is followed by 10 rounds of a game—possibly the trust game again (in the Control treatment), and possibly one of the four other games (in each case, in the treatment of the same name).

Subjects were primarily undergraduate students from University College London and Exeter University, and were recruited by a variety of methods, including physically posted announcements, postings to an university experiments website, and via a database of participants in previous experiments and others expressing interest in participating in experiments. No one took part in more than one session of this experiment.

At the beginning of a session, each subject was seated in a single room and given a set of written instructions for the first five rounds.[7] At this point, the subjects were not told how (or if) the game would differ in the last ten rounds, though the instructions stated that these five rounds made up the first part of the experiment, that the second part might be different, and that the rules for the second part would be discussed after the first part ended. The instructions for the first part were read aloud to the subjects, in an attempt to make the rules of the game common knowledge. After the fifth round of a session was completed, each subject was given a copy of the instructions for the remaining ten rounds. These were also read aloud, after which the remaining ten rounds were played.

The experiment was run on networked computer terminals, using the z–Tree experiment software package (Fischbacher (1999)). Subjects were asked not to communicate with other subjects, so the only interactions were via the computer program. Subjects were randomly assigned to roles (investor or allocator) at the beginning of a session and remained in the same role throughout the session. Investors and allocators were matched using a round–robin matching format; in Rounds 1–5, each investor would be matched to each allocator at most once (and vice versa), and in Rounds 6–15, each investor would be matched to each allocator exactly once.[8] In a round of the basic trust game (either in those sessions where it was played for all 15 rounds, or in the first 5 rounds of the other sessions), investors were prompted to choose whether they would Invest or Not Invest their 2 units. After the investors' choices were entered, each allocator would see his counterpart's decision; if it was Invest, the allocator would then be prompted to choose whether he would Split or Keep. After the allocators had entered their decisions, all subjects received feedback from the just–completed round: the investor's choice, the allocator's choice (if the investor chose Invest), and the subject's own payoff. Subjects were not explicitly told their counterparts' payoffs, though they were given enough information to be able to calculate them easily if they wished. Subjects were not given any information about the results of other pairs of subjects.

In a round of either the Escrow03 or the Escrow036 game (Rounds 6–15 of the corresponding treatments), the sequence of play was similar, except for the escrow decision. In these treatments, a round would begin with allocators' being prompted to choose which of the allowable escrow amounts would be placed into escrow. Each investor would see her counterpart's decision before making her investment decision. After investment decisions were entered, allocators received this information as in the basic trust game and were then prompted to choose whether to Split or Keep. In the Forced03 and Forced036 games, the sequence of play was identical, except that allocators did not choose the escrow amounts, but rather

[7] The instructions used in the experiment, as well as the raw data, are available from the corresponding author upon request.

[8] An implication of this matching mechanism is that over a fifteen–round session, subjects would be matched with some other subjects more than once. We tried to reduce the possibility that this would lead to repeated–game effects by not telling subjects the ID number of their counterparts, so that in the last ten rounds, each only knew that with positive probability, their current counterpart was someone with whom they were matched earlier.

were informed of them at the same time investors were. The computer program chose each possible escrow amount with probability one–half in the Forced03 treatment and with probability approximately one–third in the Forced036 treatment. Subjects' feedback at the end of a round in each of the voluntary– and forced–escrow treatments was as in the basic trust game, with the addition of the escrow amount. In all treatments, at the end of a round, subjects were asked to observe their result, write the information from that round down in a record sheet, and then click a button to continue to the next round.

At the end of Round 15 of any treatment, the experimental session ended. All subjects received a £5 show–up fee.[9] In addition, one of the first five rounds and one of the last ten rounds were randomly chosen, and each subject received his/her earnings from these two rounds, at an exchange rate of £1 per point. Subjects earned an average of about £10 for participating in a session, which typically lasted between 30 and 45 minutes.

4 Experimental results

The experiment consisted of a total of fifteen sessions, three of each treatment.

4.1 Session aggregates

Some features of the aggregate data are shown in Tables 2 and 3. Table 2 shows the relative frequencies of Invest choices by investors in the first five rounds (when subjects were playing the basic trust game in all sessions), the conditional relative frequencies of Split choices by allocators (given Invest) in these rounds, and the payoff efficiency (as in Table 1, the average joint payoff, normalized so that the maximum possible efficiency is one and the minimum possible is zero).[10] This table also shows the levels of investment,

Table 2: Aggregate results from rounds 1–5 (no escrow)

	Frequency of Invest		Conditional Frequency of Split		Efficiency
Control sessions	0.567	(85/150)	0.376	(32/85)	0.567
Escrow03 sessions	0.593	(86/145)	0.442	(38/86)	0.593
Escrow036 sessions	0.448	(56/125)	0.446	(25/56)	0.448
Forced03 sessions	0.533	(80/150)	0.325	(26/80)	0.533
Forced036 sessions	0.527	(79/150)	0.228	(18/79)	0.527
All sessions	0.536	(386/720)	0.360	(139/386)	0.536

splitting, and efficiency broken down by treatment. Since subjects were playing the same game in these

[9] At the time of the experiment, £1 was worth roughly $1.80.

[10] When escrow is not possible, efficiency is therefore simply equal to the frequency of Invest choices. We list efficiencies separately so that one could easily make comparisons with treatments in which escrow is possible and efficiency therefore depends not only on the frequency of Invest choices, but also on that of Keep choices.

rounds, regardless of the treatment (differences in the game across treatments didn't begin until round 6), and at this stage had not been given any information as to how, if at all, the second part of the experiment would differ from the first, any differences observed across treatments here could be construed as being due to random variation in trust or trustworthiness across individual subjects (and perhaps other subjects reacting to this).

Behavior in the first five rounds is substantially different from the subgame perfect equilibrium prediction, as Table 2 shows: both Invest and Split do occur with nonnegligible frequency. Investors choose Invest slightly more than half the time overall. This average hides a lot of variation across sessions—levels vary from 32% to 72%—but surprisingly little variation across treatments. Allocators choose Split about 36% of the time on average over these first five rounds. There is again substantial variation across sessions—ranging from 20% to 58%—and somewhat more variation across treatments than there was for Invest.

These aggregate results—levels of investment and returns bounded well away from both zero and one—replicate those of other trust game studies. Also in line with previous results, the average amount returned to investors (conditional on investment) in each of the five treatments is somewhat below the level that would make Invest an expected–payoff–maximizing strategy for them (though there were individual sessions in which this was not true).

With the results from the first five rounds as a benchmark, we next turn to the remainder of the experimental session, where possible escrow amounts did vary across sessions. Table 3 shows the relative frequencies of Invest and Split choices, as well as efficiencies, for the last ten rounds of each treatment—both overall and broken down by the escrow amount chosen. In the Control treatment, where escrow is not possible, results are comparable to what we saw in the first five rounds. Investment happens somewhat less than half the time; when it does, allocators choose Split somewhat less than half the time (so again, investment is not profitable for investors). Efficiency is less than what it was in the first five rounds, though this decrease is small.

In the Escrow036 and Forced036 treatments, large escrow amounts are possible, and this leads to marked changes in behavior. Following an escrow choice of 6, investors invest over 90% of the time in the Escrow036 treatment and 100% of the time in the Forced036 treatment, and conditional on investment, allocators split 97% of the time in the Escrow036 treatment and 95% of the time in the Forced036 treatment. In the Forced036 treatment, allocators cannot choose the escrow amount, but in the Escrow036 treatment, where they can, they choose to put 6 into escrow over three–quarters of the time. When allocators put less than 6 into escrow, investors seldom invest, though they do invest more often following an escrow amount of 3 (23% of the time in the Escrow036 treatment and 32% of the time in the Forced036 treatment) than following a 0 escrow amount (12% of the time in the Escrow036 treatment and 8% of the time in the Forced036 treatment). Following investment, allocators split with frequency between 25% and 40%, depending on the escrow amount and whether it is forced or voluntary. Efficiency in these treatments is close to one following an escrow amount of 6, but low when the escrow amount is anything else.

In the Escrow03 and Forced03 treatments, only low escrow amounts are possible. Overall, levels of investment and splitting in these two treatments are comparable to those in the Control treatment, but this obscures differences between play after escrow amounts of 0 and play after escrow amounts of 3. Both

Table 3: Results from rounds 6–15—aggregate and conditional on escrow amount chosen

Cell	Escrow Amount	Frequency Chosen		Conditional Freq.—Invest		Conditional Freq.—Split		Efficiency
Control	0	1.000	(300/300)	0.400	(120/300)	0.408	(49/120)	0.400
	0	0.228	(66/290)	0.136	(9/66)	0.000	(0/9)	0.136
Escrow03	3	0.772	(224/290)	0.589	(132/224)	0.394	(52/132)	0.411
	Total		—	0.486	(141/290)	0.369	(52/141)	0.348
	0	0.500	(150/300)	0.207	(31/150)	0.161	(5/31)	0.207
Forced03	3	0.500	(150/300)	0.593	(89/150)	0.427	(38/89)	0.423
	Total		—	0.400	(120/300)	0.358	(43/120)	0.315
	0	0.100	(25/250)	0.120	(3/25)	0.333	(1/3)	0.120
Escrow036	3	0.140	(35/250)	0.229	(8/35)	0.375	(3/8)	0.157
	6	0.760	(190/250)	0.921	(175/190)	0.971	(170/175)	0.895
	Total		—	0.744	(186/250)	0.935	(174/186)	0.714
	0	0.333	(100/300)	0.080	(8/100)	0.250	(2/8)	0.080
Forced036	3	0.323	(97/300)	0.320	(31/97)	0.355	(11/31)	0.216
	6	0.343	(103/300)	1.000	(103/103)	0.951	(98/103)	0.951
	Total		—	0.473	(142/300)	0.782	(111/142)	0.423

investment and splitting are substantially more frequent in the latter case than in the former—in both Escrow03 and Forced03 treatments—though neither approaches the level we saw in the Escrow036 and Forced036 treatments after an escrow amount of 6. Efficiency in both of these treatments is slightly lower overall than in the control, but again, substantially higher after an escrow amount of 3 than after an escrow amount of 0. In the Escrow03 treatment, allocators choose to put 3 rather than 0 into escrow over three–quarters of the time.

