Efficiency in the trust game: an experimental study of precommitment

Juergen Bracht
Department of Economics
University of Aberdeen Business School
Old Aberdeen AB24 3QY, UK
juergen.bracht@abdn.ac.uk

Nick Feltovich*
Department of Economics
University of Houston
Houston, TX 77204–5019, USA
nfelt@mail.uh.edu

April 4, 2007

Abstract

We experimentally test a precommitment mechanism for the *trust game*. Before the investor's decision, the allocator places an amount into escrow, to be forfeited if he keeps the proceeds of investment for himself. We vary the available escrow amounts—in particular, whether there is a high amount that gives rise to an efficient equilibrium—and whether escrow is voluntary or imposed. We find that when chosen, the high escrow amount does lead to efficient outcomes. We also find substantial investment when the high amount is unavailable or not chosen, though well below that when it is chosen, and declining over time. We find only weak evidence for behavioral theories, such as crowding out and signaling. These results are seen when escrow choices are imposed as well as when they are voluntary.

Journal of Economic Literature classifications: C72, D82, A13.
Keywords: mechanism, trust game, incentives, signal, crowding out.

*Corresponding author. Financial support from the University of Aberdeen, the University of Houston, and the British Academy is gratefully acknowledged. We thank James Andreoni, Dieter Balkenborg, Tilman Börgers, Jim Engle-Warnick, David DeMeza, Bob Hart, Steffen Huck, Oliver Kirchkamp, Tatiana Kornienko, Nat Wilcox, Rick Wilson, participants at several conferences and seminars, an associate editor, and two referees for helpful comments and suggestions. Any remaining errors are a result of badly aligned incentives.

1 Introduction

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

The goal of this paper is to examine the effects of a mechanism designed to improve efficiency in one of these situations: the *trust game* (Berg, Dickhaut, and McCabe (1995)), a simple collective–action game played between two players. One player (the *investor*) has the choice of investing or not investing in a project, which is administered by the other player (the *allocator*). With certainty, the investment is successful, in the sense that the amount invested multiplies in value. However, the allocator controls the proceeds of investment: he may keep the total amount for himself or split it evenly with the investor.

The trust game is often used as a metaphor for more complicated social situations. For example, consider the situation of foreign direct investment into a country with weak contract enforcement (or into a corrupt country whose domestic firms are politically well–connected).[1] If financial markets in the host country are poorly developed, domestic firms may be severely credit–constrained, so that there are sizeable opportunities for productive investment into these domestic firms by foreign firms. Often, however, once an investment is made, the investor (the foreign firm) is vulnerable to opportunistic behavior by the allocator (the domestic firm or the host country's government), such as asset stripping or even expropriation.

The prediction of game theory for this game is dismal indeed: the unique subgame perfect equilibrium has the investor refusing to invest, foreseeing that the allocator would keep the proceeds of investment.[2] This equilibrium is inefficient; total payoffs are higher if the investor invests, and it is possible for the allocator to split the proceeds so that both players are strictly better off than in equilibrium. However, there is no way in this game for the allocator to credibly commit to share rather than keep the proceeds.

Our mechanism is relatively simple. We add a pre–play stage, in which the allocator has the opportunity to place some amount of money into an *escrow* account. If the investor does not invest, or if the investor invests and the allocator splits the proceeds, then the entire escrow amount is returned to the allocator. However, if the allocator keeps the proceeds for himself, he forfeits the escrow amount (it is lost, not transferred to the investor). Thus, an escrow amount of a is equivalent to an enforced promise by the allocator to be penalized a contingent on his acting opportunistically, with no effect on payoffs otherwise. If a large enough amount is placed into escrow, he will subsequently have an incentive to share the proceeds of investment, as the gain from keeping would be outweighed by the loss of the escrow amount. In this case, the mechanism will achieve an efficient (and equitable) outcome.

In order to examine the effects of this mechanism, we run an experiment that looks at two versions of

[1] See, for example, Klein, Crawford, and Alchian (1978).

[2] We use terms 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 preferences concern only players' own monetary payoffs. We acknowledge that this is an abuse of terminology, as game theory itself makes no assumptions about what form preferences take; different preferences may lead to different theoretical predictions.

this escrow game, differing sharply in subgame perfect equilibrium predictions. In one version, the allocator is able to choose a large enough escrow amount for this efficient outcome to occur. In a second, escrow is possible, but there is no amount large enough to represent 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 escrow decisions are not made by the allocator, but rather imposed by the experimenter.

The predictions of standard game theory for our experiment 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 other escrow amounts were available.

On the other hand, many experimental researchers have found that behavioral theories (other–regarding preferences, imperfect rationality, or a combination of the two) can characterize aspects of decisions that standard game theory cannot. So, we also examine two behavioral sources of hypotheses. According to "crowding out" (Ostrom (2000)), individuals have an intrinsic tendency toward cooperative behavior, which is damaged by mechanisms providing financial incentives for such behavior. As a result, a mechanism that provides weak financial incentives (too small to change monetary best responses) would lead to less cooperation than if there had been no mechanism at all—in contrast to the equilibrium prediction of no effect. We also considered a "signaling" theory, according to which a choice by the allocator of the largest possible escrow amount is a signal that the allocator intends to split the proceeds of investment—even if this amount is too small to change the allocator's monetary best response after investment from keeping to splitting. Consequently, behavior following a given escrow amount should 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 into escrow—whether it was chosen or imposed—high efficiency results, as investors generally invest, correctly anticipating that allocators will split the proceeds. 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, and the frequency of their doing so declines over time. 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.

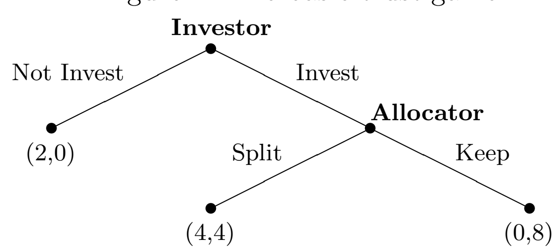
In contrast, the results show less support for our behavioral theories; typically, they describe some aspects of early–round decisions, which go away as the experiment progresses. By and large, we find only limited evidence of crowding out. If we consider a weakened version of the crowding–out hypothesis, 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 negligible (and insignificant) most of the time, particularly in later rounds. Even when we limit our focus to those subjects who had acted the most cooperatively in early rounds (when escrow was not available), we still find only qualified evidence for crowding out. 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 do seem to eventually figure this out, so that the effect eventually disappears.

2 Theoretical background

The games we consider are variants of the basic trust game 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. The allocator then decides whether to split these returns evenly with the investor or keep it all for himself; other divisions are not possible.

As mentioned in the introduction, the trust game can be considered to be a model of a number of social situations. In the game we use, an investor who trusts the allocator will choose Invest, and an allocator who is trustworthy will choose Split.[3] 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.[4]

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]In the simple version of the trust game we used, we believe that the relationships between notions such as trust and trustworthiness, and actions in the game, are uncontroversial. Other researchers have considered versions of the trust game where the investor has a continuum of possible investment levels, and the allocator a continuum of possible amounts to return to the investor. In these more complicated games, players may have varying views as to how much investment is necessary to show trust, and how much (or what portion) returned makes the allocator trustworthy.

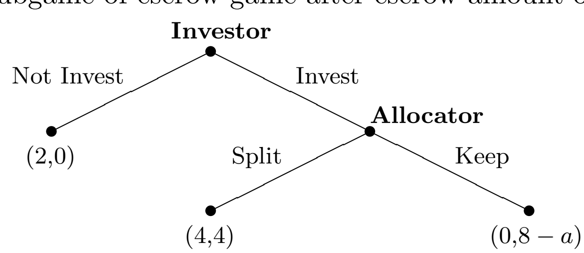
[4]This game and the games described in the next section also have non–subgame–perfect Nash equilibria. For a given game, these equilibria are all equivalent to the subgame perfect equilibrium along the path of play. For example, in all Nash equilibria of the basic trust game, the investor does not invest, and the allocator keeps with high enough probability that it would not have paid for the investor to invest.

results of these experiments have been as follows: investors invest with nonnegligible 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, and Holm and Danielson (2005) in Tanzania, in addition to the many studies in developed countries); using abstract language or various types of context; with binary 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., Barr (2003), used a stake size of half a day's wages, and Johansson–Stenman, Mahmud, and Martinsson (2004) used a stake size of two weeks' income) and the case of hypothetical payments (Holm and Nystedt (2004)). Researchers have also examined relationships between 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 trust game is inefficient—and because even though observed cooperation and efficiency in trust game experiments are typically substantially higher than predicted, they are still well below 100%—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 can place a nonnegative amount $a \in A$ in "escrow" (where the nonempty set A of allowable escrow amounts is one of our treatment variables), to be forfeited if the investor invests and the allocator keeps the returns, and returned to the allocator otherwise. Figure 2 shows the subgame

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



that results *after* the allocator's decision a is made. Note that the only difference between this subgame and the basic trust game (Figure 1) is that when the investor chooses Invest and the allocator chooses Keep, the allocator's payoff in the subgame is reduced by a relative to that in the basic trust game—the result of the allocator's forgoing the amount put into escrow. (Note that the investor's payoff is unchanged; a forfeited escrow amount simply disappears, rather than being transferred to the investor.)

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 of A. If A contains only amounts less than 4, then regardless of the allocator's choice, the investor will not invest, and the allocator would keep any returns from investment; there is no prediction for the escrow choice, since any escrow choice leads to the same payoffs. If A contains any amount(s) greater than 4, the allocator will choose an amount greater than 4 (if there is more than one such amount, there is no prediction of which will be chosen), the investor will invest, and the allocator will split.

