Social Norms, Information and Trust among Strangers: Theory and Evidence*

John Duffy[a] Huan Xie[b] and Yong-Ju Lee[c]

Revised Draft: July 2011

Abstract

Can a social norm of trust and reciprocity emerge among strangers? We investigate this question by examining behavior in an experiment where subjects repeatedly play a two-player binary "trust" game. Players are randomly and anonymously paired with one another in each period. The main questions addressed are whether a social norm of trust and reciprocity emerges under the most extreme information restriction (anonymous community-wide enforcement) or whether trust and reciprocity require additional, individual-specific information about a player's past history of play and whether that information must be provided freely or at some cost. In the absence of such reputational information, we find that a social norm of trust and reciprocity is difficult to sustain. The provision of reputational information on past individual decisions significantly increases trust and reciprocity, with longer histories yielding the best outcomes. Importantly, we find that making reputational information available at a small cost may also lead to a significant improvement in trust and reciprocity, despite the fact that most subjects do not choose to purchase this information.

JEL Codes: C72, C78, C91, C92, L14.

Keywords: Social Norms, Trust Game, Random Matching, Trust and Reciprocity, Information, Reputational Mechanisms, Experimental Economics.

---

* For helpful comments and suggestions we wish to thank the Editor, an anonymous referee, Andreas Blume, Giovanna Devetag, Guillaume Frechette, Soiliou Namoro, Lise Vesterlund, Roberto Weber, as well as seminar participants at Carnegie-Mellon University, Iowa State University, Fudan University, Shanghai Jiao Tong University, the University of Copenhagen, the 2009 CEA annual meeting and the 2009 DIME Workshop on Market and Firm Dynamics.

a Department of Economics, University of Pittsburgh, Pittsburgh, PA 15260 USA, jduffy@pitt.edu

b Department of Economics, Concordia University, Montreal, QC H3G 1M8, Canada, huanxie@alcor.concordia.ca

c Samsung Research Institute of Finance, Seoul, South Korea, yj0612.lee@samsung.com

1. Introduction

Trust is a key element in sustaining specialization and trade. In many economic transactions, trust emerges among essentially anonymous agents who have little recourse to direct or immediate punishment. For instance, in electronic commerce, it is easy to create new identities, and buyers and sellers often engage in what are, essentially, one-shot transactions. In the market for credit cards, individual card-holders frequently display little loyalty to any particular bank or card issuer, freely switching balances between credit cards. Similarly, few tourists repeatedly return to the same vacation area to consume again in the same hotel or restaurant.

Given the anonymous and infrequent nature of economic transactions in these markets, an important question is how such markets can work efficiently. In particular, what is the incentive for sellers in electronic markets to deliver the goods purchased, or of the quality promised, knowing that they are unlikely to meet the same buyer again? What is the incentive for borrowers to repay credit card debts if they can switch to another lender next time? What is the incentive for hotels and restaurants in vacation areas to provide good service, knowing that the same consumers are unlikely to ever return?

One possibility is that such incentive problems can be solved by a legal process. However, in many instances, the cost of litigation would far exceed the benefit from the transaction; in such instances legal considerations can simply be ruled out. On the other hand, we do observe that in all of these markets there exist *reputation systems* that collect and disseminate information about market participants. For instance, in most electronic markets there is an online feedback system that allows buyers to rate their prior transaction experiences with sellers and this information is publicly (and typically freely) available. In the credit card market, third party credit bureaus collect information about the customers of all banks and credit card companies and provide the information to other financial institutes, typically for a small fee. Travel guides and websites (e.g. Tripadvisor) provide feedback from tourists about hotels and restaurants in vacation areas.

In this paper we examine several mechanisms by which trust and the reciprocation of trust might be sustained by a finite population of strangers who must repeatedly and anonymously interact with one another in pairwise meetings. We first examine the hypothesis that trust is attached to the society as a whole; the fear of the destruction of that trust might suffice to enforce trustworthy behavior by *all* members of the society as shown by Kandori (1992). As an alternative (but within the same economic environment), we examine the possibility that trustworthiness resides at the *individual* rather than at the *societal* level. In particular, we ask whether the provision of information on individual reputations for trustworthiness engenders greater trust than in the case where such information is absent, and if so, how long the history of information must be. Finally, and perhaps most importantly, we explore whether the free provision of reputational information is responsible for our findings or whether trust and reciprocity can be sustained by making individual reputational information costly to acquire, as occurs, for example, when an agent must purchase a credit report to learn about the past trustworthiness of another agent.

We explore these issues both theoretically and experimentally using versions of a repeated two-player sequential "trust" (or "investment") game (due to Berg et al., 1995). In our version of this game, the first mover or "investor" decides whether to invest his endowment with the second mover, the "trustee," resulting in an uncertain payoff. Alternatively, the investor can simply keep his endowment. If the investor invests (or "trusts"), the endowment is multiplied by a fixed factor that is greater than 1 and it falls to the trustee to decide whether to keep the whole amount or to return some fraction of it to the investor (i.e. to "reciprocate"), keeping the remainder for himself. Subjects are asked to play repeated versions of this game where, in each period, subjects were randomly and anonymously paired with one another. Using this random, anonymous, pairwise matching protocol, we examine several different treatments. In our baseline treatment the trust game is played for an indefinite number of periods and is parameterized in such a way that, given the number of participants and random anonymous matching, a *social norm*[1] where all investors invest (trust) and all trustees return part of the investment (reciprocate) constitutes a sequential equilibrium without *any* information provided to investors regarding the identity of their current trustee or that trustee's past history of play. In a second treatment everything is the same as in the baseline treatment except that the trust game is played for a known finite rather than an indefinite number of periods; in that case, a social norm of full trust and reciprocity is not a sequential equilibrium. In a third treatment, everything is the same as in the baseline treatment except that, prior to making a decision, the investor can observe the trustee's action choice in the prior period (Keep or Return). In a fourth treatment, everything is the same as in the third treatment except that, prior to making a decision, the investor can observe a longer history of the trustee's prior choices as well as the frequency with which the trustee chose to return in the current supergame. Finally, in a fifth treatment, everything is the same as in the fourth treatment, except that the investor must first choose whether to pay a small cost to view the trustee's history of actions for the current supergame. If the investor does not pay this information cost then, from the investors' perspective, the game is similar to our first baseline treatment where the investor has no knowledge of the prior actions of the trustee with whom he is matched. If the investor does pay for this information, then, from the investor's perspective, the game is similar to that of our fourth treatment. Importantly, in our fifth treatment, the trustee does not know whether the investor has purchased information about the trustee's past behavior.

In the first treatment, where no individual information is available, we are able to test the theoretical possibility that a social norm of trust and reciprocity can be sustained by anonymous, randomly matched agents out of the fear that deviating from such a norm would precipitate a contagious wave of distrust and

[1] We follow Kandori (1992) and Young (2008) in defining a "social norm" as rules of behavior that serve to coordinate interactions among individuals and specify sanctions for violators.

retaliatory non-reciprocation. In this baseline treatment we find that the frequency of trust and reciprocity is low, averaging around one-third. Our second treatment examines whether the indefiniteness of the horizon plays a role as theory suggests in sustaining trust and reciprocity. We find that with a finite horizon, the frequency of trust and reciprocity is about the same as with an indefinite horizon. Our third treatment asks whether "minimal" reputational information at the individual level can improve matter relative to the baseline frequency, specifically whether additional information on the prior-period behavior of trustees (second-movers) causes these players to reciprocate (Return) more often and if so, whether this change in trustees' behavior engenders greater trust on the part of investors who move first. We find that, when minimal information on the trustee's prior-period choice is provided *following* the absence of such a reputational mechanism (treatment 1 to treatment 3), this change does indeed lead to a large and significant increase in both trust and reciprocity. However, reversing the order, when minimal information about trustees is initially provided and then removed (treatment 3 to treatment 1) we find no significant difference in the frequency of trust and reciprocity between these two treatments. In our fourth treatment, when the amount of information about trustees is increased to include the frequency with which the trustee has played return in *all* prior periods of the current supergame (i.e., there is "full" information), we find that such order effects disappear: the provision of the longer history of information about the prior decisions of trustees leads to significant increases in trust and reciprocity relative to the absence of such information, regardless of treatment order. Finally, in our fifth treatment, where investors must decide whether to purchase full information on the prior decisions of their matched trustee in the current supergame (provided freely in our fourth treatment), we find that on average, only one-fourth of investors choose to purchase this information so that the other three-fourths are in the dark about the prior behavior of their current trustee. Nevertheless, trust and reciprocity is significantly higher in this costly information treatment as compared with the baseline, no-information treatment.

We conclude that high levels of trust and reciprocity require the *availability* of individual reputational information as provided, for example, by a credit bureau and cannot be sustained by community enforcement of a social norm of good behavior. We further conclude that *longer* histories are more beneficial than shorter histories in the promulgation of reputational concerns.

2. Related Literature

We are not the first to explore the mechanisms supporting trust and reciprocity among anonymous strangers in repeated interactions. We build upon several prior theoretical and experimental studies.

2.1 Cooperation in the Infinitely Repeated Prisoner's Dilemma Game under Random Matching

Under anonymous random matching, Kandori (1992) shows that cooperation may be possible if all players adhere to a "contagious strategy" in which individuals who have not experienced a defection choose "Cooperation," and individuals who have either experienced a defection by their opponent or have defected themselves in the past choose "Defection." Specifically, he shows that for an infinite horizon and for any fixed population size, we can define payoffs for the Prisoner's Dilemma game that sustain cooperation in a sequential equilibrium.

As pointed out by Kandori (1992), there are two substantial problems associated with a "contagious equilibrium." First, when the population is large, the argument applies only to games with extreme payoff structures. Second, a single defection causes a permanent end to cooperation and this fragility may make the equilibrium inappropriate as a model for trade. Ellison (1994) extends Kandori's work and remedies these problems by introducing a public randomization device that allows for adjustment in the severity of the punishment. Compared to Kandori's (1992) results, the equilibrium in Ellison (1994) does not require excessive patience on the part of players and applies to more general payoff structures. Furthermore, given public randomizations, the equilibrium strategy supports nearly efficient outcomes even when players make mistakes with a small probability. More recently, Dal Bo (2007) has shown how social norms can be sustained by schemes that only punish deviators given sufficient information on opponent's past behavior.

Duffy and Ochs (2009) conduct an experimental test of Kandori's (1992) contagious equilibrium using groups of subjects who play an indefinitely repeated two-person Prisoner's Dilemma under different matching protocols and different amounts of information transmission. Their results show that, under fixed pairings, a social norm of cooperation emerges as subjects gain experience, while under random matching, experience tends to drive groups toward a far more competitive norm, even when some information is provided about the prior choices of opponents. Thus they conclude that random matching works to prevent the development of a cooperative norm in the laboratory. Camera and Casari (2009) address the same issue of cooperation under random matching, but focus on the role of private or public monitoring of the anonymous (or non-anonymous) players' choices and find that such monitoring can lead to a significant increase in the frequency of cooperation relative to the case of no monitoring.

In contrast to these papers, in this study we examine the indefinitely repeated "trust" game instead of the Prisoner's Dilemma game. Unlike the Prisoner's Dilemma game, the trust (or "investment") game (Berg et al., 1995) we study in this paper has 1) sequential moves and 2) no strictly dominant strategies. In particular, the first mover has an incentive to choose "trust" (rather than no trust) if he believes the second mover will reciprocate, while the second mover has an incentive to cheat (not reciprocate) if the first mover trusts him, but is indifferent between cheating and reciprocating otherwise. This game is more closely related to many real-world *one-sided incentive problems* found, for example, in credit markets or in transactions between buyers and sellers in cyberspace (e-Commerce), where two players move sequentially and only the second mover always wants to deviate from reciprocation in the one-shot game.[2] The one-sided incentive problem of the trust game may be a more promising environment for the achievement of a social norm of cooperation (trust and reciprocity) under anonymous random matching than the Prisoner's Dilemma game with its two-sided incentive problem. Furthermore we note that most real-world reputation systems are designed to monitor the behavior of "second movers". For these reasons, we think it is important to study the trust game under anonymous random matching and with various levels of information on second movers.

2.2 Repeated Trust Games

Xie and Lee (2011) theoretically extend Kandori's (1992) argument to the development of trust and reciprocity among anonymous, randomly matched players in the infinitely repeated trust game and provide sufficient conditions that support a social norm of trust and reciprocity as a sequential equilibrium in the absence of reputational information. The trust game experiment we report on in this paper satisfies the Xie and Lee conditions in all treatments, so that in the absence of any information about one's randomly determined opponents, a social norm of trust and reciprocity may be sustained by the threat to move to a contagious wave of distrust and confiscation. However, we also explore the notion that some information about opponents' prior behavior may help to sustain social norms of trust and reciprocity, as such information makes it easier for players to discern player types thus enabling reputational considerations.

There are several experimental papers on repeated trust games that relate to this study. Keser (2003) studies a trust game where "buyers" and "sellers" are randomly matched and prior to making investment decisions, buyers have either no information about their matched seller, or the seller's most recent rating (by another buyer), or the entire distribution of the seller's prior ratings (by other buyers). She reports that revelation of seller ratings significantly increases both trust and reciprocity relative to the baseline case of no information and that longer histories of ratings are more effective than shorter rating histories in increasing the levels of trust and reciprocity. Similar findings are reported by Bolton et al. (2004, 2005), Bohnet et al. (2005), Brown and Zehnder (2007) and Charness et al. (2011) using various versions of the trust game and varying the amounts and kinds of past information available to buyers and sellers as well as, in some studies, the matching protocols (strangers or partners). Engle-Warnick and Slonim (2006ab) are the only prior experimental studies examining trust and reciprocity in indefinitely repeated trust games, but they focus on the case of fixed pairings (a partners design).

These studies all differ from the one reported here in several important dimensions. First, since Keser

[2] Kandori (1992) has a formal definition of a "one-sided incentive problem" (Definition 4 on page 73). The concept requires that, only one of two parties has an incentive to deviate from the cooperative outcome, and there is a Nash equilibrium such that the payoff from the equilibrium is less than the payoff from the cooperative outcome for the party who has the incentive problem.