These aggregate data can be summarized as follows.[11] First, the *directional* predictions of subgame perfect equilibrium describe play rather well. Whenever subgame perfect equilibrium predicts a change across or within treatments that change is seen in the data, in the direction predicted. Consistent with Hypotheses 1 and 2, both investment and splitting are far more frequent in the Escrow036 and Forced036 treatments following an escrow amount of 6 than following any other escrow amount in any treatment. However, subgame perfect equilibrium's point predictions often perform poorly; for only a few treatments and escrow amounts do we see levels of investment and splitting close to zero. (In the next section, we will see that subgame perfect equilibrium fares better as a prediction of asymptotic behavior.)

Second, these aggregate data show evidence of crowding out. Recall that crowding out implies that investment and splitting should be less frequent, and efficiency lower, when a weak mechanism is imposed (in the Escrow03 and Forced03 treatments and in the Forced036 treatment following escrow of 0 or 3) than

[11]In Section 4.3, we use parametric statistics to examine these results further.

when there is no mechanism at all (in the Control treatment). In fact, the overall frequency of investment in all of the weak–mechanism cases is 0.381, the frequency of splitting is 0.360, and efficiency is 0.273, all lower than their counterpart statistics in the Control treatment (0.400, 0.408, and 0.400, respectively). This difference is fairly substantial for efficiency, but less so for investment and splitting.[12]

Third, levels of Invest and Split depend not only on how much was put into escrow, but also on what other escrow choices were available (in contrast to the equilibrium prediction that the availability of other options should be irrelevant). In particular, we see much more investment following a given escrow decision when that was the largest possible escrow amount than when it was not. For an escrow amount of 3, this is the largest possible escrow amount in the Escrow03 and Forced03 treatments, but a larger amount was possible in the Escrow036 and Forced036 treatments. Indeed, the frequency of investment following an escrow amount of 3 is 0.589 in the Escrow03 treatment and 0.593 in the Forced03 treatment but only 0.229 in the Escrow036 treatment and 0.320 in the Forced036 treatment. For an escrow amount of 0, this is the largest possible escrow amount in the Control treatment, but a larger amount was possible in each of the other four treatments; the subsequent frequency of investment is 0.400 in the Control treatment but ranges only from 0.080 to 0.207 in the other treatments. This pattern also holds for allocators, though the differences are sometimes small. Following an escrow choice of 3 and investment, allocators choose Split only slightly more often in the Escrow03 and Forced03 treatments (0.394 and 0.427, respectively) than in the Escrow036 and Forced036 treatments (0.375 and 0.355). After an escrow choice of 0 and investment, allocators choose Split more frequently in the Control treatment (0.408) than in any of the other treatments (ranging from 0 to 0.333), though the sample sizes concerned are sometimes small.

Fourth, behavior is largely unaffected by whether escrow decisions are voluntary or forced. There are essentially no apparent qualitative differences in investment, splitting, or efficiency between the Escrow03 and Forced03 data, nor between the Escrow036 and Forced036 data, either overall or when broken down by escrow amount. While consistent with subgame perfect equilibrium, this result stands in contrast to other experimental studies which show that behavior can be sensitive to such a manipulation.[13]

4.2 Round–by–round behavior

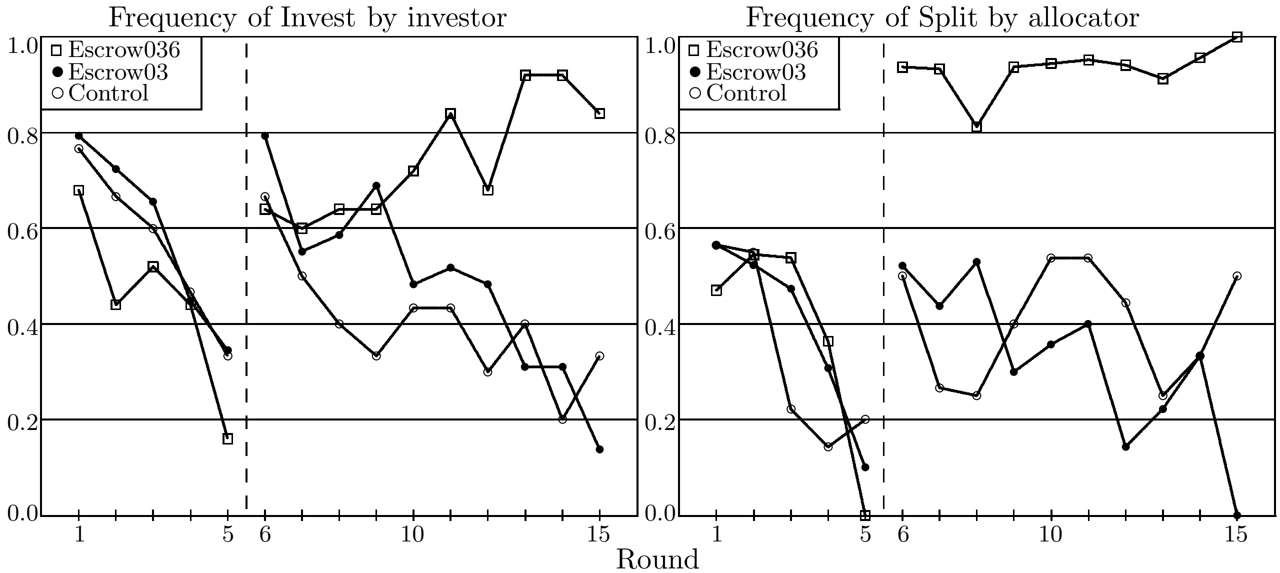
Figure 3 shows the round–by–round relative frequencies of Invest and Split for the Control, Escrow03, and Escrow036 treatments. Note that these frequencies are not broken down by escrow amount. (As mentioned earlier, sample sizes are small for escrow choices less than the maximum possible choice.) Consider first the initial five rounds, during which there is no escrow. Qualitative dynamics in these rounds are similar in all three treatments (and as we will see shortly, for the other two treatments as well). The frequency of Invest starts between about two–thirds and three–quarters, but by the fifth round has declined by half or more in

[12] One could argue that "weak mechanism" should also include those plays in the Escrow036 treatment in which the allocator chose to put 0 or 3 into escrow. Using this definition changes the weak–mechanism levels of investment, splitting, and efficiency only slightly—0.367, 0.360, and 0.264 respectively—so that again, efficiency is substantially less than in the Control treatment, while the difference is small for investment and splitting.

[13] Such sensitivity is most common when the situation is one where nonpecuniary aspects of outcomes are important, as in the trust game. For example, see Cox and Deck's (2002) results for allocators in the trust game, or Blount's (1995) results for the ultimatum game.

each of the three treatments. The frequency of Split starts at about one–half and drops reasonably steadily over these five rounds to below 20% in each treatment (and, indeed, zero in the Control treatment). Since Invest is a monetary best response for the investor only if the probability of Split is at least one–half, it appears that on average, investors are reacting rationally to their experiences of the behavior of allocators.

Figure 3: Round–by–round unconditional relative frequencies of Invest and Split
(Control and voluntary–escrow treatments)