2.2 Literature and implications

Several researchers have looked experimentally at mechanisms for improving outcomes in games where the equilibrium is inefficient. One of the simplest such "mechanisms" involves simply changing the order of play. Van Huyck, Battalio, and Walters (1995) did just this, in an experiment using the peasant–dictator game, which is very similar to the trust game. (See also Walters (1993).) In the usual version, the peasant first chooses how much to invest into a productive investment, then the dictator chooses what portion of the proceeds to take—so that the peasant and dictator correspond to the investor and allocator in the trust game. They varied the order of play, so that sometimes the dictator chose first, which is strategically equivalent to giving the dictator a way of credibly committing to a "tax rate" which is observed by the peasant prior to the investment choice. As predicted, they found that average payoffs for both the peasant and the dictator were much higher when the dictator chose first than when the investor chose first.

Andreoni and Varian (1999) examine the ability of a "compensation mechanism" (analyzed by Varian (1994)) to facilitate cooperation. In their experiment, subjects first play a prisoners' dilemma, then a modified version 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. Andreoni and Varian did indeed find more cooperation under this mechanism, but the increase was much smaller than predicted: cooperation was higher than predicted without the mechanism, but lower than predicted with the mechanism.

Houser et al. (2004) designed a trust–game experiment involving a mechanism that is similar in some ways to Andreoni's. Besides choosing an investment amount, investors choose a desired amount to be returned to them by allocators, and threaten sanctions if the allocator returns less than that amount.

They found that 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 a computer program.

Houser et al.'s main results are consistent with *crowding out*, a pattern of behavior often found in collective–action problems (including the trust game). Ostrom (2000), summarizing a large body of research, found (among other regularities) that (1) in situations like our basic trust game, levels of cooperative behavior are substantially higher than predicted by game theory, but (2) when rules are added to the game in an attempt to motivate cooperative behavior, people act approximately as predicted. Together, these results imply that "externally imposed rules tend to 'crowd out' endogenous cooperative behavior" (p. 147).[5] In the next section, we will discuss the implications of crowding out in regards to our own experiment.

Andreoni (2005) considered several versions of a "satisfaction guaranteed" mechanism, tacked on to the trust game. Using our terminology, this mechanism allows the investor—after seeing the allocator's decision—to choose whether to accept the current outcome or, instead, to impose default payoffs equal to those the players would have gotten had there been no investment. Andreoni's experiment had several treatments, varying in whether the mechanism was unavailable, required, or optional, and whether the allocator was required to honor the guarantee if it was offered. As theory predicts, he found that the highest levels of investment and returns were seen following the offer of a binding guarantee—irrespective of whether the offer was mandatory or made voluntarily. Surprisingly, investment and returns were also high when a nonbinding guarantee was offered, though not as high as when the guarantee was binding.

Falkinger et al. (2000) considered a "tax–subsidy mechanism" (proposed by Falkinger (1996)) in which a third party sets a tax/subsidy rate before the players choose contributions toward a public good. After contributions are chosen, players making above–average contributions are rewarded proportionally to how far above average their contributions were, and those with below–average contributions are punished in a similar way. In Falkinger et al.'s (2000) experiment, subjects first played a game with no mechanism followed by one with the mechanism, or played first with the mechanism then without it. Contributions were substantially higher with the mechanism than without it, but again, the increase was less than the equilibrium prediction; subjects contributed much more than predicted without the mechanism, but as much as, or less than, predicted when the mechanism was present.

Bracht, Figuières, and Ratto (2004) directly compared the mechanisms studied by Andreoni and Varian (1999) and Falkinger et al. (2000). Subjects in their experiment first played a basic public–good game, then a game with one of the two mechanisms. They 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 eventually converged to it.

[5] 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.

Under the compensation mechanism, however, contributions started at the equilibrium level but decreased over time, ending well below it (though still higher than without the mechanism).

The results of these experiments are largely consistent with each other, and carry two implications for us. First, the mechanisms' performance seems relatively robust to small changes in experimental parameters and procedures. Second, there is substantial crowding out: cooperation is well above equilibrium levels when no mechanism is in place, but under any of these mechanisms, levels of cooperation are usually no higher, and indeed are often lower, than the equilibrium prediction. Thus, when a weak externally–imposed rule does not change the equilibrium prediction, the result is a decrease in cooperative behavior under the rule compared to when no rule was present. The only exception is in Andreoni's (2005) experiment, where both investment and returns were higher after the allocator offered a nonbinding guarantee than in the basic trust game.[6] Rather, this result is consistent with a signaling hypothesis: in a treatment where no binding guarantees are possible, the choice of a nonbinding guarantee rather than no guarantee at all is read (correctly, on average) by investors to be a signal of future cooperation—even though this choice does not entail a credible commitment in the game–theoretic sense.

2.3 Experimental design and hypotheses

We consider five treatments, differing in A and in how the escrow decision is made. In our control treatment, no escrow is possible ($A = \{0\}$). In our "Escrow03" and "Escrow036" treatments, $A = \{0, 3\}$ and $\{0, 3, 6\}$, respectively. In addition, we have 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" and "Forced036" treatments, we again have $A = \{0, 3\}$ and $\{0, 3, 6\}$ respectively. (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, defined as the sum of investor and allocator payoffs, normalized so that 0 and 1 represent the minimum and maximum efficiencies. The subgame perfect equilibria of these games imply 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, leading to 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.*

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

[6] However, the very low levels of investment and returns when the allocator chose not to offer a guarantee in Andreoni's experiment are consistent with crowding out.

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

While these game–theoretic predictions are clear, there is some reason to think the actual impact of our mechanism might be different. As mentioned in the previous section, "crowding out" is often seen in games like ours. While the notion of crowding out is sufficiently broad that multiple interpretations may sometimes be possible, we consider the following interpretation. Our Control, Escrow03 (and Forced03), and Escrow036 treatments correspond to the cases of (1) no externally–imposed rules, (2) weak externally–imposed rules, and (3) strong externally–imposed rules (respectively). In our Control treatment—where no external rules are imposed—levels of investment and splitting ought to be substantially higher than game theory predicts. In the other treatments, levels of investment and splitting ought to be similar to the game–theoretic prediction, but the prediction itself will vary. 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, the rules in place are not strong enough to make cooperative behavior rational, so levels of investment and splitting should be as game theory predicts. Noting that the equilibrium prediction for the Escrow03 and Forced03 treatments is the same as that for the Control treatment, and that actual levels for the latter should be higher than the equilibrium prediction, the implication is that levels of investment and splitting should actually be 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. In the Control and forced–escrow treatments, there is no opportunity for signaling, and in the Escrow036 treatment, the implication of signaling is no different from what equilibrium predicts. In the Escrow03 treatment, however, such signaling would imply that an escrow choice of 3 leads to more cooperative behavior: investors will anticipate that allocators intend to choose Split, so they will choose Invest. Thus, 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.*

3 Experimental procedures

Subjects started all sessions by playing five rounds of the basic trust game; this was intended to familiarize them with the strategic situation and the computer interface. After these rounds were completed, subjects played ten rounds of a game that depended on the treatment. In the Control treatment, these rounds were also of the basic trust game; in the remaining treatments, they were of the corresponding game (for example, the Escrow03 game in our Escrow03 treatment).[7] All subjects in a session played the same game.

[7]Note here the distinction we draw between game and treatment in our nomenclature. Each treatment begins with five rounds of the basic trust game, then continues with ten 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).

Most sessions involved exactly 20 subjects (one Control session had only 18 subjects and one Escrow036 session had only 10 subjects). 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 a 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 written instructions for the first five rounds.[8] At this point, the subjects were not told how (or if) the game would differ in the next ten rounds, though the instructions did state 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 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 (2007)). Subjects were asked not to communicate with other subjects, other than 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 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.[9]

In a round of the basic trust game, investors were prompted to choose Invest or Not Invest. After these choices were entered, each allocator would see his counterpart's decision; if it was Invest, the allocator would then be prompted to choose Split or Keep. In a round of either voluntary–escrow game, the sequence of play was similar, except for the escrow decision. In these games, a round would begin with allocators' being prompted to choose one of the allowable escrow amounts. The investor would see her counterpart's decision before making her investment decision, after which allocators were prompted to choose Split or Keep. In the forced–escrow games, the sequence of play was identical, except that allocators did not choose the escrow amounts, but rather were informed of them at the same time investors were. The computer program chose each possible escrow amount with roughly equal frequency.[10] After the allocators had entered their decisions, all subjects received feedback from the just–completed round: the escrow amount (if applicable), the investor's choice, the allocator's choice (if the investor chose Invest), and the subject's

[8]Sample instructions can be found at <http://www.uh.edu/∼nfelt/vita.html#papers>. Other sets of instructions, as well as the raw data, are available from the corresponding author upon request.

[9]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.

[10]The assignment of escrow amounts to allocators in each round was determined in advance, rather than being drawn randomly during the round, so that results would be comparable across sessions. In the Forced03 treatment, the total numbers of 0 and 3 escrow amounts were equal (a total of 150 occurrences of each). Due to a minor programming error, the total numbers of 0, 3, and 6 escrow amounts in the Forced036 treatment were not exactly the same, but were close (a total of 100, 97, and 103 occurrences, respectively).

own payoff. Subjects were asked to observe their result, write the information down in a record sheet, and then click a button to continue to the next round. Subjects were not explicitly told their counterparts' payoffs, though they had enough information to be able to calculate them easily if they wished. Subjects were never given any information about the results of other pairs of subjects.

At the end of round 15, the experimental session ended. All subjects received a £5 show–up fee.[11] 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 fifteen sessions, three of each treatment. We begin our discussion of the results by presenting summary statistics concerning aggregate behavior in each of our treatments. Later, we will look at round–by–round behavior and examine estimation results based on parametric models.