(2003), Bolton et al. (2004, 2005), Bohnet et al. (2005), Brown and Zehnder (2007) and Charness et al. (2011) investigate *finitely* repeated games and Engle-Warnick and Slonim (2006ab) study both finite and indefinitely repeated games under *fixed* pairings, none of these prior studies can rationalize trust and trustworthiness as an equilibrium phenomenon among anonymous, randomly matched players who have no information about the history of play of their partners as is the case in our study.[3] Thus they do not address one of the main questions we pose here: whether the mechanism that supports trust and reciprocity comes about through community-wide enforcement (fear of a contagious wave of distrust and confiscation) or from the provision of information on individual behavior (that affects the behavior of both the observed and those deciding whether to trust) or possibly some combination of both. Our nested treatment design allows us to carefully assess the extent to which these two mechanisms are operative. Second, prior research does not explain why the length of reputational information on second movers should matter; Keser (2003) for example, simply documents that a longer history of reputational information generates greater trust and reciprocity than does a shorter history. By contrast, here we provide a theoretical explanation for why the history length should matter and we provide some evidence in support of that prediction. Finally all prior studies have only examined the case where information on the past behavior of second movers is *freely* provided. However, in many real-world situations, reputational data are private information gathered by third parties and are only made available to individuals at some cost, e.g., credit reports. Thus we go a step further and (in one treatment) consider how behavior is affected if information on second movers is costly and second movers don't know whether information about them has been purchased or not. This asymmetric information treatment enables us to consider whether it is the *availability* of (costly) information (and perhaps less importantly the content of that information) that may suffice to sustain cooperative behavior.

Finally, we note that our paper is also related to the literature exploring the historic development of economic institutions among strangers. Greif (1989, 1993) and Milgrom et al. (1990) model a large number of traders who are randomly paired with each other in each period. Each pair is presumed to play a game similar to the trust game, where one party has an incentive to cheat the other by supplying goods of inferior quality or reneging on promises to make future payments. In this literature, institutions are seen as a way of avoiding the inefficiency of noncooperative equilibria. Greif and Milgrom et al. argue that the exchange of information on the identity of cheaters or the development of a mechanism which strengthens the power of enforcement can help to sustain cooperation. Dal Bo (2007) shows how random matching and a community enforcement mechanism can support unequal caste-type systems among identical agents that would not be sustainable under the standard two-agent fixed-matching protocol.

[3] Many experimental studies find that trust and reciprocity prevail under the conditions of complete anonymity and one-shot interaction. As these behaviors are inconsistent with all participants being payoff maximizers, they are often explained by psychological factors such as fairness, altruism, and expectation of inequality aversion etc. See, e.g., Berg et al. (1995), Bolton and Ockenfels (2000), Fehr and Schmidt (1999); Camerer (2003) provides a survey.

3. The Model

We briefly describe the model and its predictions for our experimental design. We adopt the notation of Xie and Lee (2011). The set of players $N = \{1, 2, \ldots, 2n\}$ is partitioned into two sets of equal size, the set of investors $N_I = \{1, 2, \ldots, n\}$ and the set of trustees $N_T = \{n+1, n+2, \ldots, 2n\}$. In each period, each investor is matched with a trustee according to the uniform random matching rule, and they play the binary trust game as a stage game. This procedure is infinitely repeated, and each player's total payoff is the expected sum of his stage game payoffs discounted by $\delta \in (0,1)$.

The trust game we study is depicted in Figure 1.[4] At the beginning of the game, the investor is endowed with an amount $a \in (0,1)$. If the investor decides not to invest, the game ends. The investor's payoff is a (the value of his outside option) and the trustee's payoff is 0. If the investor chooses to invest his endowment, this choice yields an immediate gross return of 1, but the division of this gross return is up to the trustee, who moves second. If the investor has invested, the trustee decides whether to keep all of the gross return for a payoff of 1 for himself and 0 for the investor or to return a fraction $0 < b < 1$ to the investor, earning a payoff of $1 - b$ for himself. Throughout we shall assume that $0 < a < b < 1$.

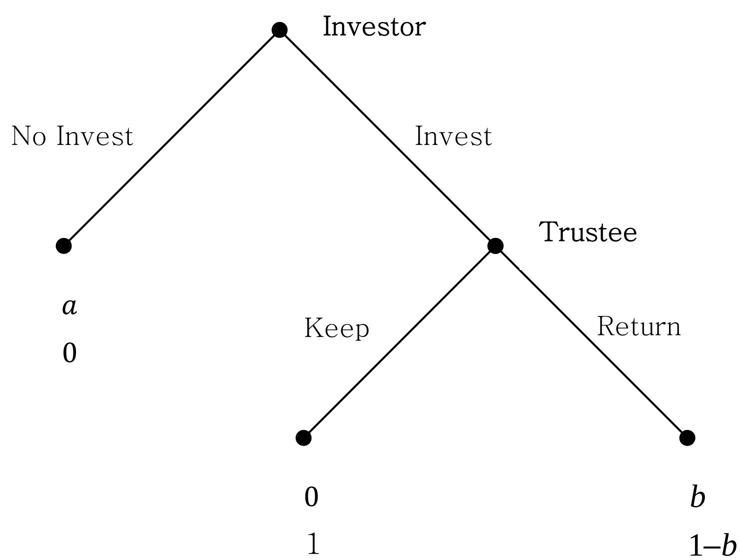

Figure 1: The Trust Game

If the game is played once, the unique subgame perfect equilibrium is for the investor not to invest as the trustee will always choose to play Keep. But since $a < 1$, this equilibrium is not efficient. The efficient outcome, where the investor invests and the trustee returns, *can* be achieved under the conditions of the "contagious equilibrium" of the infinitely repeated game, even if players are anonymously and

[4] In the trust game we study, both players have binary choice sets, a simplification necessary for the theoretical analysis that follows.

randomly re-matched after each period. We now turn to characterizing this contagious equilibrium.

3.1 Contagious equilibrium

Define the action No Invest as a "defection" by an investor and the action Keep as a defection by a trustee. Define *d-type* players as those whose history includes a defection either by themselves or by any of their randomly assigned partners. Otherwise, players are defined as *c-type* (cooperative) players.

Definition: *The "contagious strategy" is defined as follows: An investor chooses Invest if she is a c-type and No Invest if she is a d-type. A trustee chooses Return if he is a c-type and Keep if he is a d-type.*

The idea of the contagious strategy is that trust applies to the community as a whole and cannot be applied to individuals because of random anonymous matchings. Therefore, a single defection by a member means the end of trust in the whole community and a player who experiences dishonest behavior starts defecting against all of his opponents (Kandori, 1992). It is shown below that we can define payoffs for the trust game which allow trust and reciprocity to be a sequential equilibrium for any finite population.

To show that the contagious strategy constitutes a sequential equilibrium, it is sufficient to show that one-shot deviations are unprofitable after any history. In particular, Xie and Lee (2011) provide these conditions in the following lemma which puts constraints on investors' and trustees' incentives not to deviate from the contagious strategy both on the equilibrium path and off the equilibrium path.

Before stating the lemma, we first introduce the terms $f(\delta)$ and $g(\delta)$ which are functions of the period discount factor δ --for details of the construction of these terms see Appendix A. Conceptually, $f(\delta)$ represents the discounted sum of expected future payoffs – the gain-- to a trustee from *not* initiating a contagious wave of defection when all the other players in the community are c-types, and $g(\delta)$ represents the gain to a d-type trustee from deviating from defection (i.e., resuming to play Return) given that there is just one d-type investor and one d-type trustee (himself) in the current period. Thus, $f(\delta)$ and $g(\delta)$ are the discounted, expected payoffs to a trustee from avoiding triggering or slowing down the contagious strategy in the current period in different states of the world (i.e., when there are different numbers of d-type investors and d-type trustees in the community).

Lemma: *The contagious strategy constitutes a sequential equilibrium if*

$$a \geq \frac{n-1}{n} b \tag{1}$$

and

$$g(\delta) \leq b \leq f(\delta) \tag{2}$$

Condition (1) controls the investor's incentive to deviate from the contagious strategy off the equilibrium path. Due to the one-sided nature of the incentive problem, Invest is the best response to Return, so the investor has no incentive to deviate on the equilibrium path. Condition (1) requires that a d-type investor defect forever (never go back to Invest), even if she believes there is only one d-type trustee, which is the most favorable situation for investment. The left hand side of inequality (1), a, is the investor's opportunity cost from choosing Invest, and the right hand side of inequality (1) is the expected payoff to Invest given that there is only one d-type trustee so that the other $n-1$ trustees will choose Return.

An implication of condition (1) is that the existence of the contagious equilibrium requires a high outside option. For the development of a cooperative social norm, the concept of the contagious equilibrium requires a harsh punishment scheme. Not only are those who deviate from the desired behavior punished, but a player who fails to punish is in turn punished himself (Kandori, 1992). Thus an investor must defect forever once she is cheated upon. In order to prevent d-type investors from investing again off the equilibrium path, the outside option a must be sufficiently high. Notice that condition (1) becomes more restrictive when the group size n becomes larger. Outside of the laboratory, anonymous interaction typically requires a large group size. Thus condition (1) implies that the contagious equilibrium may be more difficult to sustain in the field.

Condition (2) controls the trustee's incentive to deviate from the contagious strategy both on the equilibrium path and off the equilibrium path. The first part of condition (2), $f(\delta) \geq b$, requires that the trustee's one-period gain from defection, b, must be less than or equal to the gains from *not initiating* a defection in the current period, $f(\delta)$. Thus, a trustee will not start a defection in the current period. The second part of condition (2), $g(\delta) \leq b$, implies that the period loss from attempting to slow down a contagious wave of defection, b, must be greater than the gains from slowing down the contagion when there are already other d-type players in the community. The latter restriction controls the trustee's incentive not to deviate off the equilibrium path. Finally, to show that there always exists a b between $g(\delta)$ and $f(\delta)$, Xie and Lee (2011) show that $g(\delta)$ is less than $f(\delta)$ for any δ greater than 0 given any finite population size. Intuitively, the trustee's payoff from not initiating the contagion (i.e., $f(\delta)$) is larger than the trustee's payoff from slowing down the contagious procedure (i.e., $g(\delta)$), since in the former case the trustee in consideration is the only potential d-type player in the community and the contagious procedure stops completely for the current period if this trustee chooses not to defect, while in the latter case the effect of this trustee choosing not to defect becomes much smaller with the existence of other d-type players in the community.

The lemma above is used in the proof of the following theorem, which states that we can find values for a and b in the trust game that satisfy the sufficient conditions of the lemma.

Theorem (Xie and Lee 2011): *Consider the model described above where* $2n \geq 4$ *players are randomly* *paired each period to play the infinitely repeated trust game. Then for any* δ *and* n*, there exist* a *and* b *such that (i)* $0 < a < b < 1$*; and (ii) the contagious strategy constitutes a sequential equilibrium in which (Invest, Return) is the outcome in every pair and every period along the equilibrium path.*

While other repeated game equilibria may exist under these conditions, the contagious equilibrium where (Invest, Return) is the outcome in every period is the most efficient of these equilibria, and therefore the focus of our analysis.

Remark 1: Note that the contagious equilibrium does not exist in the finitely repeated trust game by a simple backward induction argument.

3.2 Equilibria when information about trustees is available

In this paper, we consider as an alternative to anonymous, community-wide enforcement, environments where information on an individual trustee's past history of play can be observed by an investor prior to the investor making a decision to invest or not. We focus on the case of one-sided information flow (investors only view information on trustees and not vice versa) as this seems most appropriate for the trust game with its one-sided incentive problem, and because this information set-up also follows that of many real-world examples, e.g., credit markets, e-commerce. Specifically, we consider two different trustee histories that may be available to the investor: 1) "minimal information", where the investor observes only the action chosen by the trustee in the prior period (Keep, Return, or no choice) and 2) "full information", where the trustee's past history of decisions in all prior random matches with investors is revealed to the investor with whom the trustee is currently matched. We further consider an environment where full information is available to investors but at a cost, $c > 0$. The following propositions apply to such environments with costless or costly information on the trustee's history of play.

Proposition 1: *When information on the past behavior of trustees is free and full, the contagious strategy is not an equilibrium strategy.*

Proof: Consider the case where a d-type investor meets a c-type trustee in the current period. Under full information, the d-type investor can identify the trustee as a c-type player. According to the contagious strategy, the trustee should choose Return-given-Invest, and the investor, being a d-type should choose No Invest. However, given the trustee's strategy, the investor has an incentive to choose Invest since she can not only gain $b - a$ in the current period but she can also slow down the contagious process by not changing the current c-type trustee into a d-type trustee.

If the contagious strategy is no longer an equilibrium strategy, a natural question that arises is what *is* an equilibrium strategy when information is available on trustees? We propose the following:

Proposition 2: *When information on the history of a trustee's play is free and full and* $\delta \geq b$, *there exists an equilibrium in which the trustee continues to play the contagious strategy but investors play a strategy that is conditional on the information revealed about the trustee. Specifically, an investor chooses Invest if the trustee's history of play reveals the trustee to be a c-type and the investor chooses No Invest otherwise.*

Corollary 1: *When minimal information is provided freely, the strategy described in Proposition 2 is an equilibrium strategy only for a knife-edge condition* $\delta = b$.
Proof: See Appendix A.

Proposition 2 and Corollary 1 together indicate that if investors condition their investment decision on information about a trustee's prior behavior, an equilibrium involving complete trust and reciprocity will be easier to sustain in the case of full information than in the case of minimal information. Intuitively, the discount factor cannot be too high in the equilibrium under minimal information, since a d-type trustee will have an incentive to attempt to remove his bad reputation by engaging in one-shot good behavior in the current period so as to appear to be a c-type and attract investment in future periods. This problem does not arise in the case of full information because in that case it is impossible for a d-type trustee to change his type as perceived by investors.

Our final proposition applies to environments where the investor may choose to purchase full information about a trustee's past history of play at a per period cost of $c > 0$. The information purchase decision is private information; the trustee does not know whether or not his matched investor has chosen to purchase information. For this environment, we propose the following asymmetric equilibrium: only a fraction of investors choose to purchase information (or equivalently, investors choose to purchase information with some probability); a fraction of trustees always choose Return and the remaining trustees always choose Keep. For some intuition as to why there is a mixture of behavior in the equilibrium of this environment, suppose that all trustees always chose Return. Then investors would not need to purchase information, since the value of information is to distinguish trustees with a good reputation from those with a bad reputation. However if none of the investors purchased information yet they still invested with a positive probability, then trustees would have strong incentives to defect. Therefore, investors should play a mixed strategy with regard to the information purchase decision, provided the cost is small enough.