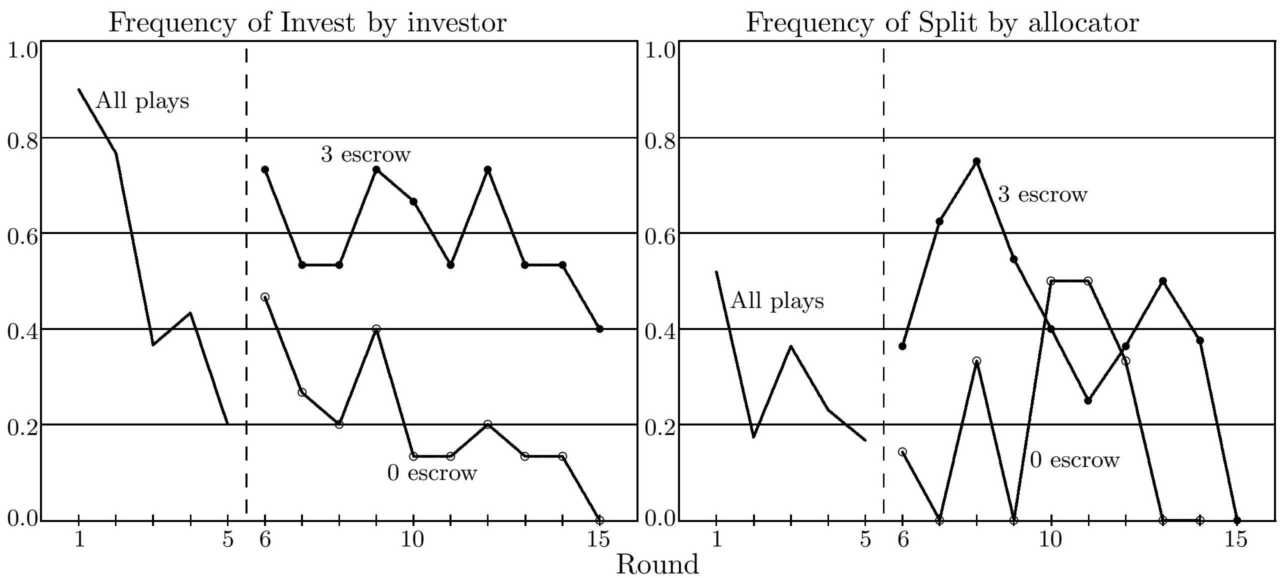


In all treatments, the levels of investment and splitting increase sharply from Round 5 to Round 6, the first round of the second part of the session. In the Escrow036 treatment, the equilibrium predictions for both investment frequency and splitting frequency go from zero in Round 5 to one in Round 6, so it is not surprising to see an increase there. In the other two treatments, the equilibrium predictions are unchanged from Round 5 to Round 6, so it is less clear what causes these increases. In the Control treatment, there is no change in the rules of the game from Round 5 to Round 6, so it is likely that the change in levels of investment and splitting is due to a "restart effect"—a change in behavior caused purely by referring to Round 6 as the first round of the second part of the session instead of one more round in the first part (see, for example, Andreoni (1988), Moxnes and van der Heijden (2003), Camerer and Fehr (2003), and Croson, Fatas, and Neugebauer (2005)). The cause of the changes in investment and splitting levels in the Escrow03 treatment may be a restart effect, or may have occurred because subjects initially perceived that some relevant aspect of the strategic environment has changed.

Dynamics in investor behavior over the last 10 rounds of the Control and Escrow03 treatments are broadly similar. Investment frequencies start out relatively high—above 60% in both treatments—but decline over time, though always remaining above zero (the equilibrium prediction). Allocator behavior differs somewhat in these two treatments; in the Control treatment, the frequency of Split varies between 20% and 60% but shows no time trend, while splitting in the Escrow03 treatment falls from about 50% to zero.

We next look at the round–by–round relative frequencies of Invest and Split in the two forced–escrow treatments. Since the theoretical predictions for these treatments depend on which escrow amount is imposed on allocators, and because there are large numbers of occurrences of each escrow amount (though sample sizes are sometimes small for allocators in cases where investment is infrequent), we disaggregate the data for Rounds 6–15 of these treatments according to the escrow amount. This is done for the Forced03 treatment in Figure 4 and for the Forced036 treatment in Figure 5.

Figure 4: Round–by–round relative frequencies of Invest and Split—Forced03 treatment (conditional on escrow amount)



Behavior in the first five rounds of the forced–escrow treatments is qualitatively similar to that in the other treatments. In both treatments, the frequency of Invest begins above 80% and the frequency of Split begins between 40% and 60%; both trend downward over the first five rounds, though there is an upward bump in Round 5 of the Forced036 treatment. In Round 6, behavior in the Forced03 treatment shows a restart effect, particularly for investment (where even after a 0 escrow amount, investment is more likely than in Round 5), though one is not apparent in the Forced036 treatment. When an escrow amount of 6 is imposed in the Forced036 treatment, Invest is chosen 100% of the time, and Split nearly 100% of the time, in all rounds. For the other escrow amounts in both treatments, investment tends to decline over time, though there is a good deal of noise in these time series. There is even more noise in the allocator data in these two treatments, due to small sample sizes in some cases, though when a time trend is apparent, it is a downward one (that is, a decline over time in the frequency of Split).

Figure 6 shows the round–by–round efficiencies for each of the four treatments. As before, these are broken down by escrow amount for the forced–escrow treatments, but not for the voluntary–escrow treatments. There is little difference across treatments in efficiency for the first five rounds (when there is no escrow in any treatment); in each treatment, efficiency falls sharply over these rounds. Over Rounds 6–15, we see a divergence across treatments, as we did for the frequencies of Invest and Split. Efficiency in the Forced036 treatment after an escrow amount of 6 is close to one in all rounds, while efficiency in the Escrow036 treatment gradually rises as investors become more likely to choose Invest and allocators

Figure 5: Round–by–round relative frequencies of Invest and Split—Forced036 treatment (conditional on escrow amount)

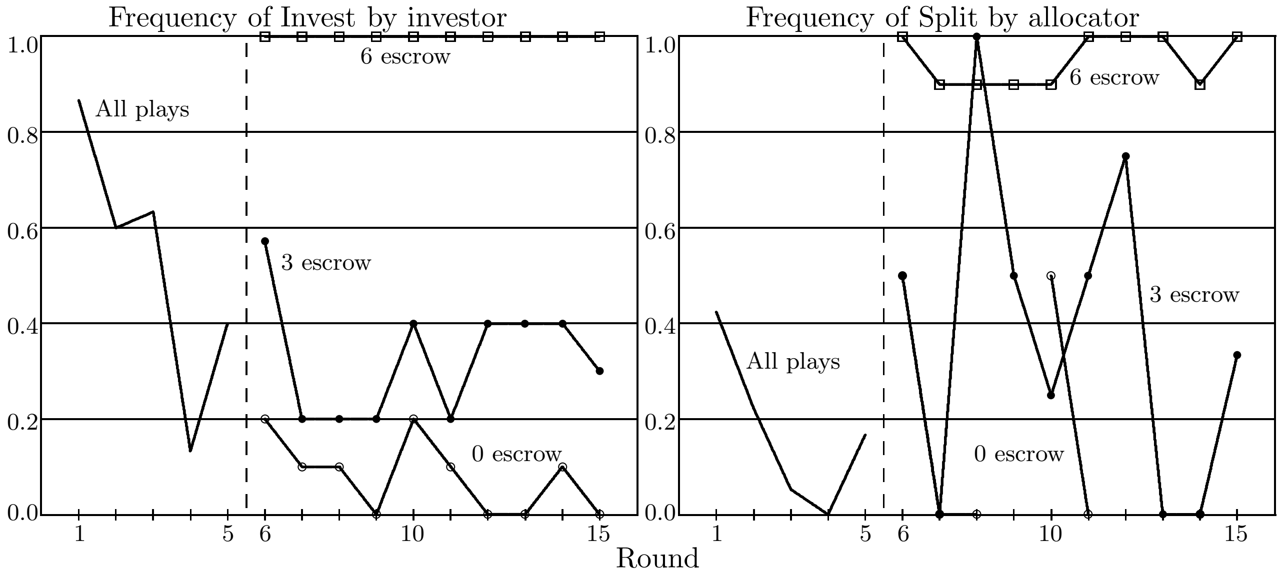
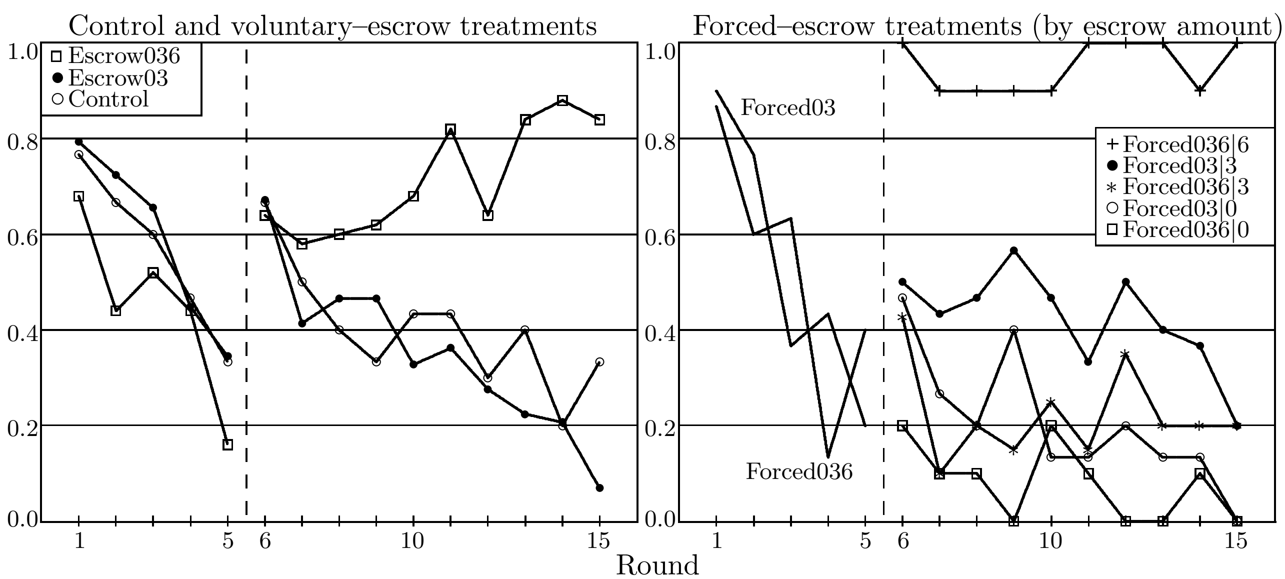


Figure 6: Round–by–round efficiencies



more likely to choose Split. In the remaining treatments, and in the Forced036 treatment when an escrow amount of 0 or 3 is imposed, there is a downward trend to efficiency, as investment, splitting, or both decrease over time.