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, conditional relative frequencies of Split choices (given Invest), and payoff efficiencies, for the first five rounds. This table also shows these quantities broken down by treatment. Since subjects

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

were playing the basic trust game in these 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 these rounds is substantially different from the subgame perfect equilibrium prediction. Investors choose Invest slightly more than half the time, and allocators choose Split about 36% of the

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

time, in these rounds. These averages hide a lot of variation across sessions (frequencies of Invest range from 32% to 72%, and those of Split from 20% to 58%) but surprisingly little variation across treatments. These aggregate results—levels of investment and returns bounded well away from both zero and one—are qualitatively similar to those of other trust game studies. Also in line with previous results, the overall proportion returned to investors in each treatment is below the level that would make Invest a 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. 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.

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

In the Control treatment, 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, in contrast, large escrow amounts are possible, leading 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; after investment, allocators split 97% of the time in the Escrow036 treatment and 95% of the time in the Forced036 treatment. When allocators can choose the escrow amount (in the Escrow036 treatment), they put 6 into escrow over three–quarters of the time. When they put less into escrow, investors seldom invest, though they do invest more often following a 3 escrow amount (23% of the time in the Escrow036 treatment and 32% of the time in the Forced036 treatment) than after a 0 escrow amount (12% and 8%, respectively). 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 otherwise. When only low escrow amounts are possible (the Escrow03 and Forced03 treatments), investment and splitting overall are comparable to the Control treatment, but this obscures differences between play after escrow amounts of 0 and after escrow amounts of 3. Both investment and splitting are substantially more frequent in the latter case than in the former, 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 data can be summarized as follows. First, the *directional* predictions of subgame perfect equilibrium describe play well. When subgame perfect equilibrium predicts a change across or within treatments, that change is seen in the data, in the predicted direction. Consistent with Hypotheses 1 and 2, investment and splitting are far more frequent in the Escrow036 and Forced036 treatments following an escrow amount of 6 than after any other escrow amount in any treatment. However, equilibrium's point predictions often perform poorly; only rarely do we see levels of investment and splitting close to zero. (In the next section, we will see that equilibrium fares better as a prediction of asymptotic behavior.)

Second, the data show evidence of crowding out. Recall that crowding out implies that cooperative behavior should be less frequent when a weak mechanism is imposed (in the Escrow03 and Forced03 treatments and in the Forced036 treatment following escrow of 0 or 3) than when there is no mechanism (in the Control treatment). In fact, the overall frequency of Invest in the weak–mechanism cases is 0.381, the frequency of Split is 0.360, and efficiency is 0.273, all lower than their counterparts in the Control treatment (0.400, 0.408, and 0.400, respectively). This difference is fairly substantial for efficiency, less so for investment and splitting.[12]

Third, levels of Invest and Split depend not only on the amount put into escrow, but also on what 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. A 3 escrow amount 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. An escrow amount of 0 is the largest possible escrow amount in the Control treatment, but a larger amount is possible in 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.

[12]One might argue that "weak mechanism" could 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.

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 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 Behavior dynamics

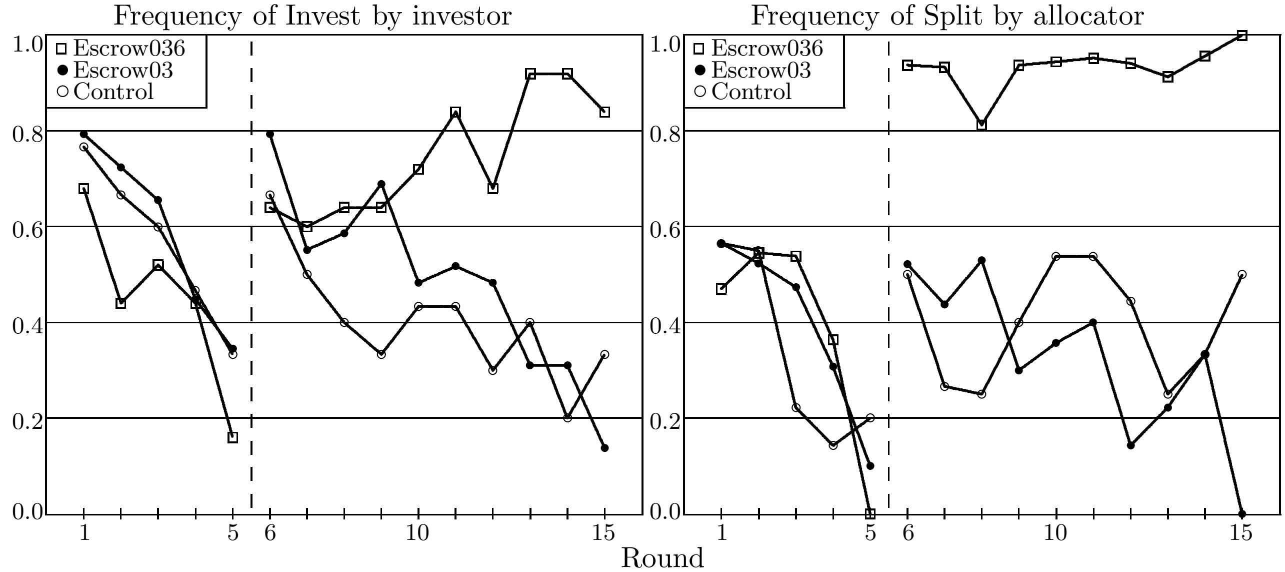
Figure 3 shows round–by–round frequencies of Invest and Split for the Control, Escrow03, and Escrow036 treatments. (These frequencies are not broken down by escrow amount, since sample sizes are small for escrow choices less than the maximum possible choice.) Qualitative dynamics in the first five rounds are similar across these treatments (and as we will see, in 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 each of the three treatments. The frequency of Split starts at about one–half and drops fairly steadily over these five rounds to below 20% in each treatment—and to 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.

In all treatments, 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 incorrectly 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 high—above 60% in both treatments—but decline over

[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.

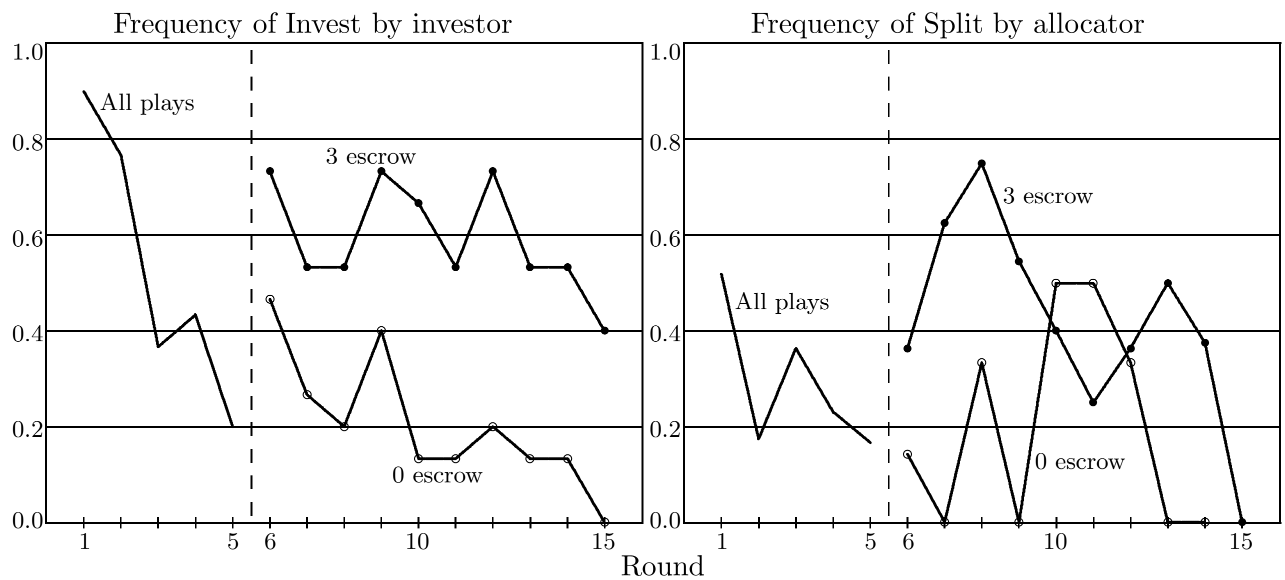
Figure 3: Relative frequencies of Invest and Split (Control and voluntary–escrow treatments)



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.