Proposition 3: *When information on the history of trustees' play is full and not too costly and* $\delta \geq b$, *there exists an equilibrium characterized by a vector of probabilities,* (q, γ, p), *where investors purchase* *information with probability* $q < 1$*, choose Invest if this information reveals the trustee to have always chosen Return, and choose No Invest otherwise. Investors who do not choose to purchase information choose Invest with probability* γ*. Fraction* $p < 1$ *of trustees always choose Return and fraction* $1 - p$ *always choose Keep. The most efficient such equilibrium obtains where* $\gamma = 1$.

Corollary 2: *When full information about trustees is available for purchase there also exists an inefficient, pure strategy equilibrium where investors never purchase information and never choose Invest and no trustee chooses Return.*

Proof: See Appendix A.

Proposition 3 says that when full information is available and not too costly (the cost conditions are given in the proof of Proposition 3), there exists an equilibrium in which only some investors purchase information about trustees and, consequently some trustees play Keep. Hence, an implication of making full information costly is that trust and reciprocity may be lower than when full information is costless. While there are many equilibria with positive levels of trust and reciprocity when information is costly (these are indexed by γ), we focus our analysis (as we have done previously) on the most efficient of these equilibria, which obtains when investors choosing not to purchase information always choose Invest ($\gamma = 1$).

Of course, as stated in Corollary 2, the inefficient equilibrium where all investors choose not to purchase information and never choose Invest and all trustees choose Keep always remains an equilibrium possibility. Thus, there is an empirical question as to whether information will be purchased in the costly information environment. We examine the latter question as well as all of our other theoretical predictions by designing and analyzing results from a laboratory experiment. We now turn to this exercise.

4. Experimental Design

Our main treatment variable concerns the information available to investors in advance of their investment decision though we also consider whether the horizon is indefinite or finite. We have five treatments. All treatments except the second one involve supergames with indefinite horizons. In the first, "no information" treatment (henceforth referred to as "No"), investors only know their own history of play and payoff in each period. Nevertheless, in this environment, full trust and reciprocity (the play of Invest and Return) can be supported under random anonymous matching via the contagious strategy. The second, no information, finite horizon treatment (referred to as "No-Finite") is similar to the first treatment in that subjects have no information about their randomly matched partner in each period but in this treatment subjects play the trust game for a known finite number of periods. In that case, trust and reciprocity cannot be supported by the contagious strategy as noted in Remark 1. In the third, "minimal information" treatment (referred to as "Min"), investors are informed of the prior-period decision of their current paired trustee, i.e., whether that trustee chose Keep or Return in the prior period of the current supergame, in the event the trustee had the opportunity to make a choice in the prior period; if the trustee did not have an opportunity to make a decision in the prior period, the information reported to the investor is "No Choice". In the fourth "full information" treatment (referred to as "Info''), investors are told the frequencies with which their currently matched trustee chose Keep or Return out of the total number of opportunities the trustee had to make either choice over *all prior periods* of the current supergame – called the Keep or Return ratios. The latter information is all that is necessary to label a trustee as either a c- or d-type, consistent with Propositions 1-3. In addition, investors in the Info treatment were also shown the trustee's actual, period-by-period history of play (Return, Keep or No Choice) for up to 10 prior periods of the current supergame.[5] Finally, in the fifth, "costly information" treatment (referred to as "Cost"), investors are not automatically provided with information on their paired trustee's previous choices as in the Info treatment; instead, individual investors can choose to purchase and privately view the same, full information record that was freely provided in the Info treatment at a small per period cost of c>0 points.

Each of our experimental sessions involved a single group of size $2n = 6$. We chose to work with groups of 6 subjects for several reasons. First, and most importantly, condition (1) for the existence of the contagious equilibrium in the trust game (where $a < b$) is more difficult to satisfy when n is large. On the other hand, we did not want the expected frequency of repeat matchings to be as high as in the minimal group size of 4. Second, we wanted to give the contagious equilibrium a chance to work; it is well known that the contagious equilibrium involving complete trust and reciprocity can collapse due to noise or trembles, and such noise is likely to increase with the size of the group.[6]

We chose a discount factor $\delta = .80$ and an indefinitely repeated *supergame* was implemented as follows. At the start of each supergame, subjects were randomly assigned a role as either the investor or trustee and they remained in that role for all periods of the supergame.[7] This design gave subjects experience with playing both roles across many supergames. In each period of a supergame, the 3 investors and 3 trustees were randomly and anonymously matched with one another for a single play of the stage game with all matchings being equally likely.[8] After playing the stage game, the results of the

---

[5] While we limited the period-by-period history of actions about a trustee to a maximum of 10 prior periods, the reported frequencies with which a trustee played Keep or Return were ***for all periods of the current supergame*** and this fact was made clear to subjects. Note further that the expected duration of a supergame, given our choice of $\delta = .80$, is just 5 periods.

[6] Camera and Casari (2009) offer a similar justification for their choice of a group size of 4. Duffy and Ochs (2009) look at groups of size 6 as well as larger groups of size 14 and find cooperation rates under random anonymous matchings are twice as high on average in groups of size 6 as compared with groups of size 14.

[7] In the instructions (Appendix B) we use neutral word "First Mover" for investor, "Second Mover" for trustee, and "sequence" for indefinitely repeated supergame. We also use "A" "B" "C" "D" to denote the investor and trustee's choices. See Appendix for instructions.

[8] This is the same matching protocol used by Duffy and Ochs (2009). Camera and Casari (2009) use a matching protocol wherein no two subjects are matched to play more than one supergame. In all treatments of our design, the assignment of roles (Investor, Trustee) was randomly determined at the start of each new supergame thereby distinguishing one supergame from the next.

game were reported to each pair of subjects and a 10-sided die was rolled. If the die came up 8 or 9, the supergame was declared over; otherwise the game continued on with another period.[9] Subjects were randomly rematched before playing the next period, though they remained in the same role in all periods of that supergame.[10] We told subjects that we would play a number of "sequences" (i.e., supergames) but we did not specify how many. For transparency and credibility purposes, we had the subjects take turns rolling the 10-sided die themselves and calling out the result. Our design thus implements random and anonymous matching, a discount factor $\delta = .80$, and the stationarity associated with an infinite horizon.

In the finitely repeated game subjects were randomly matched in every round and randomly assigned new roles at the start of each finitely repeated game just as in the indefinitely repeated game. However in the finite horizon treatment, no die was used as subjects were fully informed that each game would last exactly 5 periods. We chose a finite length of 5 periods as this is equal to the expected duration of an indefinitely repeated supergame, 1/(1-δ) under our choice of $\delta = .80$. Subjects in the finite horizon treatment played a total of eight 5-period games so that we would have approximately as many periods of data (40 periods) as in our indefinite horizon treatments.

The parameterization of the stage game used in all experimental sessions is given in Figure 2. In this figure, the terminal nodes of the tree give the number of points each type of subject earned under the three possible outcomes for each stage game played. This parameterization of the game was chosen to be consistent with our theoretical assumption that $a = 35 < b = 45 < 100$ and also satisfies the conditions (1) and (2) of the Lemma in the prior section given the choice of $n = 3$ pairs of players and the induced period discount factor $\delta = .80$ used in the sessions with indefinite horizon.[11] While other parameterizations are possible, we chose a parameterization that is *not* at the boundary of the conditions (1)-(2), but instead among randomly matched players.[12] The cost of purchasing information in the Cost treatment was set at 2 points, and satisfies restrictions given in the Proof of Proposition 3. The experiment was programmed and conducted using the z-Tree software (Fischbacher, 2007).

[9] Note that the die roll also provides a randomization device by which individuals can coordinate a halt to a contagious punishment phase as detailed in Ellison (1994).

[10] In all of the informational mechanisms discussed above, information on the trustee' behavior in previous supergames does not carry over when a new supergame begins. In the treatments where information is available, it is always available from the start of the second period of each supergame.

[11] The payoffs used in the experiment are those shown in Figure 2 which are the payoffs of Figure 1 multiplied by 100. Using a Mathematica program we calculated the value of $f(\delta) = 60.2$ and $g(\delta) = 24.2$ given our choice of $n = 3$ and $\delta = 0.8$. So the choice of b satisfies condition (2): $g(\delta) \leq b \leq f(\delta)$. The program is available upon request.

[12] In many experimental implementations of trust games, the trustee is given a positive endowment so as to avoid the possibility that the investor feels compelled (out of some sense of fairness) to invest. While this may be an issue in one-shot games, it seems less relevant in our repeated random-matching trust game, where (as we discuss below) all players are equally likely to be assigned the role of investor or trustee at the start of each new supergame, and are paid for all periods of all supergames played. Therefore, each subject in our design is effectively given the same "endowment" in expected terms. Related to the real life examples that motivate our paper, e.g., borrowing in credit markets, it also seems more reasonable to assume that only the first mover (bank) has an outside option (endowment); if a transaction does not occur, then the bank keeps its money while the borrower earns 0.

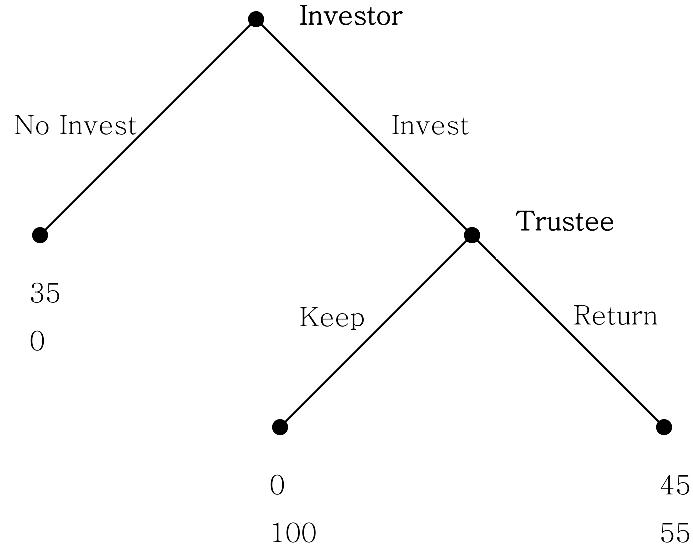

Figure 2: Stage Game Parameterization Used in the Experiment.

We used a within-subjects design in all sessions involving indefinite horizon supergames. Subjects began to play under one information condition and were switched to a second condition (and to a third in some sessions). In practice, there are at least 30 periods under each information condition – see Table 1 below. The treatment was stopped when the total number of periods under the treatment information condition exceeded 30 periods and the then-current supergame ended naturally via the die roll. Subjects were only informed of the change in an information condition when the switch took place, i.e., they did not know that a change was coming or our rule for implementing the duration of each treatment. For the sessions involving finite horizon supergames, there is only one treatment for the entire session.

We have in total 18 experimental observations (16 sessions) which we divide up into three main sets. The first set of 8 sessions (observations) examines whether providing investors with no information or minimal information on a trustee's prior behavior affects trust and reciprocity. We conducted four "No_Min" sessions (sessions that began with "no information" and later switched to "minimal information"), and four "Min_No" sessions following the opposite treatment order. In the second set of 2 sessions (4 observations) we explore whether the indefiniteness of the horizon matters for the level of trust and reciprocity observed under no information. We report results from two "No-Finite" sessions where subjects played the trust game under "no information" but for a known finite number of periods. The third set of 6 sessions (observations) investigates the effect of a longer history of information regarding trustees' prior behavior on trust and reciprocity and whether the possibility to purchase that "full" history of information at a small cost affects the frequency of trust and reciprocity. We conducted three "No_Info_Cost" sessions and three "Info_No_Cost" sessions. (Recall No means no information, Info means information and Cost means costly information). We reversed the order of the first two treatments to examine whether the treatment order matters. The Cost treatment is always the *last* treatment in this third set of sessions, as we wanted subjects to have experience with the full information ("Info") treatment before they faced a decision as to whether to purchase that same amount of information at a small cost. There are always 6 participants in all sessions involving indefinite horizon supergames (the first and third set). For the second set of sessions that involve finite horizon supergames, there are 12 participants who were divided into two groups of 6 at the beginning of the session and never had interactions with subjects in the other group during the entire session, so we can treat each group as an independent observation. The Instructions used in the "Min_No" and "Info_No_Cost" sessions are provided in Appendix B (instructions for the other treatment orderings are similar).

The motivations for this experimental design follow from our theoretical model. First, under our parameterization of the model the contagious strategy supports a social norm of complete trust and reciprocity among randomly matched anonymous players when there is an indefinite horizon and no information on trustees is available. However, we cannot exclude other equilibria, e.g., the social norm of no trust-and-reciprocity is another one. Thus it remains an empirical question as to whether community-wide enforcement suffices to support a social norm of trust and reciprocity and whether different informational mechanisms can help select different social norms. On the other hand, in the treatment that involves supergames with a finite horizon and no information on trustees, the contagious equilibrium does not exist. Therefore, comparing sessions with a finite horizon and those with an indefinite horizon under no information helps us evaluate whether indefiniteness of the horizon matters as theory suggests.

Second, since the collection, storage and dissemination of information is always costly for a society, a question of practical interest is how much reputational information is needed to significantly enhance the frequencies of trust and reciprocity. That is one motivation for why we consider both the Min and Info treatments. A second motivation comes from Proposition 2 and its Corollary which predict that full information on trustees can sustain an equilibrium involving complete trust and reciprocity under more general conditions than when there is only minimal information on trustees. Notice further, that the information reported to subjects in the Min treatment nests that of the No treatment while the Info treatment nests that of the Min treatment.

Finally, the Cost treatment recognizes that information on trustees' past history would be costly to gather and that such costs would likely be paid by the information consumers, i.e., the investors. The Cost treatment thus addresses the role of costly reputational information on trust and reciprocity—a more empirically relevant setting. The frequency of information purchases by investors is not revealed and importantly, trustees are *not informed* as to whether their paired investor purchased information about them or not. This asymmetry of information is public knowledge. Thus, on the one hand, if some fraction of investors choose to purchase information about trustees (and act according to the content of that information), their decisions can potentially provide a positive externality to the whole community due to the anonymity of matching and information purchase decisions. On the other hand, if trustees believe that some fraction of investors will not purchase information, they may behave similarly as in the No information treatment. Our Proposition 3 predicts this kind of mixed equilibrium.