4.3 Parametric statistics

In this section, we report the results and implications of several probit regressions. This gives us the opportunity not only to assess the significance of the suggestive results seen in the previous sections, but also to increase the power of our hypothesis tests by using the entire data set rather than limiting ourselves to data from individual treatments. Our first regression model specification (which we call Model specification I1) looks at investors' behavior, so our dependent variable is an indicator variable for Invest (1 if the investor chooses Invest, 0 otherwise). In order to capture the time dependence seen in Figures 3–5, we include right–hand–side variables for the round number and its square; to control for any restart effects between rounds 5 and 6, we also include an indicator variable that takes on the value of 1 when the round number is 6 or higher. To examine the effects of our voluntary– and forced–escrow treatments, as well as the actual escrow amounts, our main explanatory variables are indicators for the Escrow03, Escrow036, Forced03, and Forced036 game (so that the baseline is the Control treatment) and for escrow amounts of 3 and 6 (so that the baseline is an escrow amount of 0).[14] Finally, we include twelve additional explanatory variables, formed by taking the products of our four treatment indicators and two escrow–amount indicators with the round number and the square of the round number, in order to pick up any time variation in the effects of these variables.

Our second regression model specification (S1) looks at allocators' behavior conditional on investment, so our dependent variable is an indicator variable for Split. We use the same set of explanatory variables as in the investors' regression, but we restrict our data to the subset that follows an Invest choice by the investor.

Our third and fourth regression model specifications (I2 and S2) are similar to the first and second, but add four additional sets of variables designed to capture the interaction between the non–Control treatments and an escrow amount of 3 (so that we can assess our "crowding out" and "signaling" hypotheses). The first set consists of the product of our Escrow03 and "escrow amount of 3" indicator variables, the product of this variable and the round number, and the product of this variable and the square of the round number. The other three sets are similar, but use the Forced03, Escrow036, and Forced036 indicator variables (respectively) instead of Escrow03. To avoid perfect collinearity, we remove the "escrow amount of 3" indicator variable and its product with the round number and its square for these models. All four of our regressions were performed using Stata (version 8), and incorporate individual–subject random effects.

The results of these regressions are shown in Table 4 and continued in Table 5. These tables show the coefficient and standard error for each variable in each of our regressions. Also shown, at the bottom of each column in Table 5, is the absolute value of the log–likelihood for that regression. Before using the regression

[14] Recall that in Rounds 1–5 of all sessions, there is no escrow. In order for our Escrow03, Escrow036, Forced03, and Forced036 variables to actually pick up the effect of the difference in rules, these refer to the games rather than the treatments. For example, our Escrow03 variable was actually the product of indicator variables for the Escrow03 treatment and for a round number greater than 5.

Table 4: Coefficients from probit regressions with random effects (standard errors in parentheses)

Model specification	I1	S1	I2	S2
Dependent variable	Invest	Split (given Invest)	Invest	Split (given Invest)
	($N = 2160$)	($N = 1095$)	($N = 2160$)	($N = 1095$)
constant	1.503***	0.438**	1.508***	0.439***
	(0.159)	(0.195)	(0.160)	(0.196)
round number	−0.543***	−0.420***	−0.544***	−0.421***
	(0.055)	(0.083)	(0.055)	(0.084)
round^2	0.021***	0.019***	0.021***	0.019***
	(0.003)	(0.005)	(0.003)	(0.005)
round ≥ 6	1.324***	1.253***	1.325***	1.254***
	(0.217)	(0.343)	(0.217)	(0.344)
Escrow03	–1.070	–3.137	–0.653	–7.489
	(1.895)	(4.822)	(3.104)	(—)
Escrow03*round	0.209	0.366	–0.062	–0.135
	(0.394)	(1.089)	(0.668)	(0.428)
Escrow03*round^2	–0.017	–0.017	0.001	0.015
	(0.019)	(0.058)	(0.033)	(—)
Escrow036	–0.165	3.271	0.381	5.588
	(3.136)	(5.642)	(12.873)	(8.545)
Escrow036*round	–0.312	–0.731	–0.134	–1.041
	(0.668)	(1.253)	(3.474)	(2.031)
Escrow036*round^2	0.013	0.029	–0.016	0.038
	(0.033)	(0.066)	(0.230)	(0.113)
Forced03	−2.729*	–5.223	–2.401	–8.570
	(1.647)	(4.458)	(1.997)	(5.842)
Forced03*round	0.460	0.871	0.416	1.561
	(0.339)	(1.009)	(0.413)	(1.295)
Forced03*round^2	–0.024	–0.040	–0.022	–0.073
	(0.017)	(0.054)	(0.020)	(0.069)
Forced036	–1.659	–2.526	–1.884	2.494
	(2.054)	(4.927)	(2.664)	(8.366)
Forced036*round	–0.0005	0.340	0.175	–0.643
	(0.421)	(1.106)	(0.562)	(1.971)
Forced036*round^2	0.002	–0.017	–0.010	0.028
	(0.020)	(0.059)	(0.028)	(0.110)
3 escrow amount	–0.784	1.878	—	—
	(1.747)	(4.553)		
(3 escrow)*round	0.391	–0.033	—	—
	(0.363)	(1.033)		
(3 escrow)*round^2	–0.016	–0.004	—	—
	(0.018)	(0.055)		
6 escrow amount	0.357	4.492	0.172	1.560
	(3.533)	(5.747)	(12.905)	(8.900)
(6 escrow)*round	0.678	–0.214	0.425	0.295
	(0.741)	(1.254)	(3.482)	(2.069)
(6 escrow)*round^2	–0.025	0.012	0.007	–0.008
	(0.036)	(0.065)	(0.230)	(0.114)

* (**,***): *Coefficient significantly different from zero at the 10% (5%, 1%) level.*

Table 5: Probit regression coefficients, continued

Model specification	I1	S1	I2	S2
Escrow03*(3 escrow amount)	—	—	0.286	6.568***
			(3.506)	(2.030)
Escrow03*(3 escrow amount)*round	—	—	0.444	0.410
			(0.739)	(—)
Escrow03*(3 escrow amount)*round^2	—	—	–0.026	–0.033
			(0.036)	(0.022)
Escrow036*(3 escrow amount)	—	—	–4.304	2.437
			(13.508)	(13.110)
Escrow036*(3 escrow amount)*round	—	—	0.750	–0.761
			(3.576)	(2.947)
Escrow036*(3 escrow amount)*round^2	—	—	–0.008	0.056
			(0.234)	(0.160)
Forced03*(3 escrow amount)	—	—	–1.508	5.993
			(2.646)	(6.291)
Forced03*(3 escrow amount)*round	—	—	0.495	–0.871
			(0.541)	(1.383)
Forced03*(3 escrow amount)*round^2	—	—	–0.020	0.036
			(0.026)	(0.073)
Forced036*(3 escrow amount)	—	—	–1.019	–7.416
			(3.514)	(9.296)
Forced036*(3 escrow amount)*round	—	—	0.218	1.877
			(0.723)	(2.150)
Forced036*(3 escrow amount)*round^2	—	—	–0.003	–0.095
			(0.035)	(0.118)
–ln(L)	1037.102	496.900	1028.366	491.963

* (**,***): *Coefficient significantly different from zero at the 10% (5%, 1%) level.*

results to revisit our hypotheses from Section 2.2, we briefly note some of the other results. First of all, neither pair of regression models nests or is nested by the other pair, so we cannot use straightforward likelihood–ratio tests to compare them. However, we can use the similar "minimal prior information" posterior–odds criterion developed by Klein and Brown (1984). Given two models, Model A and Model B, the likelihood that Model A is the correct model—conditional on one or the other being the correct model—is given by

$$[N^{-(k_a-k_b)/2}][(\text{Maximized Likelihood under Model A})/(\text{Maximized Likelihood under Model B})],$$

where N is the sample size and k_a and k_b are the number of free parameters in Model A and Model B, respectively. (Like a standard likelihood–ratio test, this measure rewards goodness–of–fit but punishes free parameters.) According to this criterion, for investors, Model Specification I1 is over 100 billion times as likely to be correct as Model Specification I2, while for allocators, Model Specification S1 is over 70 billion times more likely than Model Specification S2. As a result, we will generally confine our discussion of results to the first two columns (unless, of course, we need to utilize the extra variables used in the other columns).[15]

[15] We also note here that comparison of the two pairs of columns of coefficients shows that the results are reasonably robust to the specification we use, so using the I2 and S2 model specifications instead would not change our results much.

For both investors and allocators, we find evidence of the usual restart effect: the coefficient of the "round ≥ 6" indicator variable is positive and significant. We also find evidence that behavior is time–dependent: for both player types, the coefficients for the round number and its square are significant, with the former negative and the latter positive, as well as being jointly significant ($\chi^2 = 126.70$, $d.f. = 2$, $p < 0.001$ for investors, $\chi^2 = 27.23$, $d.f. = 2$, $p < 0.001$ for allocators). The magnitudes of these coefficients imply that the point estimate for the round number t that minimizes $\beta_{\text{round}} \cdot t + \beta_{\text{round}^2} \cdot t^2$ (and hence the probability of Invest or Split) is roughly 26.3 for investors and 21.9 for allocators; 95% confidence intervals for these minimizers are (22.4,30.1) and (16.9,27.0) respectively. Thus, over the entire time scale of the Control sessions (which lasted for 15 rounds), the estimated frequency of both Invest and Split are declining in the round number. This negative slope is consistent with the standard finding for trust games that investment and splitting decrease over time. Our finding that last–movers (allocators) do change behavior over time is consistent with a small number of studies that have looked for such changes (see, for example, Cooper et al. (2003) in the ultimatum game).