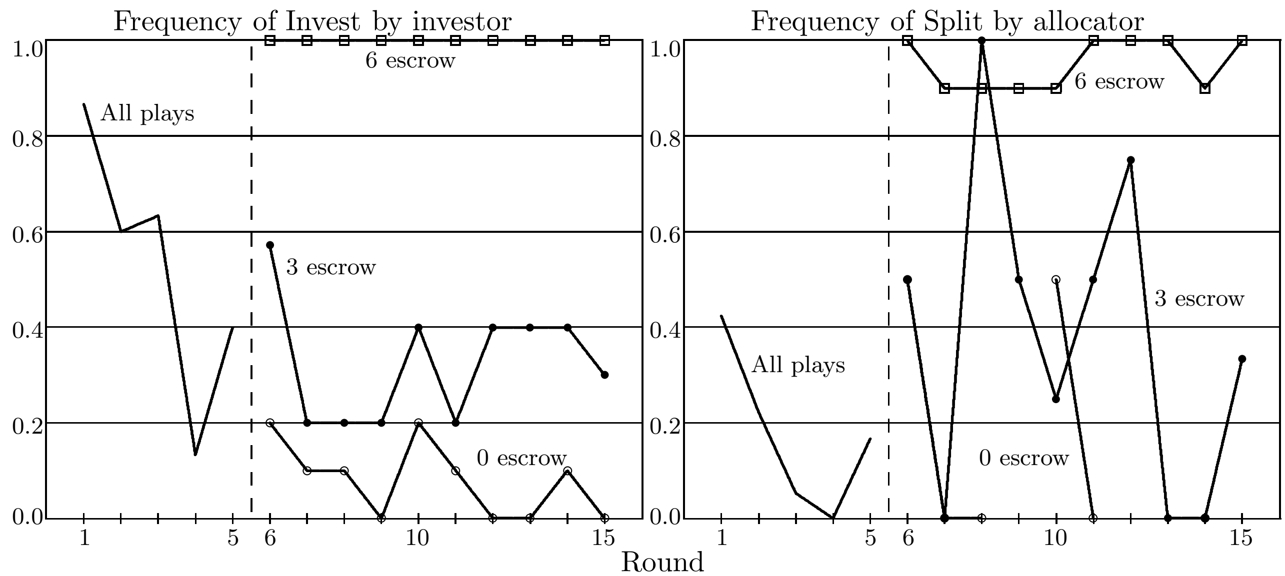
We next look at the round–by–round relative frequencies of Invest and Split in the two forced–escrow treatments. Since the theoretical predictions for these treatments depend on which escrow amount is imposed on allocators, and because there are large numbers of occurrences of each escrow amount, we 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. Behavior in the first

Figure 4: Relative frequencies of Invest and Split—Forced03 treatment (conditional on escrow amount)



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. There appears to be a restart effect for investors in the Forced03 treatment, but one is not apparent for allocators or in the Forced036 treatment. When an escrow amount of 6 is imposed in

Figure 5: Relative frequencies of Invest and Split—Forced036 treatment (conditional on escrow amount)



the Forced036 treatment, Invest is always chosen, and Split nearly always chosen, 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.

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 results seen in the previous sections, but also to increase the power of our tests by using the entire data set rather than limiting ourselves to data from individual treatments. Our first model specification (which we call Model Specification I1) looks at investors' behavior, so our dependent variable is an indicator variable for Invest. We include a right–hand–side variable for the round number, as well as an indicator variable for a round number of 6 or higher, to control for the time dependence seen in Figures 3–5, as well as any restart effects between rounds 5 and 6. Our main explanatory variables are indicators for the Escrow03, Escrow036, Forced03, and Forced036 games and for escrow amounts of 3 and 6, as well as their products with the round number. Our second regression model specification (S1) looks at allocators' behavior, so our dependent variable is an indicator variable for Split. We use the same set of explanatory variables as in Model Specification I1, but we restrict our data to the subsample that follows an Invest choice by the investor.

Our next two specifications (I2 and S2) are similar to I1 and S1, but add four pairs of variables designed to capture the interaction between the non–Control treatments and a 3 escrow amount, so that we can assess our "crowding out" and "signaling" hypotheses. The first pair consists of the product of our Escrow03 and "3 escrow amount" indicators, and the product of this variable with the round number. The other three pairs use the Forced03, Escrow036, or Forced036 variables instead of Escrow03. To avoid perfect collinearity, we remove the "3 escrow amount" variable and its product with the round number. These regressions (as well as those reported later in the paper) were performed using Stata (version 9), and incorporate individual–subject random effects.

The results of these regressions are shown in Table 4, which shows the coefficient and standard error for each variable in our four model specifications. Also shown is the absolute value of the log–likelihood, as well as a pseudo–R^2, for that model specification.[14] Before discussing the results as they pertain to our hypotheses, we briefly note some of the other results. First of all, these models are not nested, so we cannot use standard likelihood–ratio tests to compare them. However, we can use "information criteria" that, like likelihood–ratio tests, reward goodness–of–fit but punish free parameters. The Bayesian (also known as Schwarz (1978)) Information Criterion favors the two simpler models: Model Specification I1 has an BIC of 2236.580, which is lower (and thus better) than the 2264.753 of Model Specification I2, and Model Specification S1's BIC of 1119.422 is better than Model Specification S2's BIC of 1153.793. The Akaike (1974) Information Criterion, which is more forgiving of free parameters than the BIC, favors the simpler model for allocators but the less simple model for investors: I2 has an AIC of 2145.518, which is lower (and thus better) than the 2151.412 of I1, and S1's AIC of 1044.444 is better than S2's AIC of 1048.824. Because the simpler models tend to fare better under these criteria, and because comparison of the columns of Table 4 suggests that the results are reasonably robust to the specification we use, we will generally confine our discussion of results to the first two columns—unless we need the extra variables used in the other columns.

For both player types, we find evidence of declining cooperation over time and a restart effect, as the coefficient of the round number is negative and significant, while that of the "round $\geq$ 6" indicator is positive and significant. Both of these results are consistent with standard findings for trust games. The changing likelihood of Split choices over time suggests that allocators' preferences evolve (or perhaps they are learning about their own preferences, or simply are developing more understanding of the game); as there is no strategic uncertainty involved in this decision, it is hard to attribute these changes to learning about investor behavior. We note that other researchers have tested for, and found, changes in last–mover behavior over time in other games (see, for example, Cooper et al. (2003) in the ultimatum game).

The regression results suggest differences in Invest and Split frequencies across investment amounts and across treatments. The two variables for a 6 escrow amount (the "6 escrow amount" indicator and its product with the round number) are jointly significant for both investors ($\chi^2 = 223.85$, $d.f. = 2$, $p < 0.001$) and allocators ($\chi^2 = 72.89$, $d.f. = 2$, $p < 0.001$), as are those for a 3 escrow amount ($\chi^2 = 116.06$, $d.f. = 2$, $p < 0.001$ and $\chi^2 = 18.25$, $d.f. = 2$, $p \approx 0.001$, respectively). For investors, each of the four sets of treatment variables (Escrow03 and Escrow03*round, and so on) are jointly significant (with test results

[14]The pseudo–R^2 values we report were computed by rescaling the log–likelihoods into [0,1], such that a model with no right–hand–side variables other than the constant term maps to zero, and a perfect fit maps to one.

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	0.781***	–0.100	0.783***	–0.101
	(0.108)	(0.137)	(0.108)	(0.137)
round number	−0.228***	−0.132***	−0.228***	−0.133***
	(0.024)	(0.036)	(0.024)	(0.036)
round ≥ 6	1.178***	1.011***	1.179***	1.016***
	(0.214)	(0.329)	(0.214)	(0.330)
Escrow03	–0.496	−2.523**	−1.749**	–7.813
	(0.548)	(1.180)	(0.780)	(4654.371)
Escrow03*round	–0.032	0.127	0.060	0.133
	(0.053)	(0.128)	(0.082)	(719.034)
Escrow036	−2.449***	–0.833	0.498	1.874
	(0.858)	(1.372)	(2.308)	(2.275)
Escrow036*round	0.053	–0.003	–0.316	–0.281
	(0.092)	(0.148)	(0.327)	(0.299)
Forced03	−1.531***	−2.498**	−1.379**	−3.284**
	(0.499)	(1.076)	(0.571)	(1.327)
Forced03*round	0.083*	0.172	0.070	0.253*
	(0.047)	(0.116)	(0.055)	(0.137)
Forced036	−2.984***	–1.732	−1.993***	0.443
	(0.605)	(1.244)	(0.757)	(2.345)
Forced036*round	0.152***	0.085	0.080	–0.154
	(0.057)	(0.133)	(0.077)	(0.301)
3 escrow amount	0.802*	2.234**	—	—
	(0.440)	(1.034)		
(3 escrow)*round	0.056	–0.120	—	—
	(0.044)	(0.116)		
6 escrow amount	2.614***	2.685**	–0.312	0.341
	(0.927)	(1.330)	(2.283)	(2.330)
(6 escrow)*round	0.175*	0.084	0.538*	0.336
	(0.096)	(0.144)	(0.325)	(0.303)
Escrow03*(3 escrow amount)	—	—	2.724***	7.639
			(0.848)	(4654.371)
Escrow03*(3 escrow amount)*round	—	—	–0.088	–0.135
			(0.087)	(719.034)
Escrow036*(3 escrow amount)	—	—	–3.200	–3.255
			(2.505)	(3.063)
Escrow036*(3 escrow amount)*round	—	—	0.541	0.401
			(0.343)	(0.372)
Forced03*(3 escrow amount)	—	—	0.502	3.204**
			(0.667)	(1.409)
Forced03*(3 escrow amount)*round	—	—	0.080	–0.218
			(0.064)	(0.145)
Forced036*(3 escrow amount)	—	—	–0.912	–0.117
			(0.925)	(2.545)
Forced036*(3 escrow amount)*round	—	—	0.178**	0.139
			(0.091)	(0.316)
–ln(L)	1060.706	507.222	1051.759	503.412
pseudo–R^2	0.256	0.241	0.263	0.246

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

ranging from $\chi^2 = 16.56$, $d.f. = 2$, $p < 0.001$ to $\chi^2 = 52.10$, $d.f. = 2$, $p < 0.001$), but this is not always true for allocators (ranging from $\chi^2 = 3.58$, $d.f. = 2$, $p \approx 0.167$ to $\chi^2 = 14.83$, $d.f. = 2$, $p < 0.001$).