All subjects were recruited from the undergraduate populations of the University of Pittsburgh and Carnegie Mellon University. No subject had any prior experience participating in our experiment. Subjects were given $5 for showing up on time and completing the experiment and they were also paid their earnings from all periods of all supergames played. Subjects accumulated points given their stage game choices (points are shown in Figure 2, the cost of information is set 2 points). Total points from all periods of all supergames were converted into dollars at a fixed and known rate of 1 point = ½ cent. In the finite horizon treatment, subjects played half as many periods as in the indefinitely repeated game so we increased the conversion rate to 1 point =1cent.

5. Results

Observations	No. of Supergames Treat1/Treat2/Treat3	Total No. of Periods Treat1/Treat2/Treat3	Avg. Payoff Treat1/Treat2/Treat3
No_Min1	7 / 9	38 / 46	$5.38 / $6.73
No_Min2	5 / 8	34 / 37	$6.39 / $8.00
No_Min3	8 / 2	35 / 37	$5.77 / $9.03
No_Min4	9 / 8	39 / 35	$6.72 / $6.91
Min_No1	9 / 8	44 / 37	$8.67 / $8.44
Min_No2	5 / 11	38 / 41	$6.79 / $5.75
Min_No3	9 / 7	34 / 38	$4.82 / $7.93
Min_No4	13 / 5	41 / 38	$5.59 / $3.92
No-Finite1	8	40	$8.95
No-Finite2	8	40	$13.83
No-Finite3	8	40	$17.08
No-Finite4	8	40	$16.43
No_Info_Cost1	6 / 7 / 9	42 / 37/ 40	$7.20 / $8.60 / $9.86
No_Info_Cost2	11 / 6 / 7	35 / 35 / 36	$7.13 / $7.94 / $7.92
No_Info_Cost3	7 / 8 / 8	34 / 32 / 31	$6.66 / $7.03 / $7.07
Info_No_Cost1	5 / 9 / 7	44 / 43 / 32	$9.92 / $8.37 / $7.62
Info_No_Cost2	8 / 4 / 8	35 / 43 / 35	$7.78 / $4.68 / $5.03
Info_No_Cost3	7 / 9 / 7	34 / 33 / 34	$7.25 / $5.65 / $5.53
Average	N/A	N/A	N/A

Table 1: All Experimental Observations

Table 1 provides basic characteristics of all sessions, specifically the number of repeated games for each treatment, 1, 2 or 3 of the session, the total number of periods played in each of those treatments, as well as the average payoff earned by subjects in each treatment. As Table 1 reveals, in sessions with multiple treatments, the two or three treatments of each session involved roughly similar numbers of periods which (as noted earlier) was by construction. Subjects earned on average, $16.40 in addition to their $5 show-up fee. The first set of 8 sessions all finished within 1.5 hours, the second set of 2 sessions finished in about 1 hour and the third set of 6 sessions finished within 2 hours. In the following subsections, we report results from the first, second and third set of sessions respectively, and then we analyze how players made use of their personal histories and how investors made use of the various amounts of information provided to them about trustees.

5.1 No_Min and Min_No Sessions

We first analyze whether a social norm of complete (100%) trust and reciprocity emerges in the indefinitely repeated game when there is *no* information on trustees and we compare that case with treatments where "minimal" information on trustees' prior-period action –Keep or Return-- is freely provided. We first calculated the average frequencies of 1) invest, 2) return conditional on investment (return-given-invest) and 3) invest-and-return for each No_Min or Min_No session.[13] Table 2 reports the average of these session-level frequencies.[14] To aid in reading these tables where the treatment orders differ, the clear (unshaded) areas indicate average frequencies for the No treatment and the grey (shaded) areas indicate average frequencies for the Min treatment. Table 2 provides support for our first two experimental findings.

		1st treatment	2nd treatment	No. of Sessions/Obs.
Frequency of Invest	No_Min	0.487	0.699	4
	Min_No	0.467	0.508	4
Frequency of Return-given-Invest	No_Min	0.605	0.851	4
	Min_No	0.714	0.662	4
Frequency of Invest-and-Return	No_Min	0.296	0.610	4
	Min_No	0.348	0.384	4
Average Payoff per Period (in points)	No_Min	33	40	4
	Min_No	33	34	4

Table 2: Average Frequencies and Per Period Payoffs for Min_No and No_Min Sessions

Finding 1: *A social norm of complete trust and reciprocity (100% Invest-and-Return) is not achieved in*

[13] We will sometimes equivalently refer to Invest as "trust," Return-given-Invest as "reciprocate" and Invest-and-Return as "trust and reciprocity."
[14] Average frequencies in Tables 3-4 and Table C1 (Appendix C) are also constructed using the same approach.

the absence of information on trustees' behavior, contrary to a theoretical possibility described by the contagious strategy.

While the payoffs for the game were chosen so that the contagious strategy supports an equilibrium of complete trust and reciprocity, other equilibrium possibilities cannot be ruled out, for example zero trust and reciprocity by all players remains an equilibrium strategy. As Table 2 reveals, the frequency of Invest-and-Return averages around one-third across all sessions of the No treatment and is never zero. Similarly, the observed frequencies of Invest (trust), Return-given-Invest (reciprocity) are also less than 100% but considerably greater than zero across all sessions of the No treatment.

Our next finding makes use of our within-subjects design to compare behavior with and without minimal information.

Finding 2: *The free provision of minimal information to investors on the prior-period action of their matched trustee leads to a significant increase in trust and reciprocity but only if the provision of this information follows the treatment in which investors receive no information about trustees.*

For the No_Min sessions, the provision of minimal information about the trustee's prior-period play in the second half of a session leads to marginally significantly larger frequencies of Invest, Return, and Invest-and-Return compared with the corresponding 4 session-level frequencies in the NO treatment (Wilcoxon signed ranks test, p=0.0625 for all three tests).[15] However, when minimal information is provided in the first half of a session as in the Min_No treatment, then none of the frequencies of Invest, Return-given-Invest, and Invest-and-Return are significantly different from the corresponding frequencies in the No treatment (Wilcoxon signed ranks test, p>0.4 for all three tests, 4 session-level obs.).

Since the subjects were not informed of the second treatment until the change in treatment mid-way through a session, we can regard the *first* treatment of each session as an independent observation that was not influenced by other treatments. We find that regardless of whether the first treatment is No or Min information, the 8 session-level frequencies of Invest and Invest-and-Return are not significantly different from one another. However, the frequency of Return-given-Invest is significantly higher when minimal information is available in the first treatment than when it is not (Robust Rank Order test, p<0.05[16]). The latter finding implies that the provision of minimal information has a more significant effect on the behavior of trustees than on the behavior of the recipients of minimal information-- the investors!

---

[15] For this nonparametric test and all nonparametric tests that follow, we always use session-level treatment averages as our unit of observation.

[16] We use the robust rank order test rather than the more commonly used Wilcoxon-Mann Whitney test as the latter presumes that the two samples being compared come from distributions with the same second- and higher order moments and we have no a priori reason for believing this is the case.

If we restrict attention to just the second treatment of each session, after subjects have gained some experience, we find similar results: regardless of whether the second treatment is No or Min information, the 8 session-level frequencies of Invest and Invest-and-Return are not significantly different from one another, but the frequency of Return-given-Invest is marginally significantly higher when minimal information is available than when it is not (Robust Rank Order test, p=.10).

We next consider all 8 session-level observations from the No-treatment and find that there are no significant differences in the frequencies of Invest, Return-given-Invest, and Invest-and-Return when the No treatment is in the first half of the No_Min sessions or in the second half of the Min_No sessions. However, using the 8 session-level observations from the Min treatment, all frequencies of Invest, Return-given-Invest, and Invest-and-Return are significantly higher when minimal information is provided in the second half of the No_Min sessions than when it is provided in the first half of the Min_No sessions (Robust Rank Order test, $p<0.05$ for all three tests). Together, these findings suggest that when minimal information is provided to subjects who have suffered from the absence of reputational information as in our No_Min sessions, they learn to use the minimal information more effectively than subjects who begin interacting with minimal information and then lose access to that information as in our Min_No sessions.

5.2 Finite Horizon Treatment with No Information

We next focus on the no information environment of our baseline treatment and consider the effect of changing the horizon from an indefinite one to a finite one where subjects play the trust game for a known, fixed horizon of 5 periods. This finite horizon treatment rules out the possible use of contagious strategies supporting full trust and reciprocity via standard backward induction arguments. Table 3 reports the average frequencies of Invest, Return-given-Invest and Invest-and-Return for this treatment from 4 observations involving 6 subjects each who played eight 5-period trust games under no information.

	Averages over All 5 Periods of 8 Games	Averages over 1-4 Periods of 8 Games	No. of Sessions/Obs.
Frequency of Invest	0.544	0.580	4
Frequency of Return-given-Invest	0.577	0.591	4
Frequency of Invest-and-Return	0.379	0.419	4
Average Payoff per Period (in points)	35	36	4

Table 3: Average Frequencies and Per Period Payoffs for No-Finite Sessions

These averages are reported for all 40 periods (first column) or for a total of 32 periods of the 8 finitely repeated games with the exception of the fifth and final period (so as to minimize the impact of end-game effects). Regardless of which averages we use, we find none of the aggregate frequencies of Invest, Return-given-Invest and Invest-and-Return are significantly different in the "No-Finite" treatment as compared with those in the first treatment of "No_Min" and "No_Info_Cost" sessions (Robust Rank Order test, $p > 0.1$ for all three tests, 11 session level obs.). We summarize this result as follows:

Finding 3: *There is no significant difference between the aggregate frequencies of invest, return-given-invest and invest-and-return in the finite horizon no information treatment and in the indefinite horizon no information treatment. This result does not depend on whether we use data from all 5 periods or just the first four periods of the No-Finite treatment.*

5.3 No_Info_Cost and Info_No_Cost Sessions

Table 4 reports the frequencies of Invest, Return-given-Invest, and Invest-and-Return from the third set of sessions that involved three different treatments. As before, the clear (unshaded) cells of Table 4 show frequencies from the No treatment, the grey shaded cells show frequencies from the Info treatment, and now the light grey shaded cells show frequencies from the Cost treatment (always the 3rd treatment).

		1st treatment	2nd treatment	3rd treatment	No. of Sessions/Obs.
Frequency of Invest	No_Info_Cost	0.632	0.854	0.893	3
	Info_No_Cost	0.817	0.435	0.583	3
Frequency of Return-given-Invest	No_Info_Cost	0.694	0.920	0.885	3
	Info_No_Cost	0.941	0.561	0.817	3
Frequency of Invest-and-Return	No_Info_Cost	0.434	0.786	0.796	3
	Info_No_Cost	0.770	0.264	0.502	3
Average Payoff per Period (in points)	No_Info_Cost	38	45	46	3
	Info_No_Cost	44	32	37	3

Table 4: Average Frequencies and Per Period Payoffs for No_Info_Cost and Info_No_Cost Sessions

Recall that in the Info treatment the investor sees the matched trustee's aggregate Ratio of Return in all previous periods of the current supergame as well as the trustees' actions chosen in up to the 10 most recent periods of the current supergame, while in the Cost treatment that same information is only available at a cost of 2 points per period paid by the investor in advance. Similar to subsection 5.1, we start with a within-subject analysis and then move to a between-subject analysis.

Notice first that, consistent with Finding 1 for the No_Min and Min_No sessions, the frequencies of trust and reciprocity in the No treatment are all greater than zero, but the social norm of complete trust and reciprocity is not supported in that the aggregate frequency of Invest-and-Return remains low, averaging again around one-third over all sessions.

Comparing the impact of providing free and full information on trustees –the Info treatment-- with the No information treatment, we have the following:

Finding 4: *The free provision of a longer history of information to investors on the past behavior of their matched trustee (Info) leads to a significant increase in trust and reciprocity relative to the No treatment.*

Support for Finding 4 is found using the 6 session-level differences in the frequencies of Invest, Return-given-Invest and Invest-and-Return between the No and Info treatments (one-tailed Wilcoxon signed ranks test, $p<0.05$ for all three tests).

Consider next the Cost treatment where investors are provided with the possibility to purchase information at a cost of 2 points per period. Consistent with Finding 4, the frequencies of trust and reciprocity in the Cost treatment remain significantly greater than when no information is available (one-tailed Wilcoxon signed ranks test, $p<0.05$ for all three tests, 6 session level obs.). Furthermore, we also have the following (one-tailed Wilcoxon signed ranks test, 6 session level obs. at the 10% significance level):

Finding 5: *The frequencies of Invest (trust) and Invest-and-Return (trust and reciprocity) are not significantly different in the Info treatment and Cost treatment. The frequency of Return-given-Invest (reciprocity) is marginally significantly higher in the Info treatment than in the Cost treatment.*

Finding 5 provides mixed support for Proposition 3. On the one hand, trustees seem to recognize that when information is costly not all investors will purchase it and consequently they do not play Return as often as in the Info treatment. On the other hand, according to Proposition 3, investors should invest less frequently in the Cost treatment than in the Info treatment but in the data there is no significant difference in these investment frequencies. Finding 5 suggests that trustees' behavior may be more sensitive to the change in information treatment than investors' behavior, an observation that is consistent with earlier findings from the No_Min and Min_No sessions. Another explanation for the non-significance in investors' trusting behavior between the Info and the Cost treatment is that the difference in the frequency of reciprocity between these two treatments is not large enough to change the investor's best response. Indeed, when the average return rate is above around 77%, the investor's best response is Invest. We find that the frequency of return-given-invest exceeds 77% in all sessions of both the Info and Cost treatments with the exception of one session of the Cost treatment.

In our design, there is no information available in the first period of a supergame ("sequence") in any treatment, so an alternative way to examine the treatment effect is to compare the frequencies of trust and reciprocity excluding the first periods. We can confirm that Findings 4-5 continue to hold when the first periods are excluded from the data analysis -- the frequencies of Invest, Return-given-Invest, and Invest-and-Return excluding the first periods are reported in Appendix C.

Although the frequencies of trust and reciprocity are significantly increased when information is freely provided or available for purchase, one may wonder whether efficiency is enhanced by information provision, especially in the case where information is costly. The final two rows of Table 4 present the average per period payoff in points that subjects earned under each treatment (taking into account information purchase costs, if any).

Finding 6: *Players' average payoffs are significantly increased when full information is freely provided or available for purchase compared with the case where information is not available.*

Support for Finding 6 comes from analysis of the 6 session-level observations we have on payoffs in the No information treatment versus either the Cost or Info treatments (one-tailed Wilcoxon signed ranks test, $p<0.05$ for both tests).