The probit results show differences in both Invest and Split frequencies across investment amounts and across treatments. Consistent with the subgame perfect equilibrium prediction, the variables for a 6 escrow amount (that is, the "6 escrow amount" indicator and its products with the round number and its square) are jointly significant for both investors and allocators ($\chi^2 = 224.48$, $d.f. = 2$, $p < 0.001$ for investors, $\chi^2 = 70.58$, $d.f. = 2$, $p < 0.001$ for allocators). In contrast to the equilibrium prediction, the variables for a 3 escrow amount are also jointly significant for both investors and allocators ($\chi^2 = 117.16$, $d.f. = 2$, $p < 0.001$ for investors, $\chi^2 = 18.52$, $d.f. = 2$, $p < 0.001$ for allocators). For investors, the four sets of treatment variables (Escrow03, Escrow03*round, and Escrow03*round^2, and so on) are each jointly significantly different from zero (ranging from $\chi^2 = 14.17$, $d.f. = 3$, $p \approx 0.003$ to $\chi^2 = 46.86$, $d.f. = 3$, $p < 0.001$), but this is not always true for allocators (ranging from $\chi^2 = 3.64$, $d.f. = 3$, $p \approx 0.303$ to $\chi^2 = 12.10$, $d.f. = 3$, $p \approx 0.007$). Indeed, the twelve treatment variable coefficients together are jointly significant for both investors ($\chi^2 = 77.37$, $d.f. = 12$, $p < 0.001$) but not for allocators ($\chi^2 = 16.27$, $d.f. = 12$, $p \approx 0.179$).

The results shown in the table are not sufficient to test our hypotheses, as each of our hypotheses involves a combination of several coefficients. For example, determining whether investment is higher when the escrow amount is 6 than when it is 0 requires particular attention to three coefficients: those of "6 escrow amount", "(6 escrow)*round", and "(6 escrow)*round^2". In round t, the expression

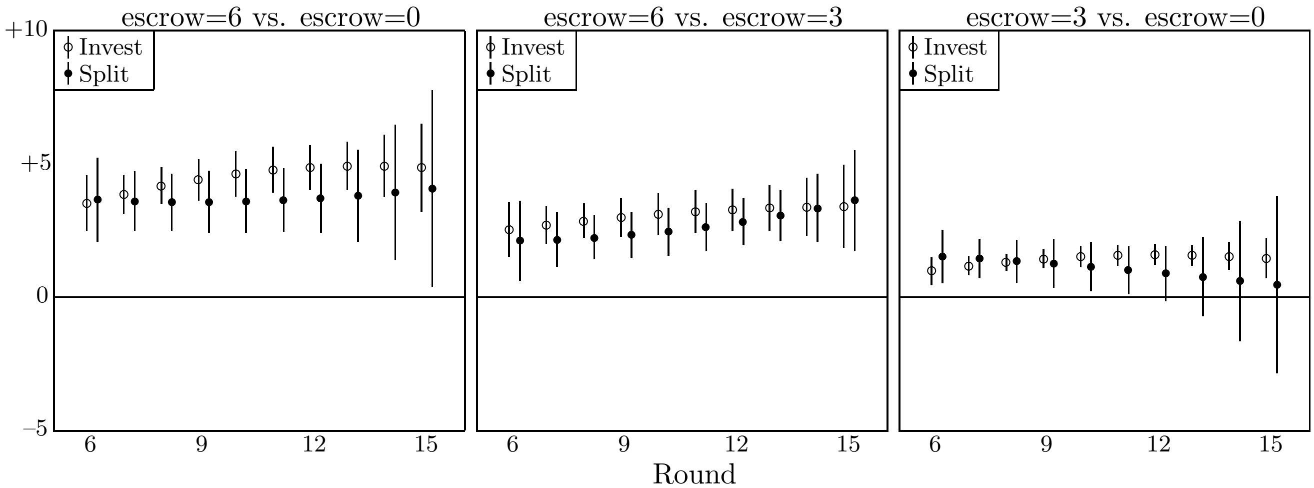
$$\beta_{\text{6 escrow amount}} + \beta_{\text{(6 escrow)*round}} \cdot t + \beta_{\text{(6 escrow)*round}^2} \cdot t^2$$

(where β_Y is the coefficient of the variable Y) is the effect of a 6 escrow amount, rather than a 0 escrow amount, on the argument of the normal c.d.f. used in the probit model. This expression will have the same sign as the marginal effect of a 6 escrow amount versus a 0 escrow amount.

Figure 7 shows, for each round and for both investors and allocators, the value of the expression above (left panel), as well as corresponding expressions for the effect of a 6 escrow amount versus a 3 escrow amount (center panel), and that of a 3 escrow amount versus a 0 escrow amount (right panel), all based on Model Specifications I1 and S1 (that is, the coefficients shown in the first two columns of Table 4). The figure shows not only the point estimates of these effects, but also estimated 95% confidence intervals.[16]

Figure 7: Estimated effects of escrow amount on Invest/Split choice based on Table 4 results

(Circles represent point estimates; line segments represent 95% confidence intervals)



As the left and center panels of the figure show, the point estimates for the effect of a 6 escrow amount versus either a 0 or a 3 escrow amount are always positive, and furthermore, the corresponding confidence intervals are entirely above zero. This means that in all rounds, both Invest and Split are significantly more likely (at the 5% level) after a 6 escrow amount than after a lower one—consistent with our Hypotheses 1 and 2 (which predicted exactly this). Moreover, these effects tend to grow slowly over time, peaking in Rounds 13 and 14 in one case (6 vs. 0 escrow amount for investors) while increasing monotonically over the length of the session in the other three cases.

The right panel shows that initially, both Invest and Split are significantly more likely after a 3 escrow amount than after a 0 escrow amount. For investors, this effect increases slightly for investors before peaking about halfway through the session. For allocators, the effect decreases monotonically, to the point where it is not significant in later rounds. To the extent that this effect is significant (for either investors or allocators), it is at odds with our Hypotheses 3 and 4, according to which frequencies of Invest and Split should be the same following either a 3 or a 0 escrow amount.

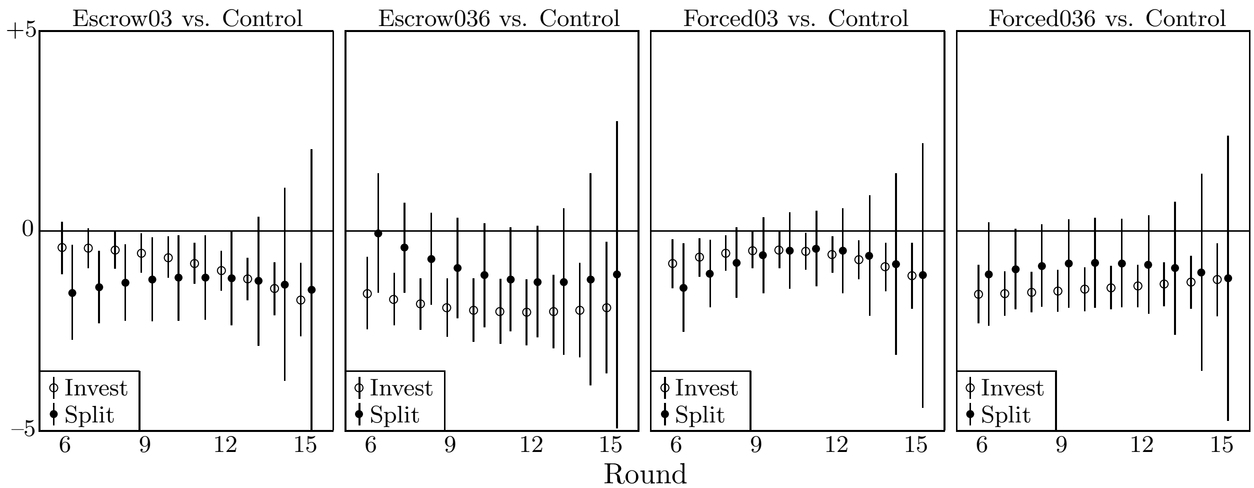
We next move to the effects of our treatments. Figure 8 shows, for each round and for both investors and allocators, the estimated effects of each of the non–Control treatments, based on Model Specifications I1 and S1 from Table 4. The four panels show, from left to right respectively, the effect of the Escrow03, Escrow036, Forced03, and Forced036 treatments versus the Control treatment, following an escrow amount of zero. As in Figure 7, both point estimates and estimated 95% confidence intervals are shown in all panels.

In each of the four panels, for both player types, and for all rounds, the point estimate for the effect of

[16]Note that these confidence intervals tend to be wider when the relevant sample sizes are smaller: for allocators—particularly when investment is rare—versus investors, for example, or for escrow amounts of 6 versus escrow amounts of 0 or 3. Also, we note here that we put results for investors and allocators in the same figure for reasons of space; we are at no time statistically testing any effect on investor behavior versus the corresponding effect on allocator behavior.

Figure 8: Estimated effects of treatment on Invest/Split choice based on Table 4 results

(Circles represent point estimates; line segments represent 95% confidence intervals)



the treatment is negative, consistent with crowding out. However, this effect is not always significant, as the 95% confidence interval contains zero in many cases.[17] It is nearly always significant at the 5% level for investors (the lone exception being round 6 in the Escrow03 treatment), while it is usually insignificant at that level for allocators, especially in later rounds.