However, these joint significance tests are unsatisfactory for two reasons. First, these sets of variables contain interaction terms, and it is well known that for nonlinear regression models (such as probits), the marginal effect of the interaction between two variables is not equal to the coefficient of the interaction term. (See Ai and Norton (2003) and Norton, Wang, and Ai (2004).) Second, even if we find that a particular set of variables is jointly significant, this alone does not tell us the actual effect of these variables; for example, we have seen that the "6 escrow amount" indicator and its product with the round number are jointly significant for investors, but we do not yet know whether Invest choices are more likely after a 6 escrow amount than after a 0 escrow amount. The total effect in round t 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 is given by $\beta_{\text{6 escrow amount}} + \beta_{\text{(6 escrow)*round}} \cdot t$ (where β_Y is the coefficient of the variable Y). So, the incremental effect (the analog to a marginal effect, for a discrete variable) of a 6 escrow amount rather than a 0 escrow amount in round t has the form

$$\Phi\left(\bar{X} \cdot B + \beta_{\text{6 escrow amount}} + \beta_{\text{(6 escrow)*round}} \cdot t\right) - \Phi\left(\bar{X} \cdot B\right),$$

where $\bar{X}$ is the row vector of the other right–hand–side variables' values, and B is the column vector of their coefficients. In the following, we will use either the unconditional sample mean or, in some cases, the mean conditional on an appropriate subsample, as values for $\bar{X}$.

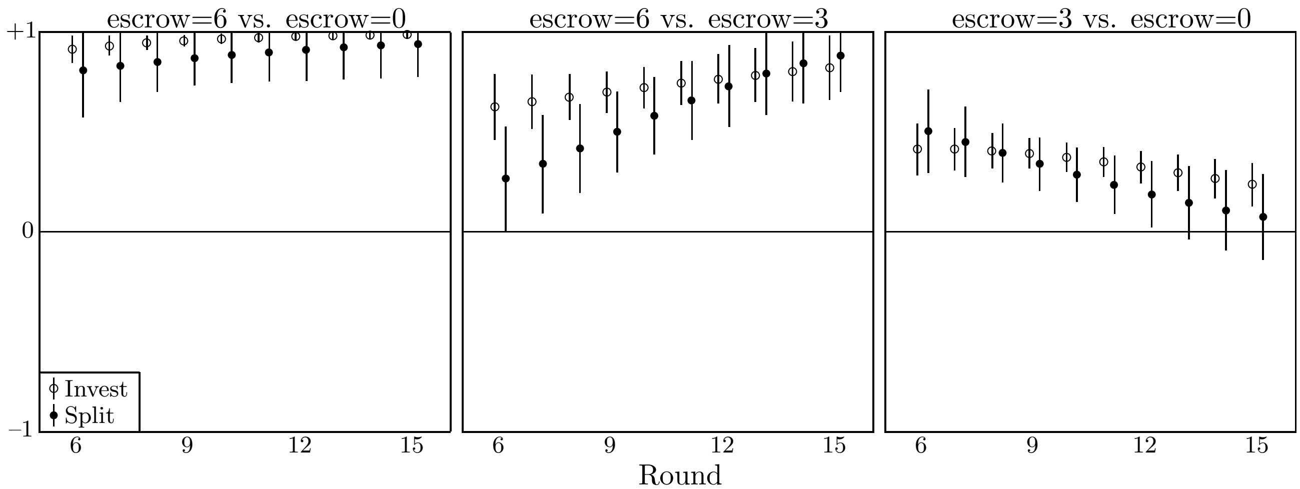
Figure 6 shows, for each round and for both player types, the estimated values of the expression above (left panel) and corresponding expressions for the effect of a 6 escrow amount versus a 3 escrow amount (center panel) and that of a 3 escrow amount versus a 0 escrow amount (right panel), based on Model Specifications I1 and S1. The figure shows point estimates and 95% confidence intervals for these effects.[15]

Not only are the point estimates for the effect of a 6 escrow amount versus either a 0 or a 3 escrow amount always positive for both player types, but the corresponding confidence intervals are as well. This means that both Invest and Split are significantly more likely after a 6 escrow amount than after a lower one—consistent with our Hypotheses 1 and 2. Moreover, the point estimates grow monotonically over the length of the session, implying that these differences become stronger as subjects gain experience. We also see that initially, both Invest and Split are significantly more likely after a 3 escrow amount than after a 0 escrow amount. However, the point estimate of this effect decreases over time for both player types; it remains significant in all rounds for investors, but by the end of the session becomes insignificant for allocators. 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.

[15] In this figure and others, we put results for investors and allocators in the same panel for reasons of space; we are at no time statistically testing any effect on investor behavior versus the corresponding effect on allocator behavior. Also, we truncate confidence intervals at +1 and –1.

Figure 6: Incremental effects of escrow amount on subject choices based on Table 4 results

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



4.4 Testing for crowding out

We next move to the effects of our treatments. Figure 7 shows, for Rounds 6–15 and for both investors and allocators, the estimated incremental effects of each of the non–Control treatments, based on Model Specifications I2 and S2.[16] The four panels show point estimates and—when available—estimated 95% confidence intervals for the effects of the Escrow03, Escrow036, Forced03, and Forced036 treatments versus the Control treatment, following an escrow amount of zero.[17]

We see some evidence of crowding out: the point estimate for the effect of each non–Control treatment is negative in all rounds for investors, and is usually negative for allocators as well. These effects are nearly always significant at the 5% level for investors (with the exception of the last two rounds of the Forced03 treatment and possibly the last two rounds of the Escrow036 treatment).[18] For allocators, these effects are significant in the Escrow03 treatment, in later rounds of the Escrow036 and Forced036 treatments, and in early rounds of the Forced03 treatment.

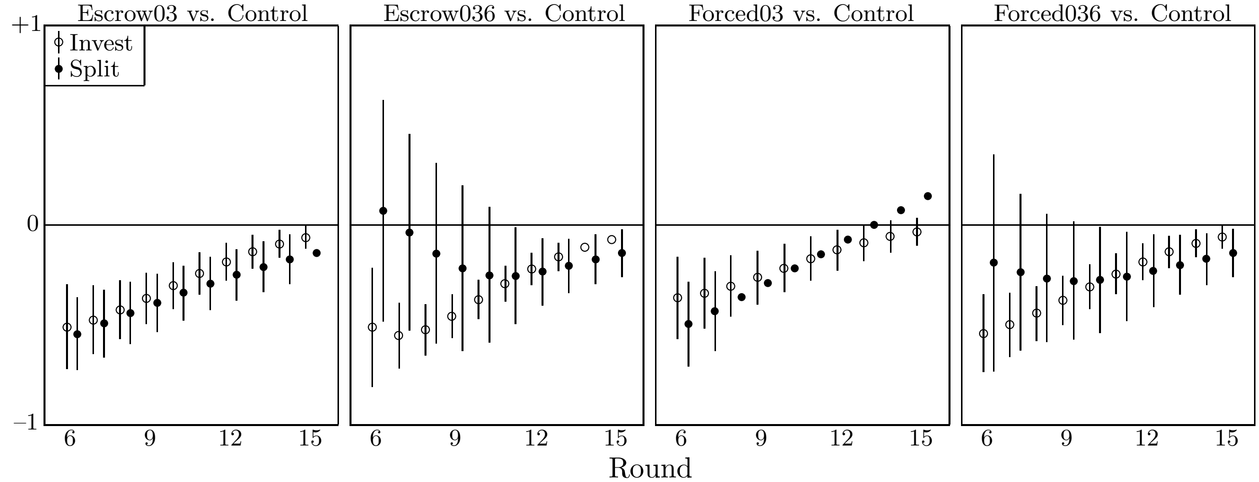
The figure also allows us to make some pairwise comparisons between treatments. For example, we can check our assertion in Section 4.1 that behavior was unaffected by whether escrow amounts were chosen by allocators or imposed on them by the experimenter, by comparing the effect of the Escrow03 variables with that of the Forced03 variables, or the Escrow036 variables with the Forced036 variables. Eyeballing the figure, we do not see large differences between either pair of treatments, for either investors or allocators. In fact, significance tests (not shown here) do find differences in investor behavior between the Escrow03

[16]Note that we are using Model Specifications I2 and S2 here rather than I1 and S1—even though we are not using the extra variables in I2 and S2—so that the results presented in Figure 7 are directly comparable to those that will be presented below in Figure 8, which does require the additional variables.

[17]In several cases (such as rounds 8–15 for allocators in Forced03 vs. Control), Stata was unable to compute standard errors for these effects; for these cases, we still report point estimates (shown as circles with no accompanying line segments), although these are for illustrative purposes only, as no tests are possible.

[18]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.

Figure 7: Incremental effects of treatment on subject choices given 0 escrow amount, based on Table 4 results *(Circles represent point estimates; segments represent 95% confidence intervals when available)*



