Does making specific investments unobservable boost investment incentives?*

Randolph Sloof,[†] Hessel Oosterbeek and Joep Sonnemans

September 12, 2005

Abstract

Theory predicts that holdup can be alleviated by making specific investments unobservable to the other party. Private information creates an informational rent that boosts investment incentives. Empirical findings, however, indicate that holdup is attenuated by fairness and reciprocity motivations. Private information may interfere with this, as it becomes impossible to directly observe whether the investor behaved fair or not. In that way unobservability could crowd out an informal fairness/reciprocity mechanism in place. This paper reports on an experiment to investigate this issue empirically. Our results are in line with standard predictions when there is limited scope for social preferences. But with sufficient scope for these motivational factors, unobservability does not boost specific investments.

1 Introduction

In a world with incomplete contracts, relationship-specific investments are vulnerable to appropriation by the trading partner. After investments costs are sunk, the other party may force a renegotiation to obtain a larger share of the ex post surplus. In that way this party may be able to obtain some of the returns on investment without sharing in the costs. Anticipating this,

*An earlier version of this paper was presented at conferences at Harvard University (ESA), in Erfuhrt (ESA) and in Madrid (EEA), and at seminars at NYU, the University of Bonn and the University of Rotterdam.

[†]Corresponding author; e-mail: r.sloof@uva.nl. All authors are affiliated with the Department of Economics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, the Netherlands.

the investor will invest less than the efficient level. This is the well-known holdup underinvestment problem (cf. Klein et al. 1978, Williamson 1985).

In an early contribution Tirole (1986, Proposition 3) shows that investment unobservability may alleviate underinvestment. Gul (2001) obtains that private information about the specific investment made may even help solving holdup completely. The general idea underlying these theoretical results is quite intuitive. Private information about the size of the ex post surplus created by the investment yields the investor an informational rent in the renegotiation stage. This increases the marginal return on investment she obtains and thus boosts her ex ante investment incentives. Therefore, by giving the investor some informational advantage, one might be able to alleviate holdup. Indeed, one objective of Gul (2001, p. 344) is:

> "...to emphasize the role of allocation of information as a tool in dealing with the hold-up problem. Audits, disclosure rules or privacy rights could be used to optimize the allocation of rents and guarantee the desired level of investment. Controlling the flow of information may prove to be a worthy alternative to controlling bargaining power in designing optimal organizations."

Rogerson (1992), Lau (2002) and Gonzalez (2004) also predict that private information rents are an effective instrument to boost investment incentives.

Existing empirical evidence, however, leads one to doubt the effectiveness of investment unobservability as an instrument. Private information may crowd out an informal fairness/reciprocity mechanism that is frequently observed in practice. To illustrate, a large number of holdup experiments have found that in complete information settings subjects on average invest significantly more than predicted.[1] Investment is typically seen as fair behavior, which is therefore rewarded by the non-investor with a larger than predicted return. This in turn makes it worthwhile to invest more than standard theory predicts.[2] Existing models of social preferences that take fairness and/or

[1] See e.g. Berg et al. (1995), Ellingsen and Johannesson (2004a, 2004b), Fehr and List (2004), Gantner et al. (2001), Hackett (1993, 1994), Königstein (2000), Oosterbeek et al. (2003) and Sonnemans et al. (2001).

[2] The empirical relevance of this informal mechanism is confirmed by the field data study of Leuven et al. (2005). They conduct a survey among a representative sample of the Dutch labor force. Among other things, the data set contains information about participation in work-related training, who paid for this training, and the worker's reciprocal attitude. The firm-sponsored training rate of workers with a high sensitivity to reciprocity appears to be around 48%, while it only equals 33% for workers with a low sensitivity to reciprocity. The difference of 15 percentage points is highly significant. Employers thus seem to actively rely on reciprocity as informal remedy against holdup.

reciprocity motivations into account, like the ones of e.g. Charness and Rabin (2002) and Dufwenberg and Kirchsteiger (2004), can explain these findings in a systematic way.

When in reality holdup is less severe, there is less need for explicit solutions like private information rents. More importantly, a serious concern is that private information may interfere with the informal fairness/reciprocity mechanism in place. This holds because parties then cannot directly observe whether the investor has behaved fair or not. Some experimental evidence that such interference indeed may happen is provided by Hackett (1994). He considers a setting in which higher investments increase the probability that the surplus up for renegotiation is large. Because in his setup this surplus is always publicly observable and thus informational rents are absent, theory predicts that investment levels are independent of whether they are observable or not. However, Hackett finds that investments levels are somewhat higher when they are observable. These results are suggestive, but in itself inconclusive about whether informational rents are (in)effective in mitigating holdup.[3] For that we need to consider a situation in which informational rents are indeed present at the renegotiation stage, i.e. one where the ex post surplus is private information to the investor. In this paper we do so, in order to establish whether investment unobservability really boosts investment incentives such as standard theory predicts.

Our experiment concerns a two-stage game between a buyer and a seller. In the first stage the buyer can make an investment that raises her valuation of the good. In the second stage the seller unilaterally determines the trading price. In one information condition the seller observes the buyer's investment decision, in another one he does not. In the latter case the seller does not know the buyer's valuation when setting his price, i.e. he does not observe the ex post surplus. Trade takes place only when the seller's price does not exceed the buyer's valuation. Otherwise the seller gets nothing and the buyer bears the cost of investment (if applicable). Within these two information conditions, (observable) investment costs can take three values: low, intermediate or high. Standard theory predicts that, compared to the observable investment case, unobservability increases investment levels by the same positive amount independent of the costs of investment. The informational rent

[3]This also applies to the experimental results reported in Oosterbeek et al. (1999) and Ellingsen and Johannesson (2005). Both papers consider situations in which the ex post surplus is always public information. The level of investment that lead to this observable surplus (i.e. the actual investment costs borne) may either be public or private information though. Because theoretically informational rents are absent at the bargaining stage, standard theory predicts that investment levels are independent of whether the actual investment costs borne are observable or not.

the buyer obtains always boosts her investment incentives. In contrast, models based on social preferences predict that with lower investment costs the effect of unobservability is smaller and may even disappear. The intuition is that lower investment costs increase the scope for motivational factors like fairness and reciprocity. And when this scope is large, investment levels are predicted to be independent of whether the investment is observable or not. If, however, this scope is small, the predictions of standard theory pertain.

Our results are in line with the social preferences predictions. When investment costs are high such that there is little scope for fairness and reciprocity, mean investments are substantially higher in the unobservable investment case than in the observable investment situation. Moreover, these mean levels are almost identical to the ones predicted by standard theory. With intermediate investment costs, mean investment levels in the two information conditions become more equal. And with low investment costs – so that there is much scope for social preferences – mean investment levels in the two conditions are the same.

The remainder of this paper is organized as follows. In the next section we present the simple game on which our experiment is based. This section also presents the standard equilibrium predictions, as well as alternative predictions based on two types of social preferences: outcome-oriented social preferences and intention-based reciprocity. Section 3 provides the details of the experimental design. Results are presented and discussed in Section 4. The final section concludes.

2 Theory

2.1 Basic setup of the model

Consider a bilateral relationship between a female buyer and a male seller who may trade one unit of a particular good. Both parties are assumed to be risk neutral. The order of play is as follows:

1. The buyer decides whether to make a specific investment ($I = 1$) or not ($I = 0$). Investment costs equal C and are immediately borne by the buyer. Without investment her valuation of the seller's good equals V, with investment this becomes $V + W$.

2. The seller makes a price demand $P \in [0, V + W]$ for which he is willing to sell the good. In case his price is weakly below the buyer's actual valuation, trade takes place at the demanded price. Otherwise, trade does not take place.

The second stage captures in reduced form a situation in which the buyer has (almost) no bargaining power at all. She is forced to accept the seller's price demand as long as this demand is weakly below her valuation. If the demanded price is above her valuation, her only option is to reject. Clearly this is not particularly realistic. In practice parties can at least indicate whether they accept or reject the terms of trade. We focus on the reduced form specification though, because it makes the equilibrium analysis under social preferences much simpler and the interpretation of the results more clear cut. In the concluding section we will return to this issue.

The seller's valuation is unaffected by the investment and normalized to zero. We assume that $0 < C < W$. This implies that making the investment is efficient. We also assume that $V > 0$, such that trade is always efficient. Maximum net overall surplus equals $V + W - C$.

Two different information conditions are considered. First, in the *observable investment* case the buyer's investment decision is publicly observable. Here the seller knows the buyer's valuation when he chooses his price demand P. Second, in the *unobservable investment* case the seller does not observe the buyer's investment choice. Then the seller does not know what the buyer's actual valuation is when he makes his price demand. This situation is formally equivalent to one in which the buyer and the seller simultaneously decide on I and P respectively. In both information conditions the setup of the game and the values of V, W and C are common knowledge.

2.2 Standard equilibrium predictions

Consider first the observable investment case. Solving the game through backward induction, the seller chooses $P_1^* = V + W$ after investment and $P_0^* = V$ after no investment. Here P_I^* denotes the seller's equilibrium price after observed investment decision $I \in \{0, 1\}$. Anticipating this pricing strategy, the buyer will not invest in order to save on the investment costs. Hence the unique subgame perfect equilibrium predicts holdup to be complete: $q_{obs}^* \equiv \Pr(I = 1) = 0$. There is no trade inefficiency, because the buyer and the seller always trade. Predicted net social surplus equals V. The efficiency loss owing to holdup is $W - C$.

In the unobservable investment case the seller cannot condition his price on the buyer's investment decision. Although he may demand any price in $[0, V+W]$, in equilibrium he will choose between $P = V$ and $P = V+W$ only. The reduced strategic form therefore corresponds to the 2×2 simultaneous-move game depicted in Table 1. This game has a unique mixed-strategy equilibrium: $q_{un}^* = \frac{V}{V+W}$ and $p^* \equiv \Pr(P = V) = \frac{C}{W}$.

Table 1: Reduced strategic form under unobservable investment

	$P = V$	$P = V + W$
$I = 0$	$0, V$	$0, 0$
$I = 1$	$W - C, V$	$-C, V + W$

Our interest lies in the effect of investment unobservability on the propensity to invest. The above analysis yields the following prediction:

Standard theory $q_{un} - q_{obs}$ is positive and independent of C; private information always boosts investment incentives.

Risk aversion does not affect this prediction. To see this, note that the buyer always chooses $q^*_{obs} = 0$, independent of her own or the seller's risk attitude. In the unobservable investment case the equilibrium is necessarily in mixed strategies. The buyer's mixing probability q^*_{un} should then make the seller indifferent, so it only depends on the risk attitude of the latter. The more risk averse the seller is, the larger q^*_{un} will be. If anything, risk aversion thus only strengthens the incentive effect of private information.

The predicted outcome in the observable investment case equals the upper-left cell in Table 1. Compared to this investment unobservability leads to more investment. This induces an efficiency gain of $q^*_{un} \cdot (W - C)$. At the same time it also introduces inefficient separations with probability $(1 - q^*_{un}) \cdot (1 - p^*)$. Inefficient separations occur when the seller demands a high price while the buyer did not invest. In that case the potential surplus of trade V is wasted. The expected gain owing to more investment and the expected loss due to inefficient separations cancel out; expected net social surplus under unobservable investment also equals V. Investment unobservability thus alters the source of the inefficiency, but it keeps the amount of inefficiency the same.[4]

Given the above conclusion, one remark is in order. In his paper Gul (2001, pp. 343-344) argues that "[N]either unobservable investment nor frequently repeated offers alleviate the holdup problem; yet, the two together completely resolve it." His solution to holdup thus contains two instruments and our experiment only considers the former. We do so for two reasons. First, our main interest lies in isolating the effect of informational rents on

[4] Lau (2002) notes that the latter result is an artifact of considering the two polar cases of no and full information asymmetry only. She shows that for intermediate degrees of information asymmetry the efficiency gains owing to better investment incentives outweigh the losses due to inefficient separations.

investment incentives. We therefore prefer to keep the bargaining stage as simple as possible. Second, Coasian dynamics (efficient trade) created by frequently repeated offers are already studied experimentally in other papers (see e.g. Cason and Sharma 2001).

The two information conditions reflect a tradeoff between efficient trade decisions and 'high-powered' investment incentives. This tradeoff may occur in various contexts. Riordan (1990) argues that vertical integration leads to a change in information structure; the downstream firm becomes better informed about upstream costs. This weakens the upstream firm's incentives to invest in cost reduction. The choice between vertical integration and market contracting is then between distorted investment incentives and distorted production decisions. Schmidt (1996) identifies a similar tradeoff between public and private ownership. Under nationalization government has precise information about a firm's costs and profits, but under privatization it has not. The costs of privatization are then a less efficient production level, while the benefits amount to better incentives for managers to save on production costs. Finally, Cremer (1995) argues that the choice of monitoring technology can be seen as a commitment device. In the context of our simple game, under unobservable investment the seller keeps the buyer at 'arm's length'. This enables him to commit to a single unconditional price P. Under observable investment such a commitment is non-credible and thus cannot be used to provide investment incentives. Without commitment the seller can always take the efficient trade decision though.

2.3 Predictions based on social preferences

The above standard predictions assume that agents solely care about their own material payoffs. By now there is a large body of empirical evidence, however, that a non-negligible fraction of people are motivated by fairness and reciprocity concerns (cf. Fehr and Gächter 2000). Models of social preferences that try to capture these alternative motivations can roughly be divided into two kinds (see Camerer 2003 and Fehr and Schmidt 2002 for comprehensive overviews). Under *outcome-oriented social preferences* agents only care about the final distribution of payoffs. The second type of social preferences theories assume that people also care about how the distribution of payoffs came about. In these models of *intention-based reciprocity* agents have a preference for rewarding kind (fair) intentions and punishing unkind (unfair) ones.

Subsections 2.3.1 and 2.3.2 below present detailed predictions of two representative models of outcome-oriented social preferences and intention-based reciprocity. The main conclusion that follows from the analysis is that when fairness/reciprocity considerations are sufficiently weak, the standard predictions prevail. However, when these alternative motivations are sufficiently strong, the buyer will always invest independent of the investment's observability. Private information then does not affect investment incentives. Now, the scope for fairness and reciprocity is larger when the costs of investment C are low relative to the return on investment W. Hence, when C decreases while keeping W constant, it becomes more likely that fairness and reciprocity considerations are 'sufficiently' strong. We thus obtain the following qualitative prediction based on social preferences:

Social preferences $q_{un} = q_{obs}$ when C is low and $q_{un} > q_{obs}$ when C is high; private information boosts investment incentives only when C is high.

These predictions also imply that, relative to the standard predictions, the increase in the investment rate owing to fairness and reciprocity motivations is larger when the investment is observable than when it is unobservable.
Two remarks remain. First, our formal analysis of social preferences assumes complete information. However, in practice subjects are heterogeneous and typically privately informed on their own preferences. Some care strongly about the payoffs of others whereas others are completely selfish. Hence even at a low cost level a fraction of subjects is likely to behave selfish. Likewise, even at a high cost level some subjects may reveal a concern for fairness and reciprocity. Yet we expect that when we aggregate over all subjects, the above qualitative predictions will pertain.

Second, our interest lies in the impact of private information on investment incentives. We do not intend to test particular models of social preferences per se. Numerous studies have already established the importance (and limits) of these motivational forces and a number of experiments have been purposely designed to discriminate between the various models, see e.g. Cox (2004), Falk et al. (2000) and McCabe et al. (2003). We did not conduct our experiment with that purpose in mind. The important thing we want to point out here is that when social preferences are effectively important (i.e. when C is low), private information is predicted to have less or even no impact on investment behavior, independent of exactly how these motivations are modeled.

2.3.1 Outcome-oriented social preferences

Following Charness and Rabin (2002) we assume that the utility of player $i = B, S$ equals the following weighted average of monetary payoffs π_i and

π_j $(i \neq j)$:[5]

$$u_i(\pi_i, \pi_j) \equiv \pi_i + \rho_i \cdot (\pi_j - \pi_i) \equiv (1 - \rho_i) \cdot \pi_i + \rho_i \cdot \pi_j \text{ when } \pi_i \geq \pi_j \qquad (1)$$
$$u_i(\pi_i, \pi_j) \equiv \pi_i + \sigma_i \cdot (\pi_j - \pi_i) \equiv (1 - \sigma_i) \cdot \pi_i + \sigma_i \cdot \pi_j \text{ when } \pi_i \leq \pi_j$$

Here ρ_i and σ_i are parameters reflecting the marginal rate of substitution between own material payoffs and the payoffs of the other. Note that the relative weight player i attaches to j's payoffs depends on player i being ahead (ρ_i) or behind (σ_i). We assume that $\sigma_i < \frac{1}{2}$. This implies that a player is more concerned about her own payoffs than about those of the other party when she is behind. We also assume that $\sigma_i \leq \rho_i$ and that $\rho_i < 1$. The latter entails that even when player i is ahead, she prefers more money to less, other things equal.

The above distributional preferences nest some particular types as special cases. For instance, competitive preferences arise when $\sigma_i \leq \rho_i < 0$. The inequality-aversion model of Fehr and Schmidt (1999) is obtained when we assume that $\sigma_i < 0 < \rho_i < 1$. Lastly, Charness and Rabin's quasi-maximin preferences require that $1 \geq \rho_i \geq \sigma_i > 0$ (and $\sigma_i < \frac{1}{2}$). This implies that players attach a positive weight to the other party's monetary payoffs, even when these payoffs exceed their own.

In both information conditions there exists a unique subgame perfect equilibrium. The main predictions about investment levels can be stated as follows.

Theorem 1 Suppose preferences are given by (1) and players know each others preferences. In the unique subgame perfect equilibrium investment behavior is characterized by:[6]

(weak) if $\rho_S < \frac{1}{2}$ and $\sigma_B < \frac{C}{C+V+W}$, then $1 > q^*_{un} > q^*_{obs} = 0$;

(medium) if $\rho_S < \frac{1}{2}$ and $\frac{C}{C+V+W} < \sigma_B < \frac{C}{C+W}$, then $1 = q^*_{un} > q^*_{obs} = 0$;

(strong) if $\rho_S > \frac{1}{2}$ or $\sigma_B > \frac{C}{C+W}$, then $q^*_{un} = q^*_{obs} = 1$.

Theorem 1 distinguishes three cases depending on the parameters ρ_i and σ_i. In the 'weak' case effectively neither player cares about the payoffs the

[5] In order to focus on distributional preferences we consider the simplified version of the Charness and Rabin (2002) model, ignoring the negative reciprocity element (i.e. shift parameter θ) of the full model. Intention-based reciprocity is formally addressed in the next subsection.

[6] We exclude the degenerate case in which $\rho_S = \frac{1}{2}$, i.e. the case the seller is indifferent between any division that weakly favors himself. Also knife-edge cases for σ_B are excluded. The proof appears in Appendix A.1.

other player gets. The 'strong' situation applies when either the seller or the buyer (or both) has a strong concern for the payoffs of the other party. The remaining 'medium' case applies when the seller does not really care about what the buyer gets, while the buyer puts some positive weight on the seller's payoffs.

The intuition behind Theorem 1 is as follows. When $\rho_S < \frac{1}{2}$ the seller's concern for the buyer's payoff is insufficient to justify own monetary sacrifice. He always asks for the complete pie under observable investment. Anticipating this, the buyer prefers to invest only if she cares sufficiently strong about the seller's payoff when the latter is ahead. Not investing yields her $\sigma_B V$, investing gives her $(1 - \sigma_B)(-C) + \sigma_B (V + W)$. We thus obtain that $\sigma_B > \frac{C}{C+W}$ is required for $q^*_{obs} = 1$.

Also under unobservable investment the seller with $\rho_S < \frac{1}{2}$ would like to claim the complete pie. Suppose that the seller is convinced that the pie is large. He then asks for $P = V + W$. In that case investing again yields the buyer $(1 - \sigma_B)(-C) + \sigma_B (V + W)$. But now no investment yields her a payoff of 0 rather than $\sigma_B V$, because the small pie is wasted. A buyer with $\sigma_B > 0$ cares about this inefficiency and is therefore more easily persuaded to invest for sure. Now only $\sigma_B > \frac{C}{C+V+W}$ is needed. Roughly put, under unobservable investment the buyer's threat to abstain from investment is less credible if she cares about the risk that the small pie is wasted. Finally, in case $\rho_S > \frac{1}{2}$ the seller's concern for the buyer is that high that he prefers to propose an equal split of the net surplus. Anticipating this, the buyer invests for sure.

From Theorem 1 it follows that the buyer always invests weakly more under unobservable investment. In case distributional preferences are weak the predicted investment rates are comparable to standard theory. But when either the buyer or the seller cares sufficiently about the payoffs of the other, the buyer invests for sure in both information conditions. In that case private information has no impact on investment incentives. Because the cutoff value for σ_B (i.e. the buyer's relative concern for the seller when she is behind) depends on the costs of investment, these predictions can be rephrased in terms of C. If C increases from a low level some buyers will move from the 'strong' to the 'medium' or 'weak' category. These buyers are now predicted to invest more when the investment is unobservable than when it is observable, creating a wedge between the average investment rates in the two information conditions. We thus obtain the qualitative prediction that private information boosts investment incentives only when C is high.

The above qualitative prediction relies on $\sigma_B > 0$ for at least some fraction of the buyers. For buyers with $\sigma_B < 0$ necessarily either the weak or the strong case of Theorem 1 applies and the case distinction is independent of C. In effect the prediction assumes that *at least some* buyers have a concern for efficiency, even when they earn less than the seller. This seems a plausible assumption. Charness and Rabin (2002) and Engelmann and Strobel (2004), for example, find that a large fraction of subjects have a concern for efficiency even when they are behind. More generally, these papers find that subjects' behavior can be well described by quasi-maximin preferences. For that type of preferences the qualitative prediction directly applies. Inequality-aversion as introduced by Fehr and Schmidt (1999) requires $\sigma_B \leq 0$. Hence, would *every subject* be driven by inequality-aversion, then Theorem 1 predicts no relationship between investment incentives and C. This is driven by our assumption of complete information about preferences though. In case subjects are privately informed on their own preferences – a key ingredient of Fehr and Schmidt's theory of inequality-aversion – equilibrium investments do vary negatively with C.[7] Overall then, our qualitative prediction is not inconsistent with a large fraction of inequality-averse subjects.

2.3.2 Intention-based reciprocity

The reciprocity concept of Dufwenberg and Kirchsteiger (2004) is not based on a payoff comparison between players. Building on Rabin (1993) they measure reciprocity with reference to the range of what one player could give the other player in principle. In their model utility functions take the following form:

$$\begin{aligned} u_B &= \pi_B + Y_B \cdot \kappa_{BS} \cdot \lambda_{BSB} \\ u_S &= \pi_S + Y_S \cdot \kappa_{SB} \cdot \lambda_{SBS} \end{aligned} \tag{2}$$

Again π_i denotes monetary payoffs of player i ($i = B, S$), while $Y_i \geq 0$ gives this player's reciprocal attitude. The higher Y_i, the more sensitive to reciprocity i is. The factor κ_{ij} represents i's kindness to j. It is positive if i is kind to j and negative if i is unkind to j. Term λ_{iji} gives i's belief about how kind j is to i. It is positive when i believes that j is kind to him, and negative when i thinks that j is unkind to him. Reciprocity is captured by the incentive to match the sign of κ_{ij} with the sign of λ_{iji}. Each player prefers being (un)kind towards the other party when the latter has been (un)kind towards him or her.

[7] To illustrate, let f be the fraction of players with $\rho_i > \frac{1}{2}$. It can be easily shown that in the observable investment case, the buyer invests whenever $\sigma_B > \frac{2C - f(W+C)}{2(1-f)(W+C)}$. The r.h.s. is an increasing function of C and negative for low values of C. So, when C increases the fraction of buyer types that are predicted to invest decreases, even when $\sigma_B < 0$.

Both factors κ_{ij} and λ_{iji} depend on player i's beliefs. Dufwenberg and Kirchsteiger (2004) provide exact definitions of how κ_{ij} and λ_{iji} are measured. Because utility now also depends on players' beliefs, psychological game theory has to be used. Within this framework Dufwenberg and Kirchsteiger define and prove the existence of a *sequential reciprocity equilibrium* (SRE). This concept requires each player to maximize his utility given correct beliefs, and also invokes a subgame perfection requirement.

The full equilibrium analysis is quite involved. However, it can be shown that for the buyer's equilibrium investment behavior the reciprocal attitude of the seller Y_S is decisive (cf. Appendix A.2). In particular, irrespective of the information condition that applies, $q^* = 1$ is possible if and only if Y_S is sufficiently high. For ease of exposition we therefore assume here that $Y_B = 0$. This implies that the buyer is not reciprocal at all.[8] In addition, we focus on the situation considered in the experiment in which the investment more than doubles the ex post surplus, i.e. $V < W$.

Theorem 2 Suppose preferences are given by (2) with $Y_B = 0$ and let $V < W$. In the observable investment case there exists a unique SRE. In the unobservable investment case the SRE is unique when $Y_S < \frac{2}{V+W-C}$ whereas for $Y_S > \frac{2}{V+W-C}$ multiple SRE exist. Equilibrium investment behavior is characterized by:

(weak) if $Y_S < \frac{2}{V+W-C}$, then $1 > \frac{V}{V+W} \geq q^*_{un} > q^*_{obs} = 0$;

(medium-low) if $\frac{2}{V+W-C} < Y_S < \frac{2}{W-C} \cdot \frac{W}{V+W}$, then $1 \geq q^*_{un} > q^*_{obs} = 0$. In the unobservable investment case there exists one SRE with $0 < q^*_{un} < \frac{V}{V+W}$ and another one with $q^*_{un} = 1$;[9]

(medium-high) if $\frac{2}{W-C} \cdot \frac{W}{V+W} < Y_S < \frac{2}{W-C}$, then $1 \geq q^*_{un} > q^*_{obs} = 0$. In the unobservable investment case there exists one SRE with $0 < q^*_{un} < \frac{V}{V+W}$, one with $\frac{1}{2} < q^*_{un} < 1$ and another one with $q^*_{un} = 1$;

(strong) if $Y_S > \frac{2}{W-C}$, then2 $1 = q^*_{obs} \geq q^*_{un} > 0$. In the unobservable investment case there exists one SRE with $0 < q^*_{un} < \frac{V}{V+W}$, one with $\frac{1}{2} < q^*_{un} < 1$ and another one with $q^*_{un} = 1$.

[8]Dufwenberg and Kirchsteiger (2000) make the same simplifying assumption in their analysis of employer-worker relationships. In Appendix A.2 we provide the complete equilibrium analysis for the more general case $Y_B \geq 0$ and show that this leads to qualitatively the same results as in Theorem 2.

[9]The SRE with $q^*_{un} = 1$ in the two 'medium' cases and in the 'strong' case in fact represents a continuum of equilibria, because $q^*_{un} = 1$ can be supported by a continuum of equilibrium pricing strategies (cf. Proposition 4 in Appendix A.2).

In the observable investment case the buyer invests for sure if the seller is sufficiently reciprocal ('strong' case). Otherwise she does not invest at all. If the investment is unobservable and Y_S is sufficiently high ('strong' and 'medium' cases), there exists a SRE in which the buyer always invests. But apart from that, for any value of Y_S there exists a SRE in which the buyer invests with low probability $0 < q^*_{un} \leq \frac{V}{V+W}$. Even a third SRE may exist side by side in which the probability of investment falls in between ($\frac{1}{2} < q^*_{un} < 1$).

The SRE with $q^*_{un} = 1$ and $q^*_{obs} = 1$ are based on positive reciprocity. Because investment is rewarded by the reciprocal seller with a return higher than predicted by standard theory, the buyer is now willing to invest more than when the seller is entirely selfish. Theorem 2 reveals that $q^*_{un} = 1$ is possible for a larger set of Y_S-values than $q^*_{obs} = 1$ is. The scope for positive reciprocity is thus larger under unobservable investment. At the same time, however, the scope for negative reciprocity is also larger. Only for the unobservable investment condition there exists a SRE in which the buyer invests less than standard theory predicts, i.e. $q^*_{un} < \frac{V}{V+W}$.[10]

The intuition behind the negative reciprocity equilibrium is as follows. For $V < W$ it holds that $\frac{V}{V+W} < \frac{1}{2}$. Therefore, would the buyer choose $q^*_{un} = \frac{V}{V+W}$ as standard theory predicts her to do, the seller considers this as unkind. His reciprocity payoff then induces him to punish the buyer by choosing $P = V + W$ for sure. To counterbalance this the buyer chooses $q^*_{un} < \frac{V}{V+W}$ to let the seller prefer $P = V$ on the basis of monetary payoffs only. Taking both the monetary and the reciprocity payoffs into account, the seller's motives cancel out and he mixes between the two prices. He does so as to make the buyer indifferent between investing or not.[11]

From Theorem 2 it follows that when reciprocity considerations are weak, the buyer invests more under unobservable investment than under observable investment. This situation becomes more likely the higher are the costs of investment C. However, when the seller is sufficiently sensitive to intention-based reciprocity, private information does not boost investment incentives. This case is likely to apply when C is relatively low.[12]

[10]The driving force behind the difference between the two information conditions is that under observable investment a player's kindness and perception of another player's kindness may differ between the various subgames. To illustrate, for a particular (mixed) investment strategy q_{obs} the seller may *at the root* (i.e. before the investment decision is made) view the buyer as unkind. Yet once the subgame after investment is reached, the seller no longer maintains this belief and views the buyer as kind.

[11]The intuition behind the equilibrium with intermediate investment probability $\frac{1}{2} < q^*_{un} < 1$ is similar. Given $q^*_{un} > \frac{1}{2}$ the seller prefers $P = V + W$ on the basis of monetary payoffs and $P = V$ on the basis of reciprocity payoffs. The equilibrium investment probability is such that the two motives cancel out.

[12]Comparing the two equilibria in which $q^*_{un} < 1$ with $q^*_{obs} = 1$ in the 'strong' case sug-

3 Experimental design

The experiment is based on a 2×3 design. For both the observable and unobservable investment case we considered three levels of investment costs: $C \in \{20, 40, 60\}$. The two other parameters always equalled $V = 50$ and $W = 80$. We ran six sessions in total. Three sessions considered the observable investment case, the other three the unobservable investment case. All subjects within a session were confronted with all three values of C. Overall 120 subjects participated, with 20 participants per session. The subject pool consisted of the undergraduate student population of the University of Amsterdam. Sixty percent were students in economics, 64 percent of the participants were male. Average earnings were 27.65 euros in about one and a half hours. Earnings varied considerably, with the minimum actual earnings equal to 7.60 euros and a maximum of 51.50 euros.

The sessions in which the investment was observable necessarily displayed a sequential game structure. We also used a sequential decision structure in the unobservable investment case; subjects knew that buyers decided on their investment before sellers chose their price demand. We did so to make both information conditions fully comparable.[13] To exclude dominated strategies in the observable investment case, the seller could never ask for more than the actual pie. Figure 1 depicts the structure of the experimental games.

[Insert Figure 1 about here]

Each session contained 36 rounds. We employed a block structure to control for learning and order effects. In particular, we divided the 36 rounds into six blocks of six rounds. Within each block the costs of investment were kept fixed. In two out of three sessions per information condition we used the 'upward' order $(20, 40, 60, 20, 40, 60)$ of investment costs. In the remaining session we employed the opposite 'downward' order of $(60, 40, 20, 60, 40, 20)$. By comparing (within a session) different blocks that consider the same value

gests that private information may even *weaken* investment incentives. This only strengthens our qualitative prediction that when C is low, private information cannot be used as an instrument to encourage investments.

[13] According to standard game theory the timing of moves does not matter in the unobservable investment case. All that matters is that the seller is uninformed about the buyer's investment when choosing his price. Experimentally the physical timing may make a difference though. For example, in coordination game experiments typically a *positional order effect* is found. Knowledge that one player makes his decision before the other player does, leads to coordination on the equilibrium that is advantageous for the first mover, see e.g. Rapoport (1997), Müller (2001) and Güth et al. (1998).

of C we can test for learning effects. By comparing the two different orders we can control for order effects.

From each block of six rounds we selected – before the experiment started – one round that was actually paid. After the final round, subjects learned which six rounds were selected and they obtained the number of points they had earned in these rounds, on top of their initial endowment of 75 points. (The conversion rate was one euro for 10 points.) Subjects were explicitly informed about this procedure at the start of the experiment. The rationale for paying only one round per block is that it strengthens the one-shot nature of each interaction.[14]

Subject roles' varied over the rounds. Within each block each subject had the role of buyer three times, and the role of seller also three times.[15] The experiment used a stranger design. Subjects were anonymously paired and their matching varied over the rounds. Within each block subjects could meet each other only once. Subjects were explicitly informed about this. Moreover, within a session we divided the subjects into two groups of ten subjects. Matching of pairs only took place within these groups. This yielded six independent observations per treatment at the aggregate group level.

The experiment was computerized. Subjects started with on-screen instructions. Before the experiment started all subjects had to answer a number of control questions correctly. They also received a summary of the instructions on paper.[16] At the end of the experiment subjects filled out a short questionnaire and the earned experimental points were exchanged for money.

4 Results

In presenting our results we pool the data from sessions that are completely similar in the order of treatments they consider, because no significant differences are found between these sessions. We also pool the results from sessions that differ only in the order of the C-values.[17] Although some order effects

[14]This holds because subjects know that they cannot compensate gains or losses within the same block. In that way our payment procedure also makes it more likely that fairness/reciprocity motivations are restricted to each interaction in isolation.

[15]We used role switching for two reasons. First, it enhances subjects' awareness of the other player's decision problem. Alternating roles provide subjects with an opportunity to see things from the other player's viewpoint and thus to understand the game better. Second, it also doubles the number of investors in the experiment.

[16]A direct translation of this summary sheet can downloaded from: www1.fee.uva.nl/scholar/mdw/sloof/papers.htm.

[17]The test results on order effects and learning effects are reported in a web-appendix available at: www1.fee.uva.nl/scholar/mdw/sloof/papers.htm.

can be detected, these are only minor. Further aggregations are not possible, because it appears that behavior evolves over time. Most findings are therefore reported separately for the first and second half of the experiment.

4.1 Investment levels

Within each block of six rounds subjects have the role of buyer three times. For each subject we calculate for each block his or her mean investment level, which equals either 0, $\frac{1}{3}$, $\frac{2}{3}$ or 1. Statistical tests can then be based on a comparison of these *individual mean* investment levels. Per treatment we have 60 individual investors. In addition we perform our tests on the group level data. As discussed in Section 3 we divided the 20 subjects within a session into two groups that were independently matched. By doing so we created six independent group observations per treatment and we can compare the *group mean* investment levels across treatments. In the sequel we base our inferences on the results of both types of tests. If not stated otherwise, a significance level of 5% is employed.

The first result compares mean investment levels across information conditions.

Result 1. *(a) With high or intermediate investment costs, mean investment levels are higher under unobservable investment than under observable investment. (b) With low investment costs, mean investment levels are independent of the information condition.*

Evidence supporting Result 1 is provided in Table 2. This table reports the mean investment levels by treatment and gives the test statistics for equality of these levels across treatments (ranksum tests). When $C = 20$ the investment rate is independent of whether the investment itself is observable or not. In case $C = 40$ we observe a significant difference only when subjects are confronted with this costs level during the second half of the experiment.[18] For $C = 60$ the difference is significant for both halves, and largest in absolute and relative magnitude.

Our second result compares mean investment levels across different costs of investment.

Result 2. *(a) In both information conditions, mean investment levels are decreasing in the costs of investment. (b) When the costs of investment are high, mean investment levels are very close to the standard predictions.*

[18] When we pool the data from the first and second halves, the difference between the unobservable and observable case is also significant; $p = .0023$ at the individual level and $p = .0247$ at the group level.

Table 2: Mean investment levels by treatment and tests for equality

	first: rounds 1-18			second: rounds 19-36		
	unobs.	obser.	$p-$values	unobs.	obser.	$p-$values
$C = 20$	.711	.772	.1304	.644	.572	.4335
	[.385]	[0]	.4192	[.385]	[0]	.2207
$C = 40$	.561	.461	.1957	.456	.189	.0000
	[.385]	[0]	.1481	[.385]	[0]	.0156
$C = 60$	.333	.122	.0001	.383	.078	.0000
	[.385]	[0]	.0215	[.385]	[0]	.0031

Remark: Standard equilibrium predictions (based on self-interest) within square brackets. $p-$values correspond to a Mann-Whitney ranksum test comparing the unobservable investment case with the observable investment case. For each level of C the upper (lower) p-value is based on individual (group) level data with $m = n = 60$ ($m = n = 6$).

Result 2 follows from comparing the mean investment levels in the different rows of Table 2. Under unobservable investment, mean investment levels fall from around 68% to around 36% when C increases from 20 to 60. With observable investment, mean investment levels fall from around 67% to around 10%. For low costs of investment the mean investment levels are well above the predicted levels of $38\frac{1}{2}\%$ and 0% respectively. But for $C = 60$ mean investment levels are fairly close to these standard predictions. This is especially true during the second half of the experiment. Table 3 reports the relevant p-values. Because comparisons are on a within-subjects basis, we make use of the Wilcoxon signed-rank test for matched pairs. In the observable investment case we observe 5 (out of 6) significant differences. In the unobservable investment case all six comparisons yield significant differences. Hence the negative relationship between investment levels and investment costs appears to be robust.

Results 1(a) and 2(b) are in line with standard equilibrium predictions, Results 1(b) and 2(a) are not. The self-interest model namely predicts that for both information conditions the propensity to invest is independent of C. Social preferences provide an explanation. As discussed in Section 2 the scope for alternative motivational factors like fairness, reciprocity and efficiency decreases with C. Therefore, when C increases, buyers should be less willing

Table 3: p-values of comparative statics tests by information condition

	first: rounds 1-18		second: rounds 19-36	
	unobs.	obser.	unobs.	obser.
$C = 20$ vs. $C = 40$	.0054	.0000	.0003	.0000
	.0273	.0277	.0350	.0273
$C = 20$ vs. $C = 60$	.0000	.0000	.0001	.0000
	.0277	.0273	.0277	.0277
$C = 40$ vs. $C = 60$	.0006	.0000	.1298	.0013
	.0277	.0277	.4593	.0345

Remark: The reported $p-$values correspond to a Wilcoxon signed-rank test. For each comparison the upper (lower) p-value is based on 60 (6) matched pairs of individual (group) mean investment levels.

to invest. This is exactly what we observe. Result 2(b) demonstrates that with C large enough, the impact of alternative motivations is likely to be weak and the predictions of standard theory and social preferences theories will coincide. Overall we conclude that unobservability of the specific investment made does boost investment incentives. But, it only does so when alternative (social preferences) motivations do not provide strong enough incentives to invest.

4.2 Pricing behavior

Although our main interest lies in buyers' investment decisions, to understand these we have to analyze sellers' price demands. In the observable investment case the seller can condition his price on the investment level observed. Here we thus consider the contingencies of no-investment and investment separately. Figures 2 and 3 depict the frequency distributions of price demands by treatment. In these figures separate demand decisions rather than the (individual or group) mean demands are the units of observation. These demands are bunched into intervals of 10 experimental points; demands that are not divisible by 10 are rounded upwards to the nearest multiple of 10. We also group the data from the first 18 and the last 18 rounds, because the shapes of the distributions are very similar over time.

[Insert Figures 2 and 3 about here]

First consider the observable investment case. When no investment is made almost always $P = 50$ is chosen. For all values of C the frequency of exactly this demand is over 90%. These demands are fully in line with standard predictions, but are much higher than those typically observed in dictator games (cf. Camerer 2003). The latter points at the importance of intention-based negative reciprocity. If only distributional preferences would play a role, we would predict no differences between the situation in which the small pie is exogenously fixed (i.e. a dictator game) and the one where it is endogenously chosen (our game). Our finding that there is a difference is in line with previous experimental results that intentions do matter, see e.g. Falk et al. (2000, 2003).

When the buyer invests the price demands are more dispersed. For all cost levels there is a large peak at $P = 130$, with a minimum mass of 39% when $C = 20$. For the higher cost levels the mass equals around 53%. In all three cases there is also a second smaller peak. This peak is at $P = 60/70$ when $C = 60$, at $P = 80/90$ when $C = 40$ and at $P = 100$ when $C = 20$. Here the frequencies are around 25% overall. Note that these second peaks roughly occur at demands $130 - C - \epsilon$, allowing the buyer to make a small return of $\epsilon \leq 10$ on investment. Both outcome-oriented social preferences and positive reciprocity provide an explanation for this.

In the unobservable investment case subjects typically choose between $P = 50$ and $P = 130$, see Figure 3. The frequency with which the low demand is chosen increases with the costs of investment: 29% when $C = 20$, 49% for $C = 40$ and 68% in case $C = 60$. These percentages are well in line with the ones of 25%, 50% and 75% predicted by standard theory. However, there are also clear indications for alternative motivations. Demands between 50 and 130 can be considered fair/reciprocal. We find that the number of these demands is modest, but decreases with C as predicted: 22% when $C = 20$, 10% when $C = 40$ and 5% when $C = 60$.[19]

The upper parts of Table 4 and 5 present the mean demands in the various treatments, together with the (expected) price predicted by standard theory. The lower parts present the $p-$values of signed-rank tests that compare the different costs situations. For the observable investment case the mean demand is largely independent of C. Although after investment the high demand of $P = 130$ is chosen with a higher probability when C is high, the second peak occurs at $130 - C$ which is lower in case C is high. Our data suggests that these two effects cancel out. In the unobservable investment

[19] Reciprocal/fair behavior of the seller can take the form of a demand between 50 and 130 or, alternatively, a mixing strategy between 50 and 130 with a lower probability of the 130-demand than standard theory predicts.

Table 4: Mean demands in observable case and tests for equality

	predictions		first: rounds 1-18		second: rounds 19-36	
	$I=0$	$I=1$	$I=0$	$I=1$	$I=0$	$I=1$
$C=20$	50	130	48.66	101.44	49.12	114.65
$C=40$	50	130	49.11	107.53	48.90	107.94
$C=60$	50	130	48.40	94	49.85	110
$C=20$ vs. $C=40$			.2701 (30)	.4059 (51)	.1765 (50)	.1115 (24)
			1.000 (6)	.1730 (6)	.6002 (6)	.8927 (5)
$C=20$ vs. $C=60$			.2288 (33)	.3923 (18)	.2376 (50)	.1322 (12)
			.7532 (6)	.2249 (5)	.2809 (6)	*n.a.*
$C=40$ vs. $C=60$			.1300 (50)	1.000 (14)	.0770 (60)	.0861 (7)
			.2489 (6)	.0431 (5)	.0277 (6)	*n.a.*

Remark: The columns labeled 'predictions' report the standard equilibrium predictions. $p-$values correspond to a Wilcoxon signed-rank test. For each comparion the upper (lower) p-value is based on individual (group) level data. Within parentheses appear the number of observations (indivual or group means) on which the test is based; these vary because (groups of) sellers may never be confronted with a particular investment choice ($I=0$ or $I=1$) in a treatment. n.a. indicates that no sensible test statistic is available, because there are too few observations.

Table 5: Mean demands in unobservable case and tests for equality

	predictions	first: rounds 1-18	second: rounds 19-36
$C = 20$	110	92.92	102.79
$C = 40$	90	80.00	85.41
$C = 60$	70	67.58	68.43
$C = 20$ vs. $C = 40$		.0019	.0000
		.1159	.0277
$C = 20$ vs. $C = 60$		.0000	.0000
		.0277	.0277
$C = 40$ vs. $C = 60$		.0057	.0000
		.1159	.0277

Remark: The column labeled 'predictions' reports the standard equilibrium predictions. $p-$values belong to a signed-rank test. For each comparison the upper (lower) p-value is based on 60 (6) matched pairs of individual (group) mean investment levels.

case mean demands are significantly decreasing in C. Moreover, actual average demands are somewhat below the expected demand of $P = 130 - C$ predicted by standard theory.

The findings of this subsection are summarized in Result 3.

Result 3. (a) *When $I = 0$ is observed, sellers almost always demand $P = 50$. In case $I = 1$ is observed, sellers demand either $P = 130$ or $P = 130 - C - \epsilon$* (with $\epsilon \leq 10$). *The mean demand does not vary with C.* (b) *When the investment decision is unobserved, sellers typically demand either $P = 50$ or $P = 130$. For higher cost levels the distribution shifts towards $P = 50$. Mean demands are decreasing in C.*

4.3 Learning

In the previous subsections we reported the results for the first and second half of the experiment separately. Comparing the two halves some learning effects can be detected. Learning can have two causes. First, some subjects may understand the subtleties of the game only after a few rounds. Second, subjects may adapt their beliefs about the population characteristics. For example, a buyer who expects sellers to act reciprocal, may be disappointed after some rounds and change the investment decisions accordingly.

First consider the observable investment case. Here sellers practically always ask the whole pie of 50 points when no investment is made. In case the buyer invests, sellers demand a larger part of the pie in the second half of the experiment (cf. Table 4). But, the difference is significant only when $C = 20$. In the first part of the experiment buyers make a modest profit on a low cost ($C = 20$) investment. In the second half this turns into a small loss. As a consequence, selfish buyers prefer not to invest in the second part of the experiment. Investments indeed decrease significantly (for all cost levels), but when $C = 20$ still 57% of the buyers do invest. Note that in this case investment is rational for buyers with a social preference parameter $\sigma_B \geq \frac{5.53}{60}$. On average buyers lose a few points (5.53), but the large increase in earnings for the sellers (65.53) is such that even relatively weak social preferences (for example, weighing own earnings 10 times as heavy as the other player's earnings) make investment worthwhile. This suggests that playing more rounds will not bring the outcome close to the one standard theory predicts.

In the unobservable investment condition the increase in demands is significant only for $C = 20$. However, mean demands stay below the standard prediction of $130 - C$. This indicates that also here social preferences remain to play a role. The decrease in investment levels between the first and second half of the experiment is insignificant for all cost levels.

Table 2 reveals that the difference in mean investment levels between the two information conditions is always higher in the second half of the experiment. This suggests that as subjects learn from previous experience, the effect of private information on investment incentives becomes more important. This appears not to be the case though. The increase in the difference between the investment rates is insignificant for all cost levels.[20] Overall we conclude that learning effects do no affect our main findings with respect to investment behavior.

[20] We formally test this by comparing the change in the unobservable investment rate over time (second half versus first half) with the change in the observable investment rate over time (i.e. we compare differences in differences). Ranksum tests reveal a significant difference only for $C = 40$ when tests are performed at the individual level ($p = .037$). Group level tests yield insignificant differences for all cost levels.

Table 6: Inefficiencies by treatment and tests for equality

	rounds	unobservable			observable	p-values	
		invest	trade	overall	inv/overall	invest	overall
$C = 20$	1-18	17.33	9.17	26.5	13.67	.4657	.0534
	19-36	21.33	11.39	32.72	25.67	.2207	.4688
		[36.92]	[23.08]	[60]	[60]		
$C = 40$	1-18	17.56	10	27.56	21.56	.1481	.2615
	19-36	21.78	13.61	35.39	32.44	.0192	.4201
		[24.62]	[15.38]	[40]	[40]		
$C = 60$	1-18	13.33	9.44	22.77	17.56	.0215	.0054
	19-36	12.33	7.77	20.11	18.44	.0031	.6242
		[12.31]	[7.69]	[20]	[20]		

Remark: Predicted inefficiencies based on standard theory appear in square brackets. p-values refer to Mann-whitney ranksum tests performed on group level data (with $m = n = 6$).

4.4 Efficiency

Finally, we take a look at efficiency. Table 6 reports the mean inefficiency losses by information condition. In the unobservable investment treatment there are two types of inefficiencies. First, the buyer may decide not to invest, leading to lower gains from trade. Second, the seller may demand too much, inducing no trade at all. The latter cannot occur in the observable investment case, because then by design the seller can never demand more than the actual pie. Standard theory predicts a lower investment inefficiency and a higher trade inefficiency under unobservable investment than under observable investment. Overall, however, these inefficiencies are predicted to cancel out. Our final result relates to this.

Result 4. *Investment inefficiency is weakly larger under observable investment, while trade inefficiency is always larger under unobservable investment. When subjects have gained experience overall inefficiencies are not significantly different from each other.*

Result 4 follows from comparing the various inefficiencies by means of ranksum tests. Because efficiency losses can only be calculated for a buyer-seller pair, tests cannot be based on individual means. We therefore only consider tests performed at the aggregate matching group level. In the observable investment case trade inefficiency is zero by design. When the investment is unobservable, the group means are always positive. For all cost levels trade inefficiency is therefore larger under unobservable investment. For the other, more interesting comparisons p-values are reported in Table 6. The second to last column reports the test statistics of comparing investment inefficiencies. The results reiterate our earlier conclusions about mean investment levels. The last column concerns the comparison of overall inefficiencies. Although these are typically smaller when the investment decision is observable, the differences are not statistically significant, with the exception of $C = 60$ in the first 18 rounds. Once subjects have gained experience overall inefficiencies do not vary with the information condition. This concurs with the predictions of standard theory.

5 Conclusion

This paper addresses the question whether making specific investments unobservable boosts investment incentives, as predicted by Tirole (1986) and Gul (2001) among others. Our experimental findings indicate that this will be the case only when there is insufficient scope for social preferences, i.e. when the costs of investment are relatively high compared to the return on investment. In case the costs of investment are relatively low, fairness and reciprocity (and efficiency concerns) are at work and these alternative motivations have a larger impact on investment incentives when the investment decision is observable. As a result, investment levels under the two information conditions are equal when the costs of investment are relatively low. Private information then does not boost investment incentives.

Our results tentatively suggest that private information is an effective instrument only when the efficiency gains of solving holdup are small. Clearly our experiment just provides a first step and a number of interesting questions remain. For instance, the reduced form price-setting stage that we employ is not realistic. In reality the buyer at least has the opportunity to accept or reject the seller's price demand. Standard predictions remain unchanged for such a setup, but the predictions under social preferences change and become much more involved. In particular, the buyer may then want to reciprocate with her acceptance/rejection decision. Anticipating this, the seller may change his demand behavior, which in turn affects investment incentives. These additional strategic issues and motives that come into play will make it more difficult to interpret observed behavior. Now that we have established that in the simplest possible setup unobservability may indeed affect investment incentives (only) when the efficiency gains of underinvestment are small, future experiments can build on this and investigate whether this result generalizes to more natural bargaining settings.

References

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

Camerer, C. (2003). *Behavioral game theory. Experiments in strategic interaction.* Princeton University Press.

Cason, T. and T. Sharma (2001). Durable goods, coasian dynamics, and uncertainty: Theory and experiments. *Journal of Political Economy 109*, 1311–1354.

Charness, G. and M. Rabin (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics 117*, 817–869.

Cox, J. (2004). How to identify trust and reciprocity. *Games and Economic Behavior 46*, 260–281.

Cremer, J. (1995). Arm's length relationships. *Quarterly Journal of Economics 110*, 275–295.

Dufwenberg, M. and G. Kirchsteiger (2000). Reciprocity and wage undercutting. *European Economic Review 44*, 1069–1078.

Dufwenberg, M. and G. Kirchsteiger (2004). A theory of sequential reciprocity. *Games and Economic Behavior 47*, 268–298.

Ellingsen, T. and M. Johannesson (2004a). Is there a holdup problem? *Scandinavian Journal of Economics 106*, 475–494.

Ellingsen, T. and M. Johannesson (2004b). Promises, threats, and fairness. *Economic Journal 114*, 397–420.

Ellingsen, T. and M. Johannesson (2005). Sunk costs and fairness in incomplete information bargaining. *Games and Economic Behavior 50*, 155–177.

Engelmann, D. and M. Strobel (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American Economic Review 94*, 857–869.

Falk, A., E. Fehr, and U. Fischbacher (2000). Testing theories of fairness - Intentions matter. working paper.

Falk, A., E. Fehr, and U. Fischbacher (2003). On the nature of fair behavior. *Economic Inquiry 41*, 20–26.

Fehr, E. and S. Gächter (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives 14*, 159–181.

Fehr, E. and J. List (2004). The hidden costs and returns of incentives - trust and trustworthiness among ceos. *Journal of the European Economic Association 2*, 743–771.

Fehr, E. and K. Schmidt (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics 114*, 817–868.

Fehr, E. and K. M. Schmidt (2002). Theories of fairness and reciprocity - Evidence and economic applications. In M. Dewatripont, L. Hansen, and S. Turnovsky (Eds.), *Advances in Economics and Econometrics - 8th World Congress, Econometric Society Monographs.* Cambridge University Press.

Gantner, A., W. Güth, and M. Königstein (2001). Equitable choices in bargaining games with advance production. *Journal of Economic Behavior and Organization 46*, 209–225.

Gonzalez, P. (2004). Investment and screening under asymmetric endogenous information. *Rand Journal of Economics 35*, 502–519.

Gul, F. (2001). Unobservable investment and the hold-up problem. *Econometrica 69*, 343–376.

Güth, W., S. Huck