This figure also allows us to make some pairwise comparisons between treatments. For example, we can check our assertion in Section 4.1 that behavior seemed unaffected by whether escrow amounts were chosen by allocators or imposed on them by the experimenter, by comparing the effect of the Escrow03 variables with that of the Forced03 variables, or that of the Escrow036 variables with that of the Forced036 variables. Eyeballing the figure, we see a relatively large difference between the Escrow03 and Forced03 treatments for allocators, and in later rounds for investors; between the Escrow036 and Forced036 treatments, we see only small differences for both investors and allocators. In fact, the effect of the Escrow03 treatment versus the Forced03 treatment is significantly different from zero at the 5% level for investors in Rounds 13 and 14, and for allocators in Rounds 10–12, while the effect of the Escrow036 treatment versus the Forced036 treatment is never significantly different from zero at the 5% level for either investors or allocators. However, we note the wide confidence intervals in the figure, especially in later rounds, so we stop short of claiming that changing between voluntary and forced escrow has no effect. Rather, we just remark that we fail to find systematic, significant differences between voluntary– and forced–escrow treatments.

Continuing with the discussion of crowding out, we note that the differences between Control and non–Control treatments implied by Figure 8, while suggestive, only give partial evidence in favor of crowding out, as the comparisons shown there assume implicitly that only zero escrow amounts are chosen in the non–Control treatments, when in fact an escrow amount of 3 is also possible and results in the same equilibrium prediction. Crowding out, as stated in our hypotheses, requires that frequencies of Invest and Split should be lower in the Control treatment not only following an escrow amount of 0, but overall:

[17] Since crowding out makes a directional prediction, we use one–tailed rejection regions here. Therefore, if the 95% confidence interval does not contain zero, the corresponding effect is significant at the 2.5% level.

including 3 escrow amounts also. Figure 9 reports exactly this overall effect on frequencies of Invest and Split, combining the effect of the treatment, the joint effect of the treatment and a 3 escrow amount, and the observed frequency of 3 escrow amounts in that treatment—conditional on an escrow amount of either 0 or 3. For example, the effect on the argument of the normal c.d.f. of a 3 escrow amount in the Escrow03 treatment versus a 0 escrow amount in the Control treatment is

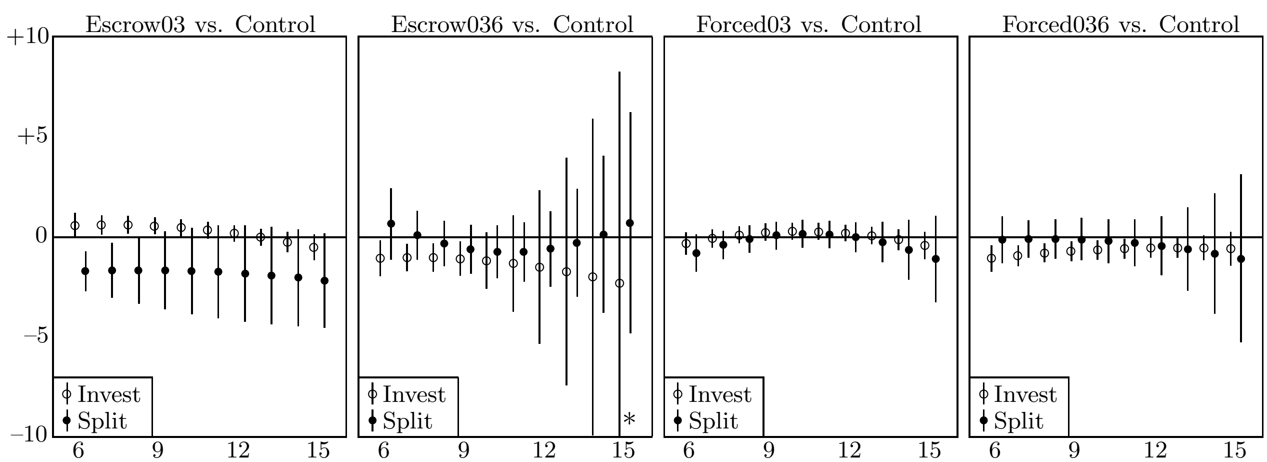
$$\beta_{\text{Escrow03}} + \beta_{\text{Escrow03*(3 escrow amount)}} + \left[\beta_{\text{Escrow03*round}} + \beta_{\text{Escrow03*(3 escrow amount)*round}}\right] \cdot t$$
$$+ \left[\beta_{\text{Escrow03*round}^2} + \beta_{\text{Escrow03*(3 escrow amount)*round}^2}\right] \cdot t^2,$$

so that the overall effect of the Escrow03 treatment versus the Control treatment is

$$\beta_{\text{Escrow03}} + p_{(3|\text{Escrow03})}\beta_{\text{Escrow03*(3 escrow amount)}}$$
$$+ \left[\beta_{\text{Escrow03*round}} + p_{(3|\text{Escrow03})}\beta_{\text{Escrow03*(3 escrow amount)*round}}\right] \cdot t$$
$$+ \left[\beta_{\text{Escrow03*round}^2} + p_{(3|\text{Escrow03})}\beta_{\text{Escrow03*(3 escrow amount)*round}^2}\right] \cdot t^2,$$

where $p_{(3|\text{Escrow03})}$ is the observed frequency of 3 escrow amounts in the Escrow03 treatment. (Since escrow amounts of 6 were not possible in this treatment, the conditional frequency of a 3 escrow amount given either a 0 or 3 amount is the same as the unconditional frequency of a 3 amount.) Note that all estimated coefficients for this figure come from Model Specifications I2 and S2 rather than I1 and S1.

Figure 9: Estimated overall effects of treatment on Invest/Split choice based on Table 4 results and observed escrow frequencies *(Circles represent point estimates; line segments represent 95% confidence intervals)*

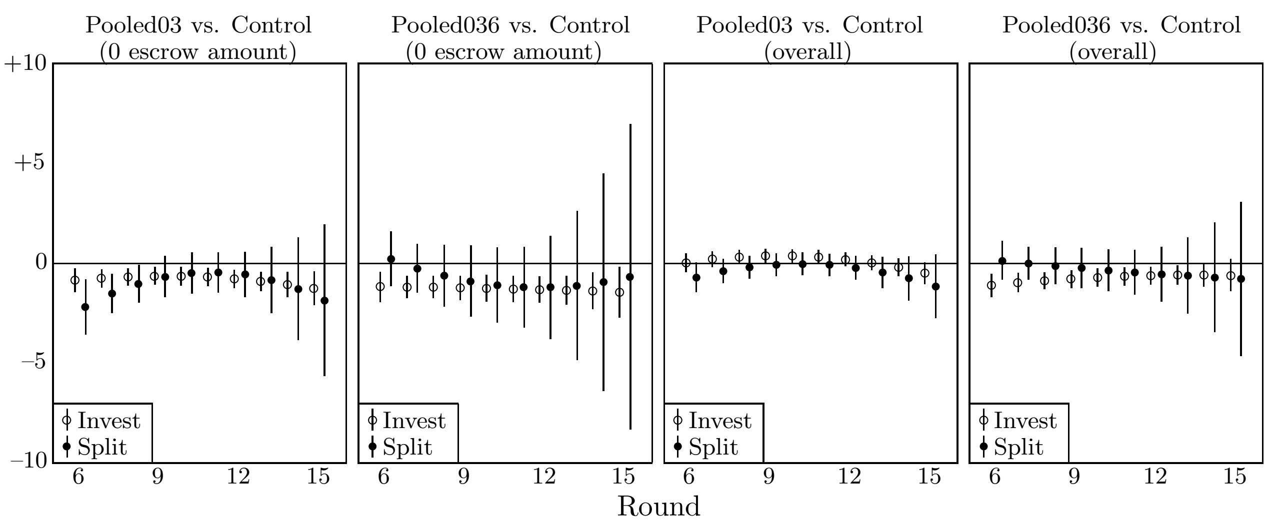


*: *Segment continues downward to –12.88.*

This figure casts some doubt on our crowding–out hypotheses. While the overall effects of the four non–Control treatments are often (though far from always) negative, these effects are typically not significant. The overall effect of either the Escrow03 or the Forced03 treatment on the frequency of Invest is significantly less than zero at the 5% level in only one case (the last round of the Escrow03 treatment), though it is sometimes significant and *positive*—the opposite of what crowding out would predict. The overall effect of either the Escrow036 or the Forced036 treatment on the frequency of Invest, on the other hand, is significant and negative (as predicted) in early rounds, but becomes insignificant by the last round—though in the case of Escrow036, the point estimate actually does continue to move away from zero in late rounds. The overall effect of any of the non–Control treatments on the frequency of Split is seldom significant and negative—the only exceptions being Rounds 6–9 and 14–15 of the Escrow03 treatment.