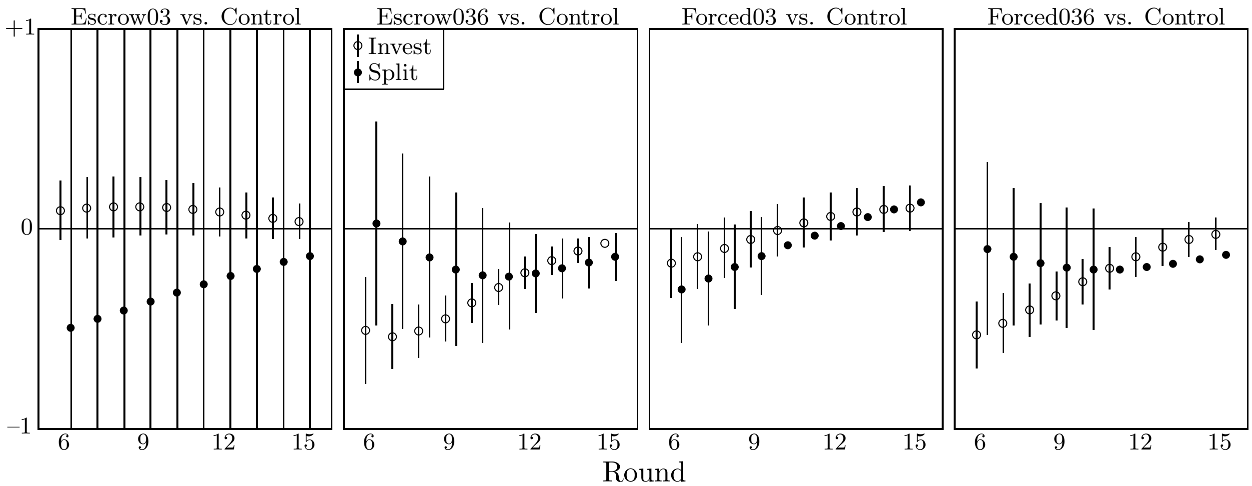
and Forced03 treatments in rounds 6–9, but not thereafter, and between the Escrow036 and Forced036 treatments in rounds 9–14, but not before or after this. For allocators, on the other hand, similar tests detect no significant differences in behavior at all, either between the Escrow03 and Forced03 treatments or between the Escrow036 and Forced036 treatments. However, the wide confidence intervals in Figure 7, especially in earlier rounds, suggest that tests between games may have low power, so we stop short of claiming that changing between voluntary and forced escrow has no effect except for investors in a few rounds. Rather, we just remark that we typically 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 7, while suggestive, only give partial evidence in favor of crowding out, because even when a negative and significant effect is shown, the associated comparison assumes 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 we defined it, requires that frequencies of Invest and Split should be lower in the Control treatment not only following a 0 escrow amount, but overall—including 3 escrow amounts also. Figure 8 reports exactly this kind of overall incremental 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, based on Model Specifications I2 and S2. For example, the overall incremental effect of the Escrow03 treatment (versus the Control treatment), as shown in Figure 8, is

$$\Phi\Big[\bar{X}\cdot B + \beta_{\text{Escrow03}} + p_{(3|\text{Escrow03})}\beta_{\text{Escrow03*(3 escrow amount)}} \\ + \Big(\beta_{\text{Escrow03*round}} + p_{(3|\text{Escrow03})}\beta_{\text{Escrow03*(3 escrow amount)*round}}\Big)\cdot t\Big] \quad - \quad \Phi\left[\bar{X}\cdot B\right].$$

This figure casts doubt on our crowding–out hypotheses. While the overall effects of the non–Control treatments are often (though far from always) negative, they are typically insignificant. In some cases—

Figure 8: Incremental overall effects of treatment on subject choices, based on Table 4 results and observed escrow frequencies *(Circles represent point estimates; segments represent 95% confidence intervals)*



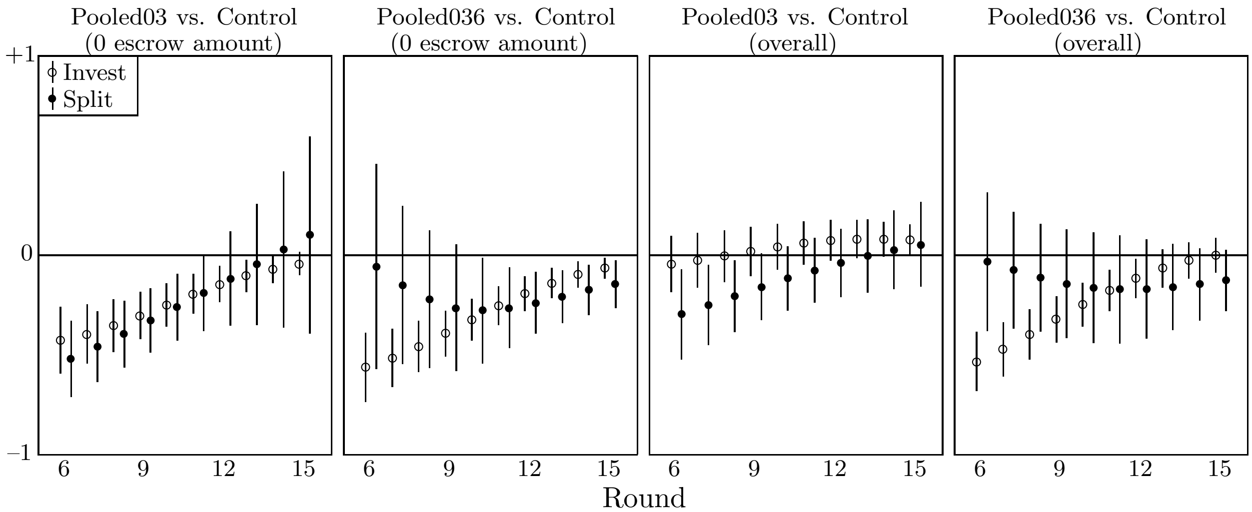
particularly allocators in the Escrow03 treatment—this lack of significance is due at least partly to high standard errors (and thus wide confidence intervals). However, except for allocators in the Escrow036 treatment, even when we do see negative point estimates, these estimates move toward zero as the round number increases, suggesting that the effects of crowding out—even when initially there—are transitory. Thus, Figure 8 suggests that the earlier evidence of crowding out we saw in Figures 3–5 (and Figure 7) tends to disappear once we account for other variables: specifically, behavior after a 3 escrow amount.

Because it is possible that the lack of significance in Figure 8 was at least partly due to large standard errors, we pool some of the data in an attempt to reduce these standard errors. Specifically, we pool treatments with identical available 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.[19] Figure 9 re–creates the results of Figures 7 and 8 using these pooled treatments. (The models that we estimated for these figures are analogous to Model Specifications I2 and S2.)

This figure suggests that, while pooling these treatments sometimes does succeed in narrowing the confidence intervals—and other times allows us to compute standard errors and therefore confidence intervals—the results are broadly similar to those of Figures 7 and 8. If only behavior following 0 escrow amounts is considered (left and left–center panels), cooperative behavior in early rounds is typically (except for allocators in the left–center panel) less likely in the non–Control treatments than in the Control treatment, but the difference shrinks over time. On the other hand, if we consider overall incremental effects (right–center and right panels), we find no significant differences in the direction implied by crowding out for either player type in late rounds of either the "Pooled03" or the "Pooled036" treatment. There are some significant differences, in the direction implied by crowding out, in early rounds, suggesting once again that to the extent that crowding out is present in our data, it dies out over time.

[19] As the discussion surrounding Figure 7 indicated, it is not completely clear that this pooling is appropriate, as there were some significant differences in behavior between corresponding voluntary– and forced–escrow treatments, so these results should be viewed with some caution.

Figure 9: Incremental effects of treatment on subject choices based on Table 4 results and observed escrow choices *(Circles represent point estimates; segments represent 95% confidence intervals)*



As another attempt to find support for crowding out, we consider the possibility that we have weakened our tests by looking at all investors and allocators, rather than just those who have some propensity to behave cooperatively. If some subjects are inherently cooperative, and others are not, then crowding out might affect only the former, in which case pooling both cooperative and uncooperative subjects might conservatively bias our tests. In order to examine this possibility, we narrow our focus to cooperative subsamples of the population. Table 5 shows the distribution of individual subjects' observed frequencies of Invest and Split choices in rounds 1–5 of all sessions of the experiment; recall that no escrow is available in any of these rounds, nor have subjects been told at this point how the game will change after round 5. We

Table 5: Distribution of individual–subject frequencies of Invest and Split in rounds 1–5

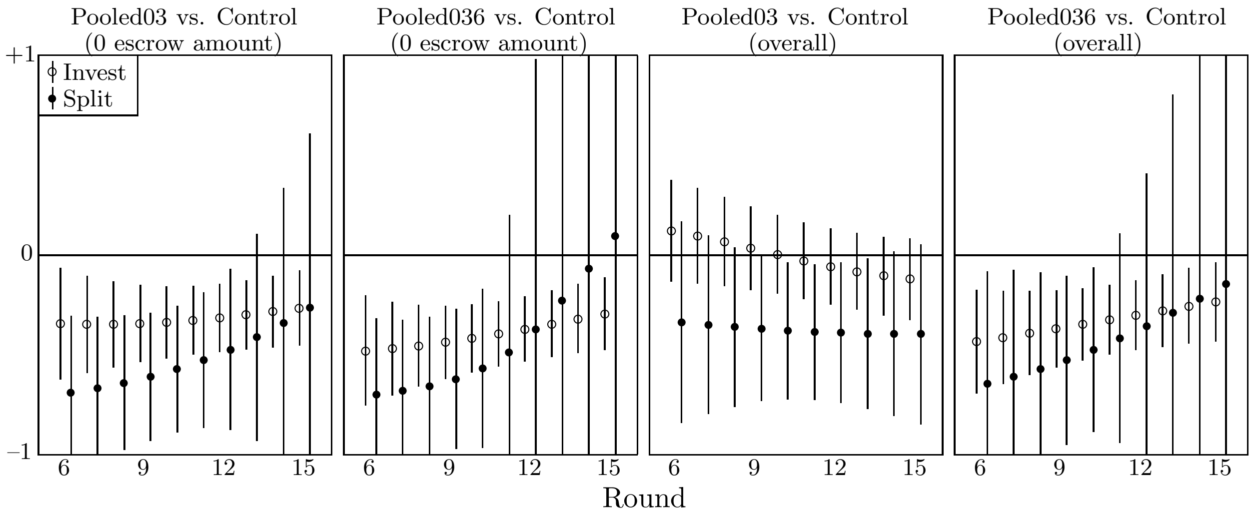
	Number of cooperative choices					
	0	1	2	3	4	5
Investors	8	9	16	25	20	11
Allocators	34	26	19	8	2	0

use these frequencies of Invest and Split as proxies for the subjects' intrinsic tendencies toward cooperative behavior. Of course, these are not perfect indicators—the number of Invest choices might reflect not only cooperative motivations, but also expectations about allocator behavior, and Split choices are only possible if the allocator's counterpart chose Invest—but they are probably the best information we have. We define "cooperative" investors as those choosing Invest at least twice (constituting 81% of investors), and "cooperative" allocators as those choosing Split at least once (constituting 62% of allocators).