As before, we can treat the first treatment in these 6 sessions as independent observations. Consistent with Finding 4, we find that providing full information in the first treatment (Info) significantly increases the frequencies of Invest, Return-given-Invest, and Invest-and-Return relative to the case where no information is provided in the first treatment (No) (one-sided Robust Rank-Order test, $p<0.05$). The same result holds if we compare the case where no information or full information is provided in the second treatment of each session; that is, experience is not the driving factor. This evidence again shows that the amount of information provided about trustees matters for enhancing trust and reciprocity.

We look for order effects by examining the third, Cost treatment of all 6 sessions. We find that none of the frequencies of Invest, Return-given-Invest and Invest-and-Return in the Cost treatment are significantly different across the 6 sessions. This finding implies that there is no significant order effect, that is, whether the subjects experience the No treatment or the Info treatment first and then switch to the other treatment does not significantly affect their behavior in the third and final Cost treatment. Finally, we also check whether subjects' behavior in the No treatment and Info treatment is different when the same treatment is the first or the second treatment in the session. We find that there is no significant difference for all the cases except that the frequency of Invest in the No treatment is marginally significantly higher when the No treatment is the first treatment than when it is the second treatment (Robust Rank-Order test, $p<0.1$). [17]

[17] These conclusions regarding order effects should be treated with some caution. Table 4 reveals a non-trivial difference in the average frequency of invest-and-return in the Cost treatment when the session starts with the No treatment as compared with when the session starts with the Info treatment. Indeed, a regression analysis (Table 9 below) reveals a significant order effect for this frequency in the Cost treatment. Thus, the insignificant non-parametric test for order effects may be due to the small number of independent observations (three per treatment order).

Finally, we also examined whether the frequencies of Invest, Return-given-Invest and Invest-and-Return are greater under free, full information as compared with the free, minimal information used in our first set of experiments. In the case where free information (minimal or full) was provided in the *first* treatment of a session, we find that the frequencies of Invest, Return-given-Invest, and Invest-and-Return are all significantly higher in the case of full information (one-sided Robust Rank-Order test, $p<0.05$ for all three tests). However, when free full or free minimal information are the second treatment of a session, following a first treatment of no information, the same three frequencies are not significantly different from one another (one-sided Robust Rank-Order test, $p>0.1$ for all three tests). Summarizing, we have:

Finding 7: *The frequencies of Invest, Invest-given-Return and Invest-and-Return under free and full information are greater than or equal to those observed under minimal information, depending on the treatment order.*

Summarizing the results of the paper to this point, we have several important findings. First, absent reputational information on the prior behavior of trustees, a social norm of complete trust and reciprocity is not achieved despite the theoretical possibility under the contagious equilibrium. Second, whether the game has a finite horizon or indefinite horizon does not affect significantly the level of trust and reciprocity under no information. Third, providing minimal information on trustee's prior behavior has some effect on Invest, Return-given-Invest and Invest-and-Return relative to the case of no information, but this effect depends on whether minimal information comes after a period of no information. Fourth, providing a longer history on trustees' prior behavior has a larger and more consistent effect on trust and reciprocity than minimal information on trustees' prior-period choices compared to the benchmark of no information. The effect of the longer history of information is robust regardless of whether that information is provided before or after the No treatment, according to both a within-subjects and between-subjects analysis of session-level data. Finally, and perhaps most importantly, making full information available at a small cost yields outcomes that are similar to those observed when full information is provided automatically and without cost.

5.4 Heterogeneous Behavior in the No Information Treatments

Our aggregate results indicate that when no reputational information is available, the frequency of Invest-and-Return (trust and reciprocity) in the indefinitely repeated game is about 33%, far below an equilibrium possibility of complete trust and reciprocity as supported by the contagious strategy. In this section we look more deeply into the data for evidence of heterogeneity in play of the trust game both at the supergame level and at the individual level.

We first ask whether the low overall frequency of trust and reciprocity in the indefinitely repeated game without information feedback is uniform across supergames or whether there is heterogeneity across supergames in the frequencies of invest-and-return. Figure 3 shows the distribution of supergames over 10 different frequencies of invest-and-return (horizontal axis), using data from the no information treatment in the "No_Min" and "No_Info_Cost" sessions where the no information treatment was first and so can be regarded as being independent of contamination from other treatments (same data source for Figure 4). The numbers above each bar in Figure 3 represents the average period length (duration) of all supergames in that frequency bin. We observe that supergames with Invest-and-Return frequencies between 30-40% are the most frequently observed arising in 25% of all supergames, however, both higher (e.g. 60-70%) and lower (e.g. 10-20%) frequencies of Invest-and-Return also have some mass in the distribution. We conclude that the distribution of the frequency of Invest-and-Return is widely dispersed, as opposed to being concentrated at the extremes (for example, 2/3 of supergames having frequencies of Invest-and-Return of 0% while the remaining 1/3 having frequencies of Invest-and-Return of 100%, which would be consistent with the contagious equilibrium being played in some of the supergames and the no trust equilibrium being played in the other supergames). We note that, more generally, many equilibria are possible in the indefinitely repeated trust game that we study, not just the extreme outcomes.

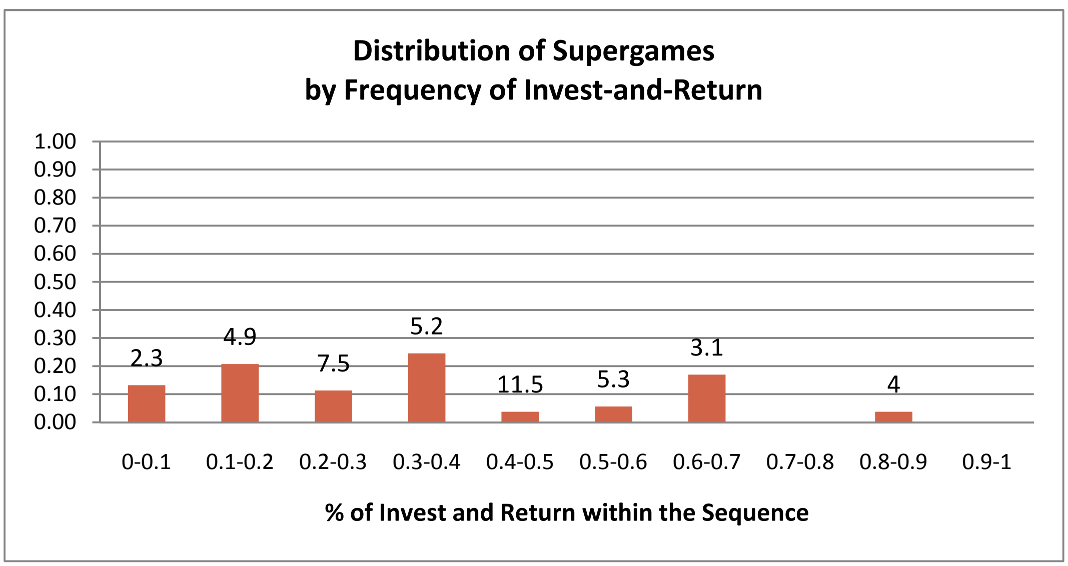

Figure 3: Distribution of Supergames by Frequencies of Invest-and-Return (All Supergames in No Treatment of No_Min and No_Info_Cost Sessions)

Figure 4 reports the average frequency of Invest-and-Return in each of the first five periods of all supergames that had five or more periods in the No information treatment of all No_Min and No_Info_Cost sessions. We chose such supergames which comprise about half (49.1%) of all supergames of the No information treatment because we wanted to allow enough periods for evidence of a contagious pattern to emerge and 5 periods is also the expected duration of a supergame. We observe that the average frequency of Invest-and-Return in the first period of these supergames is 48.7%. Given that this average is less than 100%, the contagious strategy predicts that average frequencies of Invest-and-Return in subsequent periods 2, 3, 4 and 5 should be monotonically decreasing. Figure 4 reveals a slight but statistically insignificant (*p=.30*) negative and clearly non-monotonic trend in these average frequencies over the first five periods. This provides further evidence against the notion that subjects were employing the contagious strategy in these supergames.

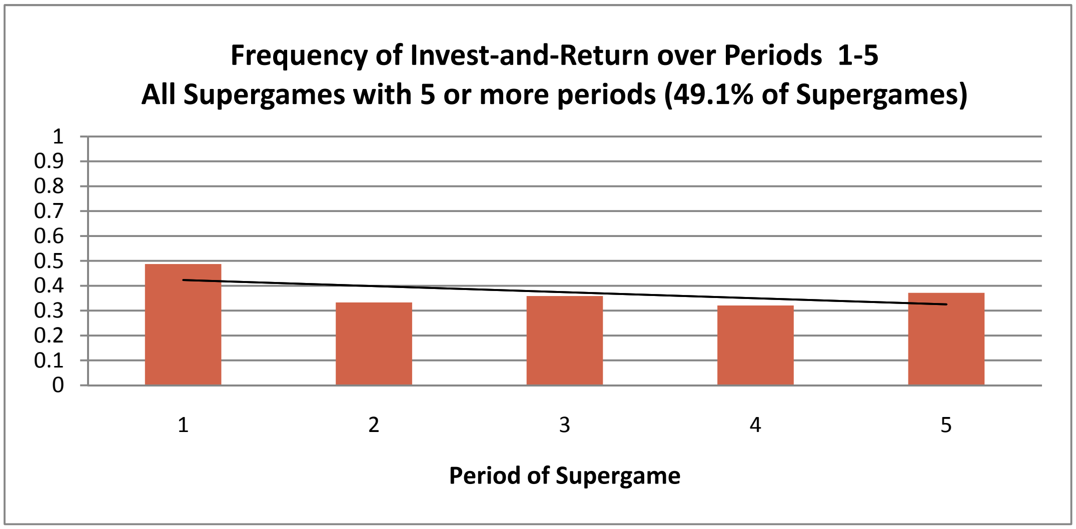

Figure 4: Frequency of Invest-and-Return by Period of Supergame
(All Supergames with 5 or more periods in No Treatment of No_Min and No_Info_Cost Sessions)

We next look for evidence of *individual* heterogeneity by examining whether and how personal histories may have played a role in subjects' decisions to invest and to return conditional on investment. For this analysis we use session-level averages from all No Information, indefinite horizon treatments (all 14 sessions). Following the logic of the contagious strategy, let us label each subject in each period of each supergame as a "c-type" or a "d-type" player according to his or her history of play so far in that supergame. Recall that d-type players are those whose history includes a defection either by themselves or by any of their randomly assigned partners. Otherwise, players are defined as c-type players. We further divide all the c-type players into c1-type players, who are automatically defined as c-type in the first period, and c2-type players, who have not yet experienced defection in later periods of the supergame. Similarly, we divide all the d-types into d1-types, who initiate a defection, and d2-types who become d-types by experiencing a defection by another player.

Table 5 reports overall average frequencies of Invest and Return-given-Invest by Investors and

Trustees respectively, based on 14 session-level averages from all NO information treatment sessions. Using a Wilcoxon signed ranks test on the session average frequencies (14 obs. per test), we find that the frequencies of Invest or of Return-given-Invest vary significantly across the different player types in all but one pairwise comparisons ($p<.05$).[18] Table 5 also reports the average frequency of the various player types. We see that in 61% of all periods, Investors can be classified as d-types while in 51% of all periods, Trustees can be classified as d-types; this amounts to a substantial departure from the equilibrium path that supports complete (100%) trust and reciprocity. Notice however, that when subjects are on the equilibrium path (when players can be classified as c-types), the frequency of invest and the frequency of return-given-invest are high, averaging in excess of 75%. The departure from 100% invest or return-given-invest among c-types appears mainly due to the behavior of c1-types who choose Invest or Return-given-Invest less than 68% of the time; c2-types, who have neither experienced nor initiated a defection, are investing or returning above 95% of the time. Among d-types, subjects who initiate a defection (d1-types) consistently continue to defect, but surprisingly, subjects who experience another's defection first (d2-types) continue to choose to invest or to return given investment with a rather high frequency: 42% for investors and 67% for trustees. We summarize these findings as in Finding 8.

Player Type	Investors		Trustees		No. of Sessions/Obs.
	% Invest	Frequency of Types (All Periods)	% Return-given-Invest	Frequency of Types (All Periods)	
C type	0.818	39%	0.788	49%	14
--C1 type	0.676	--20%	0.651	--27%	14
--C2 type	0.970	--19%	0.955	--22%	14
D type	0.309	61%	0.564	51%	14
--D1 type	0.200	--31%	0.358	--17%	14
--D2 type	0.417	--30%	0.667	--34%	14

Table 5: Behavior Conditional on Personal Histories/Types (All No Info Treatment Data)

Finding 8 *Behavior differs according to whether subjects are classified as c- or d-types. However, there are substantial departures from the contagious strategy supporting an equilibrium of complete trust and reciprocity in the No Information treatment. On the equilibrium path (when subjects can be classified as c-types) subjects initiate a defection in the first period more than 30% of the time. Off the equilibrium path (when subjects can be classified as d-types) they do not consistently punish as the contagious strategy requires.*

5.5 Use of Information in the Treatments with Information

[18] The exception is that the difference in the frequency of return-given-invest by c1- and d2-types is only marginally significant p=0.09.

This subsection focuses on how investors' behavior varied with the information they had about trustees' past decisions in the Min, Info and Cost treatments. We first analyze the effect of no versus minimal information. We then examine the effect of minimal versus full information before finally comparing the case of free versus costly full information.