In sum, Figure 9 suggests that there is little systematic evidence of crowding out in our non–Control treatments; that is, the earlier evidence we saw in Figures 3–6 disappears once we control for other variables. However, the lack of significance might in at least one case (investors in Escrow036) arise from the large standard errors involved, partly due to the low number of observations of Invest choices following escrow amounts of 0 and 3. (See also the left two panels of Figure 8.) As an attempt to reduce these standard errors, we pool voluntary– and forced–escrow treatments with identical escrow amounts: we combine the Escrow03 and Forced03 treatments to form the "Pooled03" treatment, and we combine the Escrow036 and Forced036 treatments to form the "Pooled036" treatment. Figure 10 re–creates the results of Figures 8 and 9 using these pooled treatments. As the discussion surrounding Figure 8 indicated, it is not completely

Figure 10: Estimated treatment effects on Invest/Split choice based on Table 4 results and observed escrow choices *(Circles represent point estimates; line segments represent 95% confidence intervals)*



clear that this pooling is appropriate, as there were some differences in behavior between corresponding voluntary– and forced–escrow treatments, so these results should be viewed with some caution. In any event, they are broadly similar to the results of Figures 8 and 9. If only 0 escrow amounts are considered (left and left–center panels), Invest choices are significantly less likely in the non–Control treatments in all rounds, while there is little effect on Split choices except in early rounds of the "Pooled03" treatment. If we consider overall effects (right–center and right panels), we find almost no significant effects for either player type in the "Pooled03" treatment or for allocators in the "Pooled036" treatment, and while there is a negative effect for investors in the "Pooled036" treatment, the effect decreases over time and eventually becomes insignificant.

Having found support for our subgame–perfect–equilibrium hypotheses, but very little for our crowding–out hypotheses, we now turn to our signaling hypotheses. Figure 11 shows some aspects of the interaction between the effect of the Escrow03 treatment and that of the 3 escrow amount. Recall that our signaling hypotheses involve two different types of conditional frequency: an escrow amount of 3 (versus 0) conditional on the treatment being Escrow03, and the treatment being Escrow03 (versus each of the other three non–Control treatments) conditional on the escrow amount being 3. Using the variables from Model Specifications I2 and S2 in Table 4, the effect of an escrow amount of 3 rather than 0, conditional on the treatment being Escrow03, is

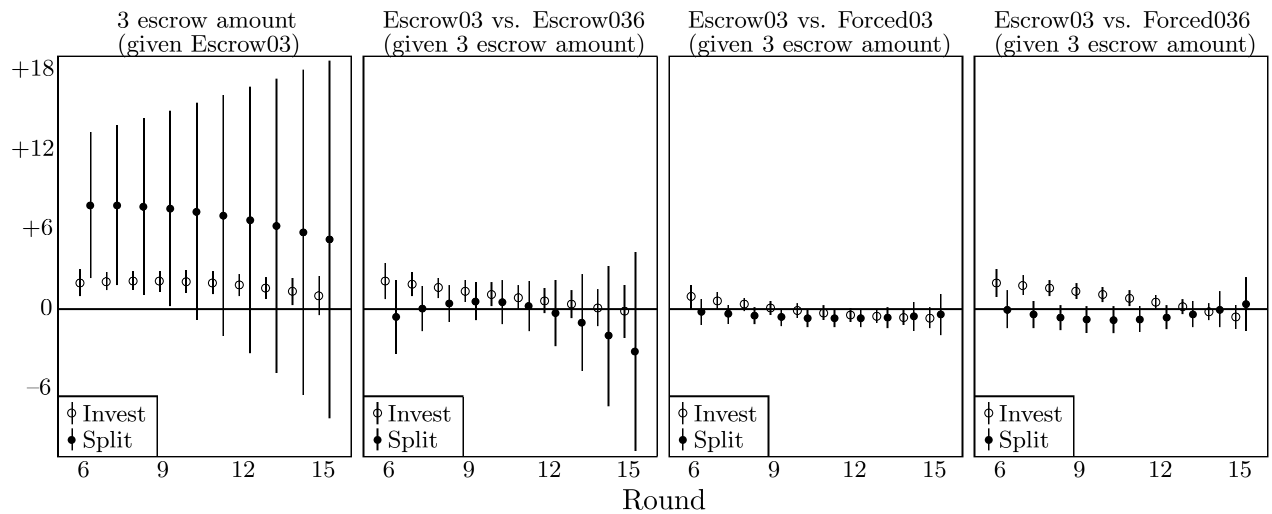
$$\beta_{\text{Escrow03*(3 escrow amount)}} + \beta_{\text{Escrow03*(3 escrow amount)*round}} \cdot t + \beta_{\text{Escrow03*(3 escrow amount)*round}^2} \cdot t^2.$$

The effect of the Escrow03 treatment versus the Escrow036 treatment, conditional on the escrow amount being 3, is

$$\begin{aligned}&\Big(\beta_{\text{Escrow03}} + \beta_{\text{Escrow03*(3 escrow amount)}} + [\beta_{\text{Escrow03*round}} + \beta_{\text{Escrow03*(3 escrow amount)*round}}]t\\&+[\beta_{\text{Escrow03*round}^2} + \beta_{\text{Escrow03*(3 escrow amount)*round}^2}]t^2\Big) -\\&\Big(\beta_{\text{Escrow036}} + \beta_{\text{Escrow036*(3 escrow amount)}} + [\beta_{\text{Escrow036*round}} + \beta_{\text{Escrow036*(3 escrow amount)*round}}]t\\&+[\beta_{\text{Escrow036*round}^2} + \beta_{\text{Escrow036*(3 escrow amount)*round}^2}]t^2\Big)\ ;\end{aligned}$$

the expressions for the effect of the Escrow03 treatment versus the Forced03 or Forced036 treatment is similar. Estimated values of these four expressions, along with estimated 95% confidence intervals are shown in Figure 11.

Figure 11: Estimates of interactions between treatments and 3 escrow amount on Invest/Split choice based on Table 4 results *(Circles represent point estimates; line segments represent 95% confidence intervals)*



The panels tell somewhat similar stories. The left panel shows that the effect of a 3 escrow amount on both investors and allocators, conditional on the Escrow03 treatment, is initially significant and positive (they are more likely to choose Invest or Split after a 3 escrow amount in this treatment than after a 0 escrow amount).[18] For allocators, the effect decreases monotonically over time so that it is statistically

[18] Again, since our hypothesis is directional, we use a one–tailed rejection region.

indistinguishable from zero by the 10th round. For investors, there is initially a slight increase, but eventually it also decreases, and is statistically insignificant by the last round. The remaining panels show that the effect on allocators of the Escrow03 treatment (versus the other non–Control treatments) conditional on a 3 escrow amount is never significantly positive (though it is significantly negative for Rounds 9–12 in the right–center panel). For investors, this effect starts out significant and positive in each of the three panels, but as it does in the left panel, decreases until eventually becoming insignificant. Thus, according to either panel in this figure, the evidence for signaling by allocators is at best ambivalent, while the evidence is reasonably strong that investors initially interpret certain allocator choices as signals, but that this belief dies out over time.

5 Conclusions

There is a growing literature studying mechanisms designed to induce cooperative behavior—and hence raise efficiency—in situations where rational, self–interested behavior is predicted to lead to inefficient outcomes. The trust game is a simple example of such a situation: the moral–hazard problem for allocators leads investors (in theory) to avoid investment, even though investment always leads to gains. We look at an escrow mechanism for this game that is predicted, in some cases, to lead to increased efficiency. Under this mechanism, the allocator puts an amount of money into an account, to be forfeited if he succumbs to the moral–hazard problem. Our experiment has five treatments, corresponding to five versions of this escrow mechanism. Our Control treatment uses only the basic trust game, with no positive escrow amount possible. In our two voluntary–escrow games, the allocator is able to choose a positive escrow amount, if he wishes, but these two games differ in which amounts are allowed. In our Escrow03 game, only low amounts (too low to lead to cooperation in any subgame perfect equilibrium) are possible, while in the Escrow036 game, one of the possible choices is high enough for equilibrium cooperation. Our remaining two games, Forced03 and Forced036, correspond to the two voluntary–escrow games, but in these, the escrow amount is exogenously imposed on the allocator, rather than chosen by him.

In our experiment, we address three sources of hypotheses. First, we consider subgame perfect equilibrium, which predicts that the players will behave cooperatively (the investor will invest and the allocator will split the proceeds) if and only if the amount put into escrow is sufficiently large. According to the subgame perfect equilibrium prediction, our mechanism may lead to increased cooperation and efficiency, but at worst will have no effect. Second, we consider "crowding out", a behavioral theory according to which externally–imposed mechanisms reduce or eliminate individuals' intrinsic tendencies to be cooperative. According to crowding out, our mechanism could actually reduce cooperative behavior and efficiency if the escrow amounts allowed are too small. Finally, we consider "signaling", a behavioral theory according to which allocators use a choice of the escrow amount to signal their intention to cooperate, so that investors, understanding this, cooperate as well—even in the Escrow03 game, where this maximum escrow amount is not high enough to make cooperation payoff–maximizing for the allocator (and thus the investor). According to signaling, our mechanism may not increase cooperation and efficiency overall, but would result in a dependence of cooperation (and efficiency) on the escrow amount when this amount is chosen voluntarily, but not when it is imposed.

We find three main sets of results. First, when the subgame perfect equilibrium predicts differences across treatments, subject behavior is consistent with these predictions. We do indeed find more investment by investors and more returns by allocators when the high amount is put into escrow than following other escrow amounts. However, subgame perfect equilibrium performs poorly in other ways; for example, there is also more investment and splitting following the low escrow amount than the zero escrow amount (subgame perfect equilibrium predicts no difference).

The other theories we considered perform less well. We tested a "crowding out" hypothesis that predicts that investment and splitting should be more likely in the Control treatment—the only one with no externally–imposed mechanism in place—than in any of the others. We found only very weak evidence in favor of crowding out: in some of our games, it didn't seem to occur at all, and in others, behavior was initially consistent with crowding out, but the effect died out quickly. If we restrict ourselves to looking only at behavior following a zero escrow amount (thus using a weaker definition of crowding out that the one usually used), we did find somewhat lower levels of investment and splitting in the non–Control treatments than in the Control treatment. However, the increases in investment and splitting following the low escrow amount (compared to the zero escrow amount) roughly cancel this out, making the overall effect negligible (and insignificant) in almost all rounds.