We then re–estimate probits similar to I2 and S2 on these subsamples of cooperative subjects. As in Figure 9, we pool voluntary– and forced–escrow treatments in order to ameliorate the problem of high standard errors (which is even more of a problem now, due to the reduced sample). We then compute point estimates and confidence intervals for the effects of our treatments on Invest and Split choices. These

Figure 10: Incremental overall effects of treatment on choices of "cooperative" subjects based on subsample estimation results from rounds 6–15

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



results are shown in Figure 10, whose four panels correspond to their counterparts in Figure 9.

It is difficult to draw many conclusions from this figure, as confidence intervals, particularly for allocators, are quite wide. However, we do find some weak evidence in favor of crowding out, even when we focus on the overall effects of the treatments (the right–center and right panels). Both "cooperative" investors in the pooled Escrow036 and Forced036 treatments and "cooperative" allocators in the pooled Escrow03 and Forced03 treatments are less likely to behave cooperatively than in the Control treatments, and the difference is at least sometimes significant in late rounds. (On the other hand, crowding out among "cooperative" allocators in the pooled Escrow036 and Forced036 treatments goes away over time, and there appears to be no crowding out among "cooperative" investors in the pooled Escrow03 and Forced03 treatments.) If we consider choices following escrow amounts of zero only (left and left–center panels), we do see crowding out among the "cooperative" investors that persists over the length of the experimental session, though among allocators, it again dies out by the end of the session.

4.5 Testing for signaling

Having found support for our subgame–perfect–equilibrium hypotheses, but at best only qualified support for our crowding–out hypotheses, we next turn to our signaling hypotheses. Figure 11 shows some interactions 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: a 3 (versus 0) escrow amount conditional on the Escrow03 treatment, and the Escrow03 (versus another non–Control) treatment conditional on a 3 escrow amount. Based on Model Specifications I2 and S2, the incremental effect of a 3 rather than a 0 escrow amount on the probability of Invest or Split, conditional on the Escrow03 treatment, is

$$\Phi\left(\bar{X}\cdot B+\beta_{\text{Escrow03*(3 escrow amount)}}+\beta_{\text{Escrow03*(3 escrow amount)*round}}\cdot t\right)-\Phi\left(\bar{X}\cdot B\right).$$

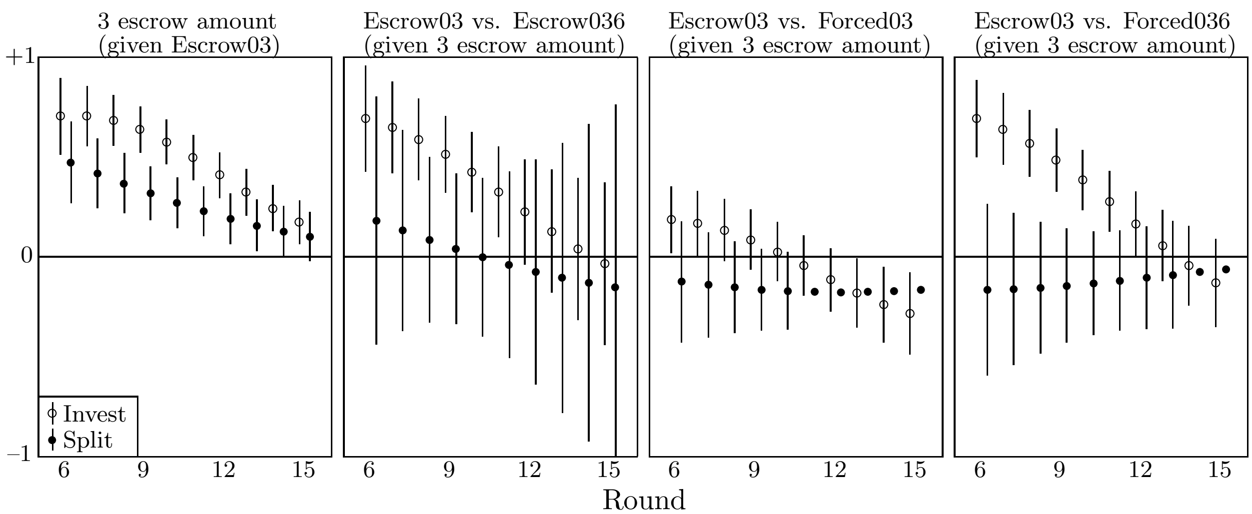
The effect of the Escrow03 versus the Escrow036 treatment, conditional on a 3 escrow amount, is

$$\begin{aligned}&\Phi\left(\bar{X}\cdot B+\beta_{\text{Escrow03}}+\beta_{\text{Escrow03*(3 escrow amount)}}+[\beta_{\text{Escrow03*round}}+\beta_{\text{Escrow03*(3 escrow amount)*round}}]t\right)\\ -\ &\Phi\left(\bar{X}\cdot B+\beta_{\text{Escrow036}}+\beta_{\text{Escrow036*(3 escrow amount)}}+[\beta_{\text{Escrow036*round}}+\beta_{\text{Escrow036*(3 escrow amount)*round}}]t\right),\end{aligned}$$

and the expressions for the Escrow03 treatment versus the Forced03 and Forced036 treatments are analogous. Point estimates of these four expressions, along with estimated 95% confidence intervals when available, are shown in Figure 11.

Figure 11: Incremental effects of selected variables on subject choices based on Table 4 results

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



The panels tell somewhat similar stories. The left panel shows that the effect of a 3 escrow amount, conditional on the Escrow03 treatment, is initially significant and positive (in this treatment, subjects are more likely to choose Invest or Split after a 3 than after a 0 escrow amount). For allocators, the effect decreases monotonically over time so that it is statistically indistinguishable from zero by the 15th round. For investors, there is initially a slight increase, but eventually it also decreases, though remaining statistically significant even in the last round. The remaining panels show that there is no persistent increase in cooperative behavior following 3 escrow amounts in the Escrow03 treatment versus the other three non–Control treatments. For allocators, there is no significant effect at all. For investors, there is a significant increase versus each of these other treatments (as our signaling hypotheses imply) in early rounds, but this effect declines over the length of the session, becoming insignificant by the end of the session. Thus, we conclude that—as we found for crowding out—the evidence for signaling is present mainly in early rounds. As the session progresses, both signaling by allocators and the belief by investors that allocators are signaling seem to die out over time.

4.6 Individual learning

Many of our results suggest that behavior changes over time. A reasonable conjecture is that over the course of the session, investors and allocators are learning about the strategic structure of the games, the behavior of the other type of player, or both. We now a look at how some aspects of past play affect the decisions subjects make in the current round.

We begin by looking at investors, and consider the natural possibility that—other things equal—they are less likely to choose Invest after getting burned by such a choice: that is, after choosing Invest in the previous round and then seeing the allocator choose Keep. To test this possibility, we define a new variable called "hurt", which is set to one if in the previous round, the investor chose Invest and the allocator chose Keep. We then estimate a new probit model (Model Specification I3) using the indicator variable Invest as the dependent variable. Our explanatory variables include "hurt" in addition to all of the variables that were used in Model Specification I1 (see Table 4).

We next consider allocators. Unlike investors, allocators sometimes have more than one decision in a round: the escrow decision and the Keep/Split decision. It is also less clear how the outcome of the previous round might affect the allocator's play in the current round. Unlike the investor's decision, the allocator's Keep/Split decision is not made under strategic uncertainty, so there is nothing about investor behavior the allocator can learn from. The escrow decision is made under strategic uncertainty, but it is not obvious what the counterpart for allocators is to "hurt" for investors. One possibility is that the allocator's tendency to choose positive escrow amounts might be lessened by such a choice being followed by an investor choice of Not Invest, as if the allocator's efforts to commit to cooperative behavior were being ignored. We thus define a variable called "ignored" which is one if in the previous round, the allocator chose a positive escrow amount and the investor chose Not Invest.

Our new regression for the Keep/Split decision (Model Specification S3) uses the Split indicator as our dependent variable. Our explanatory variables are "ignored" as well as the right–hand–side variables from Model Specification S1. For our regressions involving the escrow decision, we limit our sample to rounds 7–15 of the voluntary–escrow treatments; these are exactly the situations in which allocators had an escrow decision in both the current and the previous round. Our dependent variable is an indicator for a positive escrow amount, and "ignored" is our main explanatory variable. In our Model Specification E1, we pool the Escrow03 and Escrow036 data and use as additional explanatory variables the round number, the Escrow03 indicator, and their product. Our E2 and E3 specifications are similar, except that they look at (respectively) just the Escrow03 or just the Escrow036 treatments, so the only right–hand–side variable other than "ignored" is the round number.

Table 6 shows selected results of these regressions: coefficients and standard errors for our new variables, as well as the log–likelihood and the pseudo–R^2 for each model specification. Also shown are the p–values from significance tests for the round number and from χ^2 tests of joint significance for all right–hand–side time variables—the round number as well as its products with the various treatment and escrow–amount indicators—for each model specification. The results in this table indicate that, in decisions involving strategic uncertainty (investment and escrow, but not Split/Keep), subjects are learning from the feedback they receive, and in particular, the information present in our "hurt" and "ignored" variables. Investors

Table 6: Selected coefficients from probit regressions with random effects (standard errors in parentheses)

Model specification	I3	S3	E1	E2	E3
Dependent variable	Invest (rounds 2–15: $N = 2016$)	Split (given Invest, rounds 2–15: $N = 979$)	Positive escrow amount (rounds 7–15) Voluntary–escrow sessions: $N = 432$	Escrow03 only: $N = 232$	Escrow036 only: $N = 200$
"hurt" (Invest, Keep in prev. round)	−0.408*** (0.083)	—	—	—	—
"ignored" (positive escrow, No Invest in prev. round)	—	0.070 (0.241)	−0.438** (0.213)	−0.506* (0.280)	–0.405 (0.325)
Other right–hand–side variables	Same as Model I1	Same as Model S1	Round number, Escrow03, Escrow03*round	Round number	
Significance of round number	$p < 0.001$	$p \approx 0.081$	$p \approx 0.008$	$p \approx 0.033$	$p \approx 0.012$
Joint significance of all RHS time variables	$p < 0.001$	$p \approx 0.008$	$p \approx 0.003$	$p \approx 0.033$	$p \approx 0.012$
–ln(L)	972.385	428.872	180.997	121.014	59.221
pseudo–R^2	0.265	0.269	0.056	0.024	0.076

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

are significantly less likely to choose Invest in the next round if they chose Invest in the current round, and the allocator's response was Keep. Based on the pooled Escrow03 and Escrow036 data, we see that allocators, for their part, are significantly less likely to choose a positive escrow amount in the next round if they did so in the current round, but the investor chose Not Invest, though this association becomes less significant, or even insignificant, if we look at the Escrow03 and Escrow036 data separately. Perhaps not surprisingly, we do not see any significant relationship between the "ignored" variable and the likelihood of allocators' choosing Split; as mentioned before, the Keep/Split decision is not made under any strategic uncertainty, so it's difficult to imagine allocators' learning anything about investors that would be useful in making this decision. We also see that in all five columns, the collection of all right–hand–side time variables are significant (at least at the 5% level), as is the round number itself (at least at the 10% level), suggesting that these "hurt" and "ignored" variables do not fully capture the learning investors are doing. It could be that investors and allocators are also learning via information other than that present in these variables (possibly including results from rounds other than the most recent), or they might be learning partly through their own introspection between rounds, or both.

5 Conclusions

There is a growing literature devoted to mechanisms designed to induce cooperative behavior—and hence raise efficiency—in situations where 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 but not others, 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 and efficiency. Our Forced03 and Forced036 games correspond to the two voluntary–escrow games, but in these, the escrow amount is exogenously imposed on the allocator, rather than chosen by him.

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 an implication of the "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 after zero or low escrow amounts in any of the others. We found a small amount of evidence in favor of this hypothesis. In some of our games, we failed to find any evidence at all of crowding out, 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. Support for crowding out did improve somewhat when we concentrated on subjects who had behaved particularly cooperatively in early rounds.

We also tested a "signaling" hypothesis, according to which a maximum escrow choice can be construed as a signal that the allocator intends to split. If so, then given a low escrow amount, investment and splitting should be more likely if that was the largest amount possible—and chosen voluntarily—than when either of these is 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, standard game theory is quite useful for describing behavior in our experiment. Not only does play move in the direction of the subgame perfect equilibrium prediction—so that it is a good prediction of asymptotic behavior—but it often fares well as a point prediction by the final round of each treatment. We cannot completely ignore non–equilibrium phenomena, as we did find 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 decay quickly over time, usually to the point of disappearing completely by the end of a session. The effectiveness of standard game theory, and relative lack thereof of our other sources of hypotheses, stand in contrast to results found in experimental tests of mechanisms by other researchers (see Section 2.3); the reason for this is unclear.

We next make a note regarding our use of sequential–move games rather than simultaneous–move games. In simultaneous–move games, all decisions are made under strategic uncertainty, so it can be difficult to disentangle players' preferences from their beliefs about others' 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 faced with the Keep/Split decision, in our games) faces no strategic uncertainty. Thus, any choice observed at that stage of the game should be the result of only preferences, so that if these decisions change over time, this should be due to changing preferences—or perhaps learning about what one's preferences are. (Of course, allocators' escrow choices are made under strategic uncertainty, so that changes in escrow choices need not arise from changing preferences.) On the other hand, in cases where we do not find significant changes over time in allocators' decisions, we can take this as evidence 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.[20]

References

Ai CR, Norton EC (2003) Interaction terms in logit and probit models. Econ Letters 80:123–129

Akaike H (1974) A new look at the statistical model identification. IEEE Transactions on Automatic Control 19: 716-723

Akerlof GA (1970) The market for lemons: quality uncertainty and the market mechanism. Quarterly J Econ 84:488–500

[20] Andreoni's (2005) nonbinding "satisfaction–guaranteed" mechanism, discussed in Section 2.3, is one possibility that deserves further study. In addition, Bracht and Feltovich (2007) examine costless and costly signaling mechanisms and their effects on investment and splitting in the trust game.

Anderson L, Mellor J, Milyo J (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. J Pub Econ 37:291–304.

Andreoni J (2005) Trust, reciprocity, and contract enforcement: experiments on satisfaction guaranteed. Working paper, University of Wisconsin.

Andreoni J, Varian H (1999) Preplay contracting in the prisoners' dilemma. Proc Nat Acad Sci 96:10933–10938.

Barr A (2003) Trust and expected trustworthiness: experimental evidence from Zimbabwean villages. Econ J 113:614–630.

Berg J, Dickhaut J, McCabe K (1995) Trust, reciprocity, and social history. Games Econ Beh 10:122–142.

Blount S (1995) When social outcomes aren't fair: the effect of causal attributions on preferences. Org Beh and Human Decision Processes 63:131–144.

Bohnet I, Zeckhauser R (2004) Trust, risk and betrayal. J Econ Beh Org 55:467–484.

Bracht J, Feltovich N (2007) Observation, cheap talk, and cooperation in the trust game: an experimental study. Working paper, University of Houston.

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

Camerer C, Fehr E (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: 55–95.

Cooper D, Feltovich N, Roth A, Zwick R (2003) Relative versus absolute speed of adjustment in strategic environments: responder behavior in ultimatum games. Exp Econ 6:181–207.

Cox J, Deck CA (2002) The impact of trembling on behavior in the trust game. Working paper, University of Arkansas.

Croson R, Buchan N (1999) Gender and culture: international experimental evidence from trust games. Amer Econ Rev 89:386–391.

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

Deck C (2001) A test of game–theoretic and behavioral models of play in exchange and insurance environments. Amer Econ Rev 91:1546–1555.

Eckel CC, Wilson RK (2002) Conditional trust: sex, race and facial expressions in a trust game. Working paper, Rice University.

Eckel CC, Wilson RK (2004a) Is trust a risky decision? J Econ Beh Org 55:447–465.

Eckel CC, Wilson RK (2004b) Detecting trustworthiness: does beauty confound intuition? Working paper, Rice University.

Engle-Warnick J, Slonim RL (2004) The evolution of strategies in a repeated trust game. J Econ Beh Org 55:553–573.

Falkinger J (1996) Efficient private provision of public goods by rewarding deviations from average. J Pub Econ 62:413–422.

Falkinger J, Fehr E, Gächter S, Winter-Ebmer R (2000) A simple mechanism for the efficient provision of public goods: experimental evidence. Amer Econ Rev 90:247–264.

Fehr E, Falk A (1999) Wage rigidity in a competitive incomplete contract market. J Pol Econ 107:106–134

Fehr E, Kirchsteiger G, Riedl A (1993) Does fairness prevent market clearing? An experimental investigation. Quarterly J Econ 108:437–459

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

Fischbacher U (2007) Z–Tree: toolbox for readymade economic experiments. Forthcoming, Exp Econ

Glaeser EL, Laibson DI, Scheinkman JA, Soutter CL (2000) Measuring trust. Quarterly J Econ 115:811–846

Guerra G, Zizzo DJ (2004) Trust responsiveness and beliefs. J Econ Beh Org 55:25–30

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

Holm HJ, Danielson A (2005) Tropic trust versus Nordic trust: experimental evidence from Tanzania and Sweden. Econ J 113:505–532

Holm HJ, Nystedt P (2005) Trust in surveys and games—a matter of money and location? Working paper 26, Lund University

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

Johansson-Stenman O, Mahmud M, Martinsson P (2004) Does stake size matter in trust games? Mimeo, Göteborg University

Klein B, Crawford RG, Alchian AA (1978) Vertical integration, appropriable rents, and the competitive contracting process. J Law Econ 21:297–326

Lazzarini, SG, Miller GJ, Zenger TR (2004) Order with some law: complementarity versus substitution of formal and informal arrangements. J Law Econ Org 20:261–298

Moxnes E, van der Heijden E (2003) The effect of leadership in a public bad experiment. J Conflict Res 47:773–795

Norton EC, Wang H, Ai CR (2004) Computing interaction effects and standard errors in logit and probit models. STATA J 4:154–167

Ostrom E (2000) Collective action and the evolution of social norms. J Econ Perspectives 14:137–158

Scharlemann JPW, Eckel CC, Kacelnik A, Wilson RK (2001) The value of a smile: game theory with a human face. J Econ Psych 22:617–640

Schwarz G (1978) Estimating the dimension of a model. Annals Stat 6:461–464

Siegel S, Castellan Jr. NJ (1988) Nonparametric statistics for the behavioral sciences 2nd ed, McGraw Hill, New York

Van Huyck JB, Battalio RC, Walters MF (1995) Commitment versus discretion in the peasant–dictator game. Games Econ Beh 10:143–170

Varian H (1994) A solution to the problem of externalities when agents are well–informed. Amer Econ Rev 84:1278–1293

Walters MF (1993) Announcements and reputation in the peasant–dictator game: an experimental analysis. Unpublished masters thesis, Texas A&M University