5.5.1 Minimal Information versus No Information

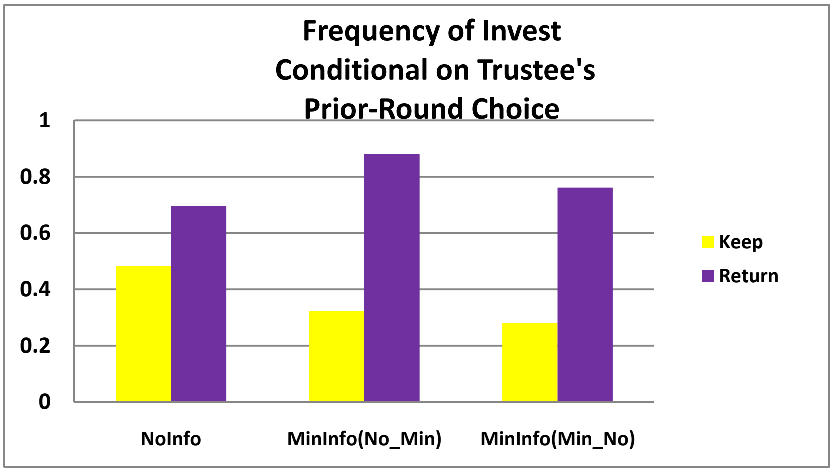

Figure 5: Frequency of Invest Conditional on Trustee's Prior-Period Choice

Figure 5 presents the frequency of Invest conditional on the prior-period choice of the investor's current matched trustee (i.e., Keep or Return) in the No and Min treatments of the No_Min and the Min_No sessions, respectively. In the No information treatment, there is no significant difference in these conditional frequencies of Invest (Wilcoxon signed ranks test on 8 session-level observations $p>0.1$); while Figure 5 suggests there is a difference, this is owing to investment being higher in sessions where the return rate is also higher, but within the same sessions the difference is not significant. By contrast, in the Min information treatment, the frequency of Invest is significantly higher when the prior period behavior of the trustee is Return than when it is Keep (regardless of the treatment order – one-sided Wilcoxon signed ranks test $p<.10$ for both Min treatments, 4 observations each).

Finding 9: *The frequency of Invest is significantly higher when the matched trustee's prior-period choice is "Return" than when it is "Keep" in the Min treatment but not in the No treatment.*

Finding 9 has to be treated with some caution as the prior behavior of trustees is not an experimental manipulation and may well represent an endogenous response to the prior trust exhibited by investors. Indeed, in the No information treatment we observe higher frequencies of Invest when the current matched trustee returned in the prior period as compared with when that trustee chose Keep, though this difference is not statistically significant. Relatedly, Figure 5 also suggests that investors' behavior is not 100% dependent on the minimal information they receive about trustees' prior period behavior. Indeed, we observe that investors still invest with a positive probability even if the minimal information shows that their current partner defected (kept) in the prior period, and symmetrically, investors do not always invest when their partner is revealed to have cooperated (returned) in the prior period.

5.5.2 Minimal Information versus Information

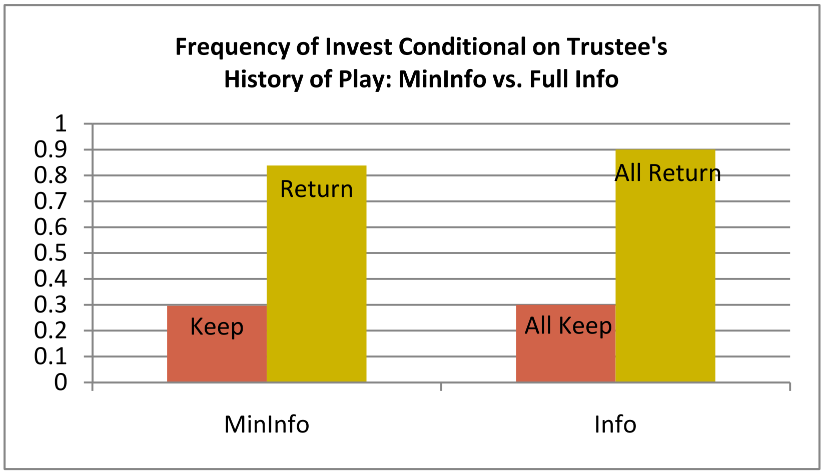

Figure 6: Frequency of Invest Conditional on Trustee's History

In the case of free information on the full histories of trustees' behavior, Proposition 2 states that there exists an equilibrium in which the trustee continues to play the contagious strategy but investors play a strategy that is conditional on the information revealed about the trustee. In particular, complete trust and reciprocity is an equilibrium possibility in this information environment. In Figure 6 we compare the conditional frequency of Invest decisions by investors in the Min Info treatment with that in the Info treatment. For the Min treatment, the conditional frequency is based on whether the investor's currently matched trustee was revealed to have kept or returned in the prior period. For the Info treatment, the conditional frequency is based on whether the investor's currently matched trustee was revealed to have always kept (All Keep) or always returned (All Return) in all prior periods of the current supergame. Figure 6 provides support for the following:

Finding 10: *The frequency of Invest, conditional on a history of "All Return" (Info treatment) is significantly greater than the frequency of Invest conditional on a prior period history of "Return" (Min treatment). There is no significant difference in the frequency of Invest conditional on a history of "All Keep" (Info treatment) or a prior period history of "Keep" (Min treatment).*

Finding 10 is based on the results of a robust rank order test using 14 session-level observations, $p \leq .05$. Based on Proposition 2 (and Corollary 1), we would expect a high frequency of return conditional on a prior history of Return in the Info treatment but not in the Min treatment. While the very high frequency of return conditional on investment in the Info treatment is consistent with this prediction, the lower but still high frequency of return in the Min treatment is not. We note further that while there is a significant difference in the conditional frequencies of Invest between these two treatments, the difference is not very large.

5.5.3 (No) Information versus Costly Information

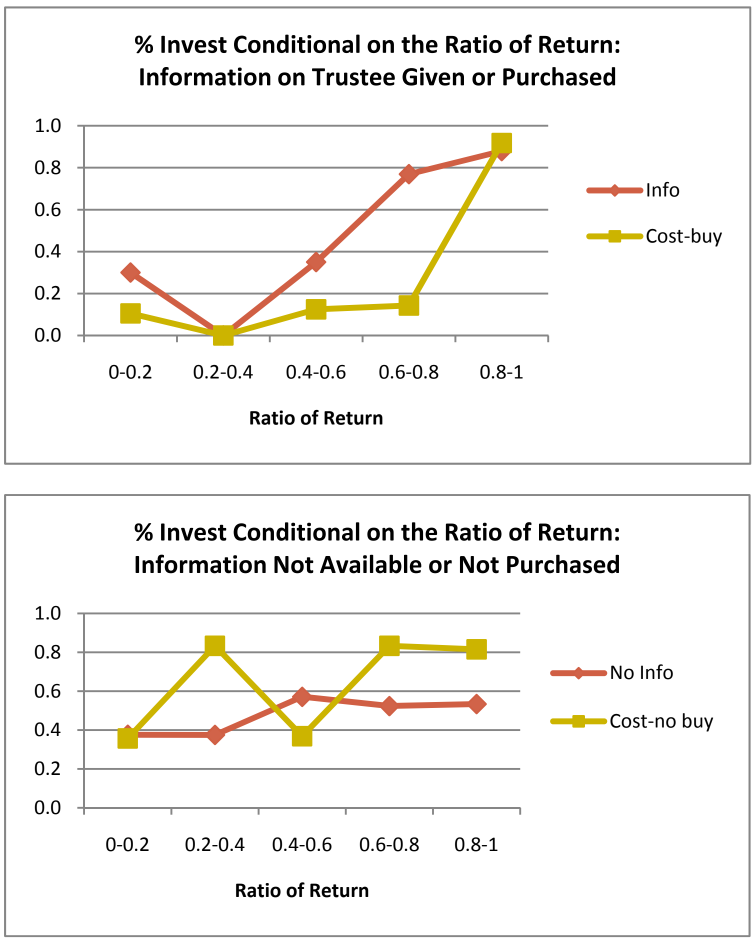

Figure 7: Frequencies of Invest Conditional on Trustee's Aggregate Ratio of Return

Figure 7 shows the frequency of Invest conditional on the matched trustee's Ratio of Return in the current supergame for the (free) Info and Cost treatment-when information was purchased (top panel) and for the No and Cost treatment-when information was not purchased (bottom panel). Recall that this ratio, as well as the trustee's history of play in the 10 most recent periods, was revealed to investors in the Info treatment and was available for purchase in the Cost treatment. Here we compare the frequency of Invest in these two treatments when the trustee's Ratio of Return is allocated into 5 non-overlapping bins for the Ratio of Return. Figure 7 provides support for the following finding:

Finding 11: *When investors have information on trustees, frequencies of Invest are generally increasing in the Ratio of Return by the matched trustee.*

We further observe from the top panel of Figure 7 that investors purchasing information do not choose Invest with a high frequency unless the trustee's Ratio of Return is also very high (0.8-1), which is likely owing to the cost these investors had to pay for information about their trustee.

How does behavior in the Cost treatment compare with the equilibrium prediction of Proposition 3? Table 6 provides some answers.[19] First, the most efficient equilibrium— the benchmark we use in our analysis of the data— has $\gamma = 1$, i.e., investors who do not purchase information choose Invest 100% of the time. As Table 6 reveals, this prediction comes close to holding true in 4 of our 6 sessions.

	% of Invest conditional on No Information Purchase		% of Information Purchase		% of Return given Invest	
Session	Data	Equil. pred.	Data	Equil. pred.	Data	Equil. pred.
No_Info_Cost1	1.000	1	0.226	0.563	0.992	0.943
No_Info_Cost2	0.907	1	0.368	0.563	0.775	0.943
No_Info_Cost3	0.882	1	0.261	0.563	0.889	0.943
Info_No_Cost1	0.905	1	0.013	0.563	0.989	0.943
Info_No_Cost2	0.164	1	0.247	0.563	0.838	0.943
Info_No_Cost3	0.407	1	0.333	0.563	0.625	0.943
Averages	0.725	1	0.245	0.563	0.877	0.943

Table 6: Comparison with Equilibrium Prediction for Cost Treatment

Conditional on $\gamma = 1$, the efficient equilibrium prediction calls for investors to purchase information on trustees 56.3% of the time. As Table 5 reveals, the actual frequency of information purchase by investors is significantly lower, averaging 24.49% (t-test p<.01).[20] Combined with the finding that the frequency of trust and reciprocity is not significantly different between the Info and Cost treatment, we conclude that the *mere availability* of information creates a large positive externality for investors. While the overall frequency of Invest in the Cost treatment is 65.55% when information is purchased, as Table 6

[19] All analysis here excludes the first period of each supergame, where information is not available to purchase.

[20] In our experimental design, trustees were not informed ex-post as to whether or not their matched investor had purchased information about their (the trustee's) history of play, as such choices are usually private information in real-world settings. However, this design choice may have led investors to purchase information less frequently than in the mixed strategy equilibrium of the theory.

reveals this frequency averages 72.48% when information is *not* purchased. Finally, we note that despite the lower-than-predicted frequency of information purchase by investors, trustees are nevertheless playing Return with high frequency in the Cost treatment. In fact we cannot reject the null hypothesis that the frequency of return given investment is not different from the efficient equilibrium prediction of 94.3% (t-test $p>.10$).

In Table 7 we take a closer look at which subjects bought information and how often they bought information. Table 7 shows the frequency of information purchase for each of the six subjects when that subject was assigned the role of investor in the Cost treatment. The frequencies over 0.5 are shown in bold face.

Notice that 9/36=1/4 of subjects purchased information more than 50% of the time (9 entries are in boldface). Approximately 1/2 of subjects (17/36) never chose to buy information. One subject was never assigned as an investor in the Cost treatment. The remaining 1/4 of subjects bought information at least once, but less than 50% of the time.

Session	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6
No_Info_Cost1	0/14	0/9	0/7	**21/21**	0/18	0/24
No_Info_Cost2	4/27	**2/2**	5/24	2/6	**19/26**	0/2
No_Info_Cost3	0/8	2/20	**11/11**	**3/5**	**2/2**	0/23
Info_No_Cost1	0/23	N/A	0/4	0/23	1/9	0/16
Info_No_Cost2	0/19	0/14	0/19	4/12	**1/1**	**15/16**
Info_No_Cost3	1/7	5/20	1/3	0/7	**20/20**	0/24

Table 7: The Frequency of Information Purchase for Each Subject

We summarize these results as follows:

Finding 12: *In the Cost treatment, there is heterogeneity in the frequency with which subjects purchase information, with some investors never purchasing information, while others frequently purchasing information. On average, the frequency with which investors purchase information is less than half of the predicted level in the most efficient equilibrium. Nevertheless, the frequency with which uninformed investors play Invest and the frequency with which trustees play Return-given-Invest are both high, and close to the efficient equilibrium predictions.*

5.6 Investor Strategies

In this section we explore the types of *strategies* that were adopted by investors in making investment decisions in the three sets of experiments. Our method is to use a random-effects Probit regression model on pooled data from all sessions of a treatment, where the standard errors have been adjusted to allow for clustering of observations by session number.[21] In all specifications of this model, we always include several variables that capture fixed effects (sequence, period, order). Here, "sequence" refers to the supergame number and "period'' to the period number in a sequence. "Order" is a dummy variable that equals 1 if the session begins with minimal or full information and equals 0 otherwise. Following Camera and Casari (2009), we also introduce a set of regressors (dummy variables) that are used to trace out investor's strategies. In particular, we include a "grim trigger" dummy that equals 1 in all periods of a supergame following the first period of that supergame in which the investor experienced a defection by a trustee and equals 0 otherwise. We also include two "tit-for-tat" dummies that equal 1 in the first and second period respectively following a defection and equal 0 otherwise. Third, we include dummies that represent the information shown to the investor regarding to the current paired trustee's past behavior in the current sequence (Last return, All return, and All no return). "Last return" is a dummy variable that is equal to 1 if information reveals that the current paired trustee returned last period. "All return" has a value of 1 if information reveals that the trustee returned every time the investor invested, and "All no return" has a value of 1 if information reveals that the trustee never returned given investment. We also consider some interactive dummy variables that are the product of dummy variables, e.g., Grim and Last return. Finally, Buy and All return, Buy and All no return, Buy and Last return are dummy variables that are relevant to the Cost treatment only and are conditional on the investor having bought information.

Table 8: Random-Effect Probit Regression on Individual Choice to Invest
(For No_Min and Min_No and Finite Horizon Sessions)

Dependent variable: 1 = Invest, 0 = No Invest	No Information (Baseline Treatment)	No Information Finite Horizon	Minimal Information
Constant	.994 (.217)***	1.18 (.316)***	-.548 (.541)
Period	-.020 (.012)*	-.144 (.078)*	-.009 (.018)
Sequence	.021 (.041)	.007 (.0359)	.038 (.030)
Order	-.366 (.347)		-.289 (.318)
Grim trigger	-1.188 (.335)***	-.978 (.287)***	-.765 (.391)**
Tit-for-tat with lag 1	-.050 (.225)	-.329 (.293)	-.305 (.246)
Tit-for-tat with lag 2	-.130 (0.233)	-.269 (.293)	-.229 (.191)
Last return			2.254 (.207)***
Grim and Last return			-1.102 (.498)**
Observations	720	384	747

The Probit Regression results in Table 8 are for the No_Min, Min_No and No-Finite experimental

[21] For this estimation we used the gllamm package in Stata version 11.

sessions and reveal several important findings. First, a case might be made that subjects in the No treatment were playing in accord with the predictions of the contagious strategy, as evidenced by the significance of the negative coefficient associated with the grim trigger dummy and insignificance of all other right hand side variables (other than the constant term). However, we see that the grim trigger dummy is also significant in the No-finite treatment, which suggests that subjects who employed such a strategy are not conditioning on whether the horizon is indefinite or finite. By contrast, in the Min treatment, subjects were playing according to a mixture of motives, as indicated by the positive and significant coefficient on the last return dummy variable and the negative and significant coefficients on the grim and grim-and-last-return dummy variables.

Similar findings emerge from the probit regressions conducted using the third set of experimental sessions, the No_Info_Cost and Info_No_Cost data as revealed in Table 9. In particular, we see that investors in the No treatment continued playing a grim-type strategy, with tit-for-tat also showing up significantly. Investors in the Info and Cost treatments were more likely to choose Invest if the trustees' history they had bought showed they had always returned or returned in the last period and were less likely to choose Invest otherwise. By contrast with the Min treatment, in the Info treatment, we observe that the coefficient on the grim-and-last-return interaction dummy is positive; when more information is available, it appears that investors rely less on the global, grim punishment strategy in favor of a more local, individual information-contingent strategy.

Table 9: Random-Effect Probit Regression on Individual Choice to Invest
(For No_Info_Cost and Info_No_Cost Sessions)

Dependent variable: 1 = Invest, 0 = No Invest	No Information (Baseline Treatment)	Full Information	Costly Information
Constant	.761 (.477)	.588 (.853)	3.903 (2.887)
Period	-.002 (.054)	-.011 (.007)	-.102 (.043)**
Sequence	.030 (.038)	.036 (.069)	-.075 (.129)
Order	-.465 (.370)	-.307 (.515)	-1.338 (.561)**
Grim trigger	-1.227 (.382)***	-1.549 (.210)***	-1.120 (.670)*
Tit-for-tat with lag 1	-.973 (.111)***	-1.088 (.622)*	-.467 (.267)*
Tit-for-tat with lag 2	-.419 (.289)	-.515 (.488)	.199 (.463)
All return		.468 (.229)**	
All no return		-1.350 (.449)***	
Last return		.935 (.409)**	
Grim and Last return		1.127 (.337)***	
Buy and All return			1.325 (.605)**
Buy and All no return			-1.341 (1.035)
Buy and Last return			.838 (.428)*
Grim and Buy and Last return			.119 (.305)
Observations	552	528	486

We summarize these results as:

Finding 13: *When investors lack information on trustees they behave in a manner that resembles the contagious strategy, investing less frequently when their history has included a defection (play of Keep by the trustee) than when it has not. When investors have information on trustees, they condition their behavior on this information. With longer histories of information, the importance of the contagious-grim strategy in sustaining trust is diminished in favor of a conditionally trustworthy strategy where investors choose Invest if the trustee's history indicates the trustee is likely to return.*

6. Conclusion

We have studied the development of a social norm of trust and reciprocity among strangers in an indefinitely repeated trust game with random matching. By contrast with the Prisoner's Dilemma game, the trust game captures the one-sided incentive problem that arises in many everyday economic transactions. Indeed, the main focus of most reputation systems is on the behavior of trustees (second movers). The baseline treatment involves no information on trustees' (second mover) behavior but the environment is one that can support complete trust and reciprocity as an equilibrium of the indefinitely repeated game. Other treatments involve the amount of information provided on a trustee's prior history of play. A further treatment concerns whether information provision is free or costly. We also examine the role of the indefiniteness of horizon in supporting the contagious equilibrium. We have provided a simple theoretical model examining these various informational assumptions and have examined the predictions that follow from our model in a controlled laboratory experiment.

Although the parameters of the game were chosen to support a social norm of complete (100%) trust and reciprocity as an equilibrium of the baseline indefinitely repeated game with no information about trustees, we find that the frequency of trust and reciprocity averages only about 33% in this case. A similar finding obtains in the finitely repeated games under no information about trustees. Providing reputational information about trustees leads to an increase in the frequencies of trust and reciprocity over the baseline No information case; the difference in the frequency of trust and reciprocity becomes significantly greater as the length of the history of the prior behavior of trustees is increased. Most importantly, providing the possibility to purchase information at a small cost also significantly increases the frequencies of trust and reciprocity compared with the baseline case where such information is not available. The latter result suggests that it may be the *availability* of information that matters even more than the actual content of that information. After all, only 24% of investors in our Cost treatment purchased information, and most who did not purchase information chose Invest anyway.

Our findings help us to comprehend the emergence and prevalence of reputation systems in economic interactions involving “strangers” and one-sided incentive problems such as the online feedback system found on eBay or credit reports provided by credit bureaus. The significant contribution of this paper is that we have identified the importance of individual, reputational information over community-wide enforcement in increasing and sustaining trust and reciprocity and improving efficiency. This understanding is of obvious importance to the design and operation of economic institutions.

For future research, we are interested in designing and comparing different reputational mechanisms that improve the efficiency of the markets with random and anonymous players. For instance, one comparison is between an online, peer-to-peer feedback system (where information need not be truthful) and a third party credit bureau (where information may be presumed to be more truthful). Online feedback systems involve decentralized voluntary contribution of information about transactions with free dissemination to other users. On the contrary, the traditional credit bureau collects information in a centralized and mandatory method, as in our Min and Info treatments and charges users for access to this information. Although we have shown in this paper that providing reputational information significantly increases efficiency, we still have not solved the more practical issue of how to decentralize or finance the reputational system. We leave such an analysis to future research.

References

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

Bohnet, I., Harmgart, H., Huck, S., Tyran, J-R.: Learning trust. J Eur Econ Assoc **3**, 322-329 (2005)

Bolton, G.,Katok, E., Ockenfels, A.: How effective are electronic reputation mechanisms? An experimental investigation. Manag Sci **50**, 1587-1602 (2004)

Bolton, G., Katok, E., Ockenfels, A.: Cooperation among strangers with limited information about reputation. J Public Econ **89**, 1457-1468 (2005)

Brown, M., Zehnder, C.: Credit reporting, relationship banking, and loan repayment, J Money, Credit, Banking **39**, 1883-1918 (2007)

Camerer, C.F.: Behavioral Game Theory: Experiments in Strategic Interaction. Princeton University Press, Princeton (2003)

Camera, G., Casari, M.: Cooperation among strangers under the shadow of the future. Amer Econ Rev **99**, 979-1005 (2009)

Charness, G., Du, N., Yang, C-L.: Trust and trustworthiness reputations in an investment game. Games Econ Behav **72**, 361-375 (2011)

Dal Bo, P.: Social norms, cooperation and inequality. Econ Theory **30**, 89-105 (2007)

Duffy, J., Ochs, J.: Cooperative behavior and the frequency of social interaction. Games Econ Behav **66**, 785–812 (2009)

Ellison, G.: Cooperation in the prisoner's dilemma with anonymous random matching. Rev Econ Stud **61**, 567-588 (1994)

Engle-Warnick, J., Slonim, R.L.: Learning to trust in indefinitely repeated games. Games Econ Behav **54**, 95-114 (2006a)

Engle-Warnick, J., Slonim, R.L.: Inferring repeated-game strategies from actions: evidence from trust game experiments. Econ Theory **28**, 603-632 (2006b)

Fehr, E., Schmidt, K.: A theory of fairness, competition and cooperation. Quart J Econ **114**, 817-868 (1999)

Fischbacher, U.: z-Tree: Zurich toolbox for ready-made economic experiments. Exper Econ **10**, 171-178 (2007)

Greif, A.: Reputation and coalitions in medieval trade. J Econ Hist **49**, 857-882 (1989)

Greif, A.: Contract enforceability and economic institutions in early trade: The Maghribi traders' coalition, Amer Econ Rev **83**, 525-548 (1993)

Keser, C.: Experimental games for the design of reputation management systems. IBM Syst J **42**, 498-506 (2003)

Kandori, M.: Social norms and community enforcement. Rev Econ Stud **59**, 63-80 (1992)

Xie, H., Lee, Y.-J.: Social norms and trust among strangers. Working paper, Concordia University (2011)

Milgrom, P., North, D., Weingast, B.: The role of institutions in the revival of trade: The law merchant, private judges, and the Champagne fairs. Econ Politics **2**, 1-23 (1990)

Okuno-Fujiwara, M., Postlewaite, A.: Social norms and random matching games. Games Econ Behav **9**, 79-109 (1995)

Young, H.P.: Social norms. In: Durlauf, S.N., Blume, L.E. (eds.) The New Palgrave Dictionary of Economics, Second Ed., Palgrave Macmillan, New York (2008)

Appendix A: Definition of $f(\delta)$ and $g(\delta)$ and Proof of Propositions 2 and 3

Definition of $f(\delta)$ and $g(\delta)$: To provide a formal definition of $f(\delta)$ and $g(\delta)$, further notation is necessary. Let X_t be the total number of d-type investors and let Y_t be the total number of d-type trustees at the beginning of period t. The state of the world in period t, Z_t, contains information about the number of d-type investors and d-type trustees in the current period and is defined as a one-to-one and onto function from (X_t, Y_t) to the set of natural numbers $\{1, 2, \ldots, n(n+2)\}$:

$$Z_t = (n+1)X_t + Y_t \quad for \quad X_t + Y_t > 0.$$

Let A be an $n(n+2)\times n(n+2)$ transition matrix when all players follow the contagious strategy. It has elements

$$a_{ij} = \Pr\{Z_{t+1} = j \mid Z_t = i \text{ and all players follow the contagious strategy}\}.$$

For example, $a_{12} = \Pr\{Z_{t+1} = 2 \mid Z_t = 1\} = \Pr\{(X_{t+1}, Y_{t+1}) = (0,2) \mid (X_t, Y_t) = (0,1)\}$ denotes the probability that there are two d-type trustees and no d-type investors in next period given that there is one d-type trustee and no d-type investors in the current period. Similarly, let B be an $n(n+2)\times n(n+2)$ transition matrix when the d-type trustee in consideration deviates from the contagious strategy while all other players continue to follow the contagious strategy, with elements

$$b_{ij} = \Pr\{Z_{t+1} = j \mid Z_t = i \text{, one d-type trustee deviates from the contagious strategy and all other players follow the contagious strategy}\}.$$

Thus, matrix $B - A$ characterizes how the diffusion of d-type players is delayed if one d-type trustee unilaterally deviates from the contagious strategy. Define ρ as an $n(n+1)\times 1$ column vector with ith element equal to the conditional probability that the d-type trustee meets a c-type investor when the state is i in period t. Finally, let e_i be a $1\times n(n+2)$ row vector with the ith element equal to 1 and all other elements equal to 0.

Define $f(\delta) \equiv \delta e_1 (B - A)(I - \delta A)^{-1} \rho$ and $g(\delta) \equiv \delta e_{n+2} (B - A)(I - \delta A)^{-1} \rho$.

$f(\delta)$ is the increase in the sum of the expected probability (and also the expected payoff) of meeting a c-type investor for the d-type trustee in all the future periods when this d-type trustee chooses to deviate from defection (i.e. continues to return) given that the d-type trustee is the *only* d-type player in the community. Similarly, $g(\delta)$ is the increase in the sum of the expected payoff (and the probability of meeting a c-type investor) in future periods for the d-type trustee from slowing down the contagious process given that currently there is just one d-type trustee and one d-type investor.[22]

Proof of Proposition 2: Consider the following strategy: the trustee chooses Keep if he is a d-type and chooses Return if he is a c-type; the investor chooses Invest in the first period and in subsequent periods if the information reveals the current trustee to be a c-type and chooses No Invest if the information reveals the current trustee to be a d-type.

We show next that a one-shot deviation is not profitable for the investor or the trustee both on the equilibrium path and off the equilibrium path.

For the trustee on the equilibrium path, the payoff from following the strategy above is $\frac{1-b}{1-\delta}$, while the payoff from a deviation is 1, with no payoff in all future periods. So the trustee has no incentive to deviate if $\delta \geq b$.

For the trustee off the equilibrium path, the payoff from following the strategy is 1, the payoff while the payoff from a deviation is $1-b$, so the trustee has no incentive to deviate.

For the investor on the equilibrium path (when the information reveals the current trustee to be a c-type), the payoff from following the strategy is $b + E\pi_1$, and the payoff from deviation is $a + E\pi_2$, where $E\pi_1$ and $E\pi_2$ are the future payoffs given that the investor chooses Invest and No Invest in the current period respectively. If the investor chooses No Invest instead of Invest this period, the current matched trustee will become a d-type instead of remaining a c-type player, so the investor's future payoff given No Invest in the current period, $E\pi_2$, is smaller than the future payoff given Invest in the current period, $E\pi_1$. Therefore the investor has no incentive to deviate on the equilibrium path.

For the investor off the equilibrium path (when the information reveals the current trustee to be a d-type), the payoff from following the strategy is $a + E\pi_3$, and the payoff from deviation is $E\pi_4$, where $E\pi_3$ and $E\pi_4$ are the investor's future payoffs given No Invest and Invest in the current period respectively. Since the current trustee is a d-type, the investor's choice in the current period will not affect the number of d-type trustees in the future. Thus, $E\pi_3$ equals $E\pi_4$ and the investor has no incentive to deviate.

Proof of Corollary 1: We will show that the following strategy under minimal information is an equilibrium strategy, but only for the knife-edge condition where $\delta = b$. Specifically, the trustee chooses Keep if he is a d-type and chooses Return if he is a c-type; the investor chooses Invest in the first period or if the minimal information reveals the current trustee returned last period (is a c-type) and chooses No Invest otherwise.

[22] The formal derivation of $f(\delta)$ and $g(\delta)$, as well as the formula for each element of matrix *A* and *B* can be found in Xie and Lee (2011).

It is easy to verify that the investor will follow the proposed strategy (similar to the proof of Proposition 1). In particular, consider the incentive constraint for a d-type trustee given investment by the investor in the current period. The d-type trustee's payoff from choosing Keep in the current period is 1. However, by choosing Return in the current period and waiting to play Keep in the next period, his payoff is $1 - b + \delta$ since the investor in the next period will choose Invest given the trustee's record of Return in the current period. Therefore a d-type trustee will not deviate from playing Keep only if $b \geq \delta$.

Next, consider the incentive constraint for a c-type trustee given investment by the investor in the current period. His payoff from choosing Return is $\frac{1-b}{1-\delta}$ and his payoff from choosing Keep is 1. So he will not deviate from playing Return only if $\delta \geq b$.

Therefore under minimal information, the proposed strategy is an equilibrium strategy only if $\delta = b$. Intuitively, in the minimal information case, once a trustee gets investment for the current period, for the future history, the past history is irrelevant and hence trustees carrying different minimum information will face the same incentives. Different types might behave differently only in the knife-edge case where they are both indifferent between remaining true to their type and switching to the other type.

Proof of Proposition 3:

Consider the following strategies: In the first period, the investor always chooses Invest. In subsequent periods, the investor purchases information with probability q. If he does not purchase information, he chooses Invest with probability γ, where $0 \leq \gamma \leq 1$, otherwise if the information purchased reveals the trustee has always played Return, the investor plays Invest and otherwise he plays No Invest. Trustees adopt an asymmetric strategy, with a fraction p of trustees choosing Return given investment in every period, and fraction $1 - p$ of trustees choosing Keep given investment in every period.

The trustee's payoff from choosing Return every period is $(1 - b) + \frac{\delta}{1-\delta}[q + (1 - q)\gamma](1 - b)$.

The trustee's payoff from choosing Keep every period is $1 + \frac{\delta}{1-\delta}(1 - q)\gamma$.

These two payoffs must be equal in the first period, so

$$(1 - b) + \frac{\delta}{1-\delta}[q + (1 - q)\gamma](1 - b) = 1 + \frac{\delta}{1-\delta}(1 - q)\gamma,$$

which implies $q = \frac{b - \delta b + \gamma\delta b}{\delta - \delta b + \gamma\delta b}$.

$$\frac{\partial q}{\partial \gamma} = \frac{(\delta - b)\delta b}{(\delta - \delta b + \gamma\delta b)^2}$$

Notice that q is increasing in γ if $\delta > b$. Thus, if the investor chooses Invest with a higher probability when no information is available, in equilibrium the probability for the investor to purchase information must be higher. The intuition for this finding is that if the investor increases the probability of invest under no information, she must at the same time decrease the probability of no information, otherwise all trustees will switch to always choosing Keep in every period.

In order to show that this is an equilibrium strategy, we also need to show that, given investment, the trustee will not switch from playing Return to playing Keep or from playing Keep to playing Return in every period, i.e., that p is constant. This is easy to verify since once the trustee chooses Keep in one period, then the best he can do (under full information) is to continue to choose Keep in every period. Therefore, a trustee who chooses Keep in the first period has no incentive to switch to the other strategy in later periods. For a trustee who chooses Return in the first period, in every period he is facing the same constraint he faced in the first period, so in every period he will be indifferent between always choosing Return and always choosing Keep given investment.

Next we show that investors have no incentive to deviate from the proposed strategy. Consider first the investors who do not purchase information. These investors will follow a mixed strategy only if their payoff from Invest is equal to their payoff from No Invest, i.e., if $a = pb$. Since investors also randomize as to whether or not they will purchase information, the following equation must hold,

$$a = pb = pb + (1-p)a - c,$$

where the last term of the equation is investor's payoff from purchasing information and c is the cost of purchasing information.

Therefore, in equilibrium the fraction of good trustees is $p = \frac{a}{b}$, and the cost of information is $c = \frac{a}{b}(b-a)$.

In particular, we are interested in two special cases: Investors who do not purchase information either choose Invest with probability zero (i.e., choose No Invest) or they choose Invest with probability one.

Case 1: If the investors choose No Invest with probability one when no information is purchased ($\gamma = 0$), it must hold that $a = pb + (1-p)a - c \geq pb$. Given the cost c chosen in the experiment, $a > pb$ and therefore the investors will choose No Invest in the first period. So in all the subsequent periods, investors should choose not to purchase information and No Invest with probability one. Therefore, the equilibrium reduces to the most inefficient equilibrium.

Case 2: If the investors choose Invest with probability one when no information is purchased ($\gamma = 1$), the investors purchase information with probability $q = \frac{b}{\delta}$ and it must hold for that $pb = pb + (1-p)a - c \geq a$. So the fraction of trustees who always choose Return given investment is $p = \frac{a-c}{a}$, where $c \leq \frac{a}{b}(b-a)$. From the parameters chosen in the experiment, $q = 9/16$ and $p = 33/35$.

Next we turn to the efficiency comparison.

(1) When $0 < \gamma < 1$, according to the equilibrium outcome, the investor's payoff in each period is a. The trustee's average payoff in the first period is $p(1-b) + (1-p) = 1-a$ given that $p = a/b$, and the trustee's average payoff in each subsequent period (denoted by AU) is $AU = p[q + 1-q\gamma 1-b+1-p[(1-q)\gamma]$, where $q+1-q\gamma 1-b$ is the expected payoff for a trustee who always chooses Return and $(1-q)\gamma$ is the expected payoff for a trustee who always chooses Keep. In the equilibrium, $q = \frac{b-\delta b+\gamma\delta b}{\delta-\delta b+\gamma\delta b}$ is increasing in γ. It is easy to verify that $(1-q)\gamma$ is also increasing in γ. Therefore, the most efficient outcome obtains when $\gamma \to 1$ and $AU \to 0.52$ given $p = \frac{a}{b} = 7/9$.

(2) When $\gamma = 1$, according to the equilibrium outcome in Case 2, the investor's payoff in each period is $pb > a$ given $c = 2$ as we chose in the experiment. The trustee's average payoff in the first period is $p(1-b) + (1-p)$, and the trustee's average payoff in each subsequent period $AU = p[q + 1-q\gamma 1-b+1-p1-q\gamma=0.54>0.52$ given $p=33/35$. Therefore the most efficient equilibrium outcome obtains when $\gamma = 1$.

Proof of Corollary 2: As noted in the text, when information is costly, there also exists an inefficient equilibrium where no investors purchases information or chooses Invest and no trustee chooses Return. It is easy to show that deviations from this equilibrium either on or off the equilibrium path are not profitable for either the trustee or the investor.

Not for Publication

Appendix B: Instructions Used in the Experiment

In this appendix we provide the instructions given to subjects in the Min_No sessions. In italics, we show the additional instruction that was provided to subjects in the Info_No_Cost sessions.

Instructions

Overview

This is an experiment in decision-making. The department of economics has provided funds for this research. During the course of this experiment, you will be called upon to make a series of decisions. If you follow these instructions carefully and make good decisions, you can earn a considerable amount of money which will be paid to you in cash at the end of the experiment. We ask that you not talk with one another for the duration of the experiment.

Specifics

The experiment is divided into a series of sequences. A sequence will consist of an indefinite number of rounds. At the beginning of each sequence, you will be randomly assigned as a First Mover or a Second Mover. Your role will appear on your computer screen and will not change during the sequence. At the beginning of each round you will be randomly paired with another person who is assigned to the other role from your own. That is, if you are a First Mover (Second Mover), in each round you will be randomly paired with a Second Mover (First Mover) with all possible pairings being equally likely.

In each round, you and your paired player will play the game described in the following graph. First, the First Mover chooses between A and B. If the First Mover chooses A, the round is over. The First Mover receives 35 points and the Second Mover receives 0 points. If the First Mover chooses B, then the Second Mover must make a choice between C and D. If the Second Mover chooses C, then the First Mover receives 0 points and the Second Mover receives 100 points. If the Second Mover chooses D, then the First Mover receives 45 points and the Second Mover receives 55 points. Following the first round of a sequence, the First Mover will be told the decision that his/her paired Second Mover has made in the last round, if that player was able to choose between C and D. The First Mover never knows the identity of his/her paired Second Mover.

(*Following the first round of a sequence, at the beginning of each round the First Mover will be told the past choices that his/her paired Second Mover has made in each of the most recent rounds of the current sequence (up to 10 rounds), provided that this Second Mover had a choice to make (the First Mover chose B). The First Mover will be also told the total number of times that his/her paired Second Mover has chosen C or D in the entire sequence, again conditional on having the opportunity to make a choice (First Mover chose B). The First Mover never knows the identity of his/her paired Second Mover*.)

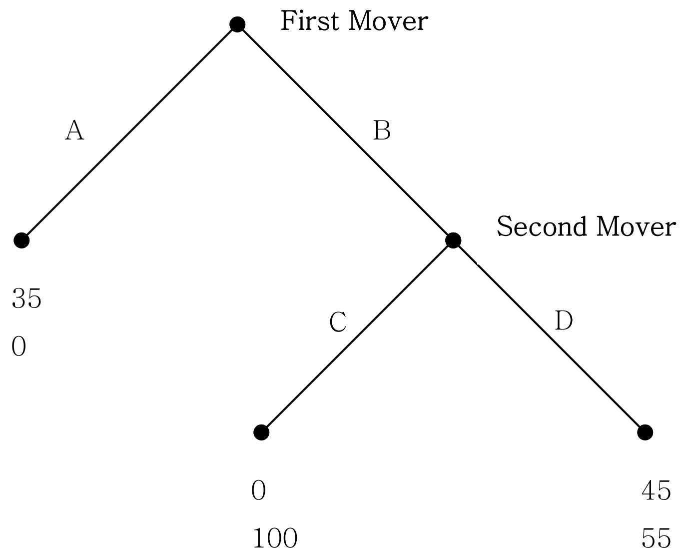

Figure A1: Decisions and Earnings (in Points)

To complete your choice in each round, click on the decision button and then the OK button. The Second Movers need to wait for the First Movers to make a choice between A and B before making their own choice. Then the Second Mover will be told that the round is over (if the First Mover chooses A), or will be asked to make a choice between C and D (if the First Mover chooses B).

The computer program will record your choice and the choice made by the player paired with you in this round. After all players have made their choices, the results of the round will appear on your screen. You will be reminded of your own choice and will be shown the choice of your match, as well as the payoff you have earned for the round. Record the results of the round on your RECORD SHEET under the appropriate headings.

Immediately after you have received information on your choice and the choice of the player with whom you are randomly paired for the round, a ten-sided die with numbers from 0 to 9 on the sides will be thrown by one of you (you will take turns throwing the die) to determine whether the sequence will continue or not. The experimenter will announce loudly the result of each die roll. If a number from 0 to 7 appears, the sequence will continue into next round. If an 8 or 9 appears, the sequence ends. Therefore, after each round there is 80% chance that you will play another round and 20% chance that the sequence will end.

Suppose that a number less than 8 has appeared. Then you will play the same game as in the previous round, but with an individual selected at random whose role is different from yours. You record the outcome and your earnings for the round. Then another throw is made with the same die to decide whether the sequence continues for another round.

If an 8 or 9 appears, the sequence ends. The experimenter will announce whether or not a new sequence will be played. If a new sequence is to be played then you will be randomly reassigned as a First Mover or a Second Mover. The new sequence will then be played as described above.

If the experiment does not end within 2 hours, you will be invited to continue the experiment in the next several days.

Earnings

Each point you earn is worth 0.5 cent ($0.005). Therefore, the more points you earn the more money you earn. You will be paid your earnings from all the rounds of all sequences in cash, and in private, at the end of today's experiment as well as $5 show-up fee.

Final Comments

First, do not discuss your decisions or your results with anyone at any time during the experiment.

Second, your ID# is private. Do not reveal it to anyone.

Third, since there is 80% chance that at the end of a round the sequence will continue, you can expect, on average, to play 5 rounds in a given sequence. However, since the stopping decision is made randomly, some sequences may be much longer than 5 rounds and others may be much shorter.

Fourth, your role as a First Mover or a Second Mover will be randomly assigned when a new sequence begins. Your role will not change for the duration of that sequence.

Finally, remember that after each round of a sequence you will be matched randomly with a player whose role is different from yours. Therefore, the probability of you being matched with the same individual in two consecutive rounds of a game is 1/3 since there are 3 First Movers and 3 Second Movers in the room.

Questions?

Now is the time for questions. Does anyone have any questions?

Quiz

If there are no more questions, please finish the quiz. Your answers to this quiz will not affect your earnings. The purpose of the quiz is to help you understand the instruction better. After everyone has completed the quiz the answers will be reviewed.

Continuation Instruction (Min_No sessions)

From now on until the end of today's experiment, everything is the same as in the original instruction except that there is no information available anymore to the First Mover about the decision that his/her paired Second Mover has made in the last round.

Continuation Instruction (Info_No_Cost sessions)

From now on everything is the same as in the original instructions except that there is no information available anymore to the First Mover about the past choices that his/her paired Second Mover has made in the most recent rounds of the current sequence, or about the frequency with which his/her paired Second Mover has chosen C or D in all rounds of the current sequence.

Continuation Instruction (Info_No_Cost sessions)

From now on until the end of today's session, at the beginning of each round of a sequence except the first round, the First Mover may choose to buy information about the past choices that his/her paired Second Mover has made in the current sequence, at a price of 2 points. The information provided is the same that was provided in the instructions for the first part of today's session: the First Mover is informed of the past choices that his/her paired Second Mover has made in the most recent rounds of the current sequence and about the frequency with which his/her paired Second Mover has chosen C or D in all rounds of the current sequence. The First Mover is free to decide whether or not to buy this information every round. If the First Mover chooses to buy the information, the 2 point cost will be deducted from his/her payoff from the round; otherwise there will be no deduction (and no information on the Second Mover will be provided to the First Mover). As in all previous rounds, the Second Mover will only be informed of his own payoff, and will not be able to observe whether or not the First Mover has chosen to purchase information about the Second Mover.

Not For Publication

Appendix C: Aggregate Frequencies in No_Info_Cost and Info_No_Cost Sessions (first period excluded)

		1st treatment	2nd treatment	3rd treatment	No. of Obs.
Frequency of Invest	No_Info_Cost	0.594	0.846	0.867	3
	Info_No_Cost	0.810	0.379	0.545	3
Frequency of Return-given-Invest	No_Info_Cost	0.722	0.928	0.887	3
	Info_No_Cost	0.925	0.541	0.783	3
Frequency of Invest-and-Return	No_Info_Cost	0.424	0.786	0.776	3
	Info_No_Cost	0.751	0.228	0.463	3

Table C1: Average Frequencies for No_Info_Cost and Info_No_Cost Sessions (First Period Excluded)