We also tested a "signaling" hypothesis that assumes that choosing a low positive escrow amount can be construed as a signal that the allocator intends to split. If this were true, then given a low positive escrow amount, investment and splitting would be more likely if that was the largest amount possible—and chosen voluntarily—than when either of these were not true. While the low positive amount was chosen frequently (more than three–quarters of the time) in the Escrow03 treatment, there is little evidence that this was actually a signal of cooperation by allocators, as we did not find significantly more splitting here than in the other cases. We did find, however, that investors initially seem to interpret a low positive escrow amount as a signal of intention to split, but over time and with evidence to the contrary, they learn to ignore such "signals".

Our results lead us to several conclusions. First, as the above makes clear, standard game theory is quite useful for describing behavior in our experiment. Not only does play by both investors and allocators move in the direction of the subgame perfect equilibrium prediction—so that it is a good prediction of asymptotic behavior—but subgame perfect equilibrium also fares well as a point prediction by the final round of each treatment. We cannot completely ignore the possibility of non–equilibrium phenomena, as we did find at least some evidence of both crowding out and signaling, as well as the perception of signaling by others. However, these phenomena are by and large transitory; all three disappear relatively quickly.

We next make a note regarding our use of sequential–move games rather than simultaneous–move games. In simultaneous–move games, all players make decisions under strategic uncertainty, so it can be difficult to disentangle players' preferences from their beliefs about other players' behavior, based on the actions they choose. In particular, when changes in behavior over time are observed, it is often unclear whether players are adapting to their opponents' play or evolving their own preferences. In sequential–move games, on the other hand, the player moving last (the allocator, in our games) faces no strategic uncertainty. Thus, any decision observed should be the result of only preferences, and if these decisions change over time, this should be due to changing preferences. When (as in the current paper) we do not find significant changes over time in allocators' decisions, we can infer that changes in behavior over time by other subjects (such as investors in this experiment) or in subjects in similar experiments by other researchers (such as Bracht, Figuières, and Ratto (2004)) are due to learning rather than changing preferences.

Finally, we wish to encourage more work on mechanisms for increasing efficiency. Our results suggest that these mechanisms will work only to the extent that they provide unambiguous monetary incentives for cooperative behavior. An implication of this is that two mechanisms that look roughly similar may have vastly different effects, based on the predictions made by standard game theory. A mechanism that "gets the incentives right" could lead to high levels of cooperation and efficiency—even in an environment, such as ours, where individuals have previously played a game that typically leads to betrayal and frustration. On the other hand, a mechanism that gets the incentives wrong could lead to levels of cooperation and efficiency no better than, and possibly even worse than, when there is no mechanism at all. We acknowledge the possibility that there exist other mechanisms that do lead to increased cooperation and efficiency beyond what is theoretically predicted, but further research is necessary to determine whether this is true, and if so, what form they take.[19]

[19] As a next step, Bracht and Feltovich (in preparation) examine costless and costly signaling mechanisms and their effect on investment and splitting in the trust game.

References

Akerlof, G. A. (1970), "The market for lemons: quality uncertainty and the market mechanism," *Quarterly Journal of Economics* 84, pp. 488–500.

Anderson, L., J. Mellor, and J. Milyo (2004), "Do liberals play nice? The effects of party and political ideology in public goods and trust games," Department of Economics working paper #7, College of William and Mary.

Andreoni, J. (1988), "Why free ride? Strategies and learning in public goods experiments," *Journal of Public Economics* 37, pp. 291–304.

Andreoni, J. and H. Varian (1999), "Preplay contracting in the prisoners' dilemma," *Proceedings of the National Academy of Sciences* 96, pp. 10933–10938.

Barr, A. (2003), "Trust and expected trustworthiness: experimental evidence from Zimbabwean villages," *The Economic Journal* 113, pp. 614–630.

Berg, J., J. Dickhaut, and K. McCabe (1995), "Trust, reciprocity, and social history," *Games and Economic Behavior* 10, pp. 122–142.

Blount, S. (1995), "When social outcomes aren't fair: the effect of causal attributions on preferences," *Organizational Behavior and Human Decision Processes* 63, pp. 131–144.

Bohnet, I. and R. Zeckhauser (2004), "Trust, risk and betrayal," *Journal of Economic Behavior and Organization* 55, pp. 467–484.

Bracht, J. and N. Feltovich (in preparation), "Observation, cheap talk, and cooperation in the trust game: an experimental study."

Bracht, J., C. Figuières, and M. Ratto (2004), "Relative performance of two simple incentive mechanisms in a public good experiment," CMPO working paper 04/102, University of Bristol.

Camerer, C. and E. Fehr (2003), "Measuring social norms and preferences using experimental games: a guide for social scientists," in J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr and H. Gintis, eds., *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small–Scale Societies*, Oxford, Oxford University Press, pp. 55–95.

Cooper, D., N. Feltovich, A. Roth, and R. Zwick (2003), "Relative versus absolute speed of adjustment in strategic environments: responder behavior in ultimatum games," *Experimental Economics* 6, pp. 181–207.

Cox, J.C. and C.A. Deck (2002), "The impact of trembling on behavior in the trust game," working paper, University of Arkansas.

Croson, R. and N. Buchan (1999), "Gender and culture: international experimental evidence from trust games," *American Economic Review* 89, pp. 386–391.

Croson, R., E. Fatas, and T. Neugebauer (2005), "Reciprocity, matching and conditional cooperation in two public goods games," *Economics Letters* 87, pp. 95–101.

Deck, C. (2001), "A test of game–theoretic and behavioral models of play in exchange and insurance environments," *American Economic Review* 91, pp. 1546–1555.

Eckel, C.C. and R.K. Wilson (2002), "Conditional trust: sex, race and facial expressions in a trust game," working paper, Rice University.

Eckel, C.C. and R.K. Wilson (2004a), "Is trust a risky decision?" *Journal of Economic Behavior and Organization* 55, pp. 447–465.

Eckel, C.C. and R.K. Wilson (2004b), "Detecting trustworthiness: does beauty confound intuition?" Working paper, Rice University.

Engle–Warnick, J and R.L. Slonim (2004), "The evolution of strategies in a repeated trust game," *Journal of Economic Behavior and Organization* 55, pp. 553–573.

Falkinger, J. (1996), "Efficient private provision of public goods by rewarding deviations from average," *Journal of Public Economics* 62, pp. 413–422.

Falkinger, J. (2004), "Noncooperative support of public norm enforcement in large societies," CESIFO working paper 1386.

Falkinger, J., E. Fehr, S. Gächter, and R. Winter-Ebmer (2000), "A simple mechanism for the efficient provision of public goods: experimental evidence," *American Economic Review* 90, pp. 247–264.

Fehr, E. and A. Falk (1999), "Wage rigidity in a competitive incomplete contract market," *Journal of Political Economy* 107, pp. 106–134.

Fehr, E., G. Kirchsteiger, and A. Riedl (1993), "Does fairness prevent market clearing? An experimental investigation," *Quarterly Journal of Economics* 108, pp. 437–459.

Flood (1952), "Some experimental games," Research Memorandum RM–789, RAND Corporation.

Fischbacher, U. (1999), *Z–Tree: Toolbox for Readymade Economic Experiments*, IEW Working Paper 21, University of Zurich.

Glaeser, E.L., D.I. Laibson, J.A. Scheinkman, and C.L. Soutter (2000), "Measuring trust," *Quarterly Journal of Economics* 115, pp. 811–846.

Guerra, G. and D.J. Zizzo (2004), "Trust responsiveness and beliefs," *Journal of Economic Behavior and Organization* 55, pp. 25–30.

Hardin, G. (1968), "The tragedy of the commons," *Science* 162, pp. 1243–1248.

Holm, H. and P. Nystedt (2005), "Trust in surveys and games–a matter of money and location?" working paper 26, Lund University.

Houser, D., E. Xiao, K. McCabe, and V. Smith (2005), "When punishment fails: research on sanctions, intentions, and non–cooperation," working paper, George Mason University.

Johansson–Stenman, O., M. Mahmud, and P. Martinsson (2004), "Does stake size matter in trust games?" mimeo, Göteborg University.

Klein, R.W. and S.J. Brown (1984), "Model selection when there is 'minimal' prior information," *Econometrica* 52, pp. 1291–1312.

Lazzarini, S.G., G.J. Miller, and T.R. Zenger (2004), "Order with some law: complementarity versus substitution of formal and informal arrangements," *Journal of Law, Economics, & Organization* 20, pp. 261–298.

Moxnes, E. and E. van der Heijden (2003), "The effect of leadership in a public bad experiment," *Journal of Conflict Resolution* 47, 773–795.

Ostrom, E. (2000), "Collective action and the evolution of social norms," *Journal of Economic Perspectives* 14, pp. 137–158.

Scharlemann, J.P.W., C.C. Eckel, A. Kacelnik, and R.K. Wilson (2001), "The value of a smile: game theory with a human face," *Journal of Economic Psychology* 22, pp. 617–640.

Siegel, S. and N.J. Castellan, Jr. (1988), *Nonparametric Statistics for the Behavioral Sciences*, McGraw–Hill, New York.

Varian, H. (1994), "A solution to the problem of externalities when agents are well–informed," *American Economic Review* 84, pp. 1278–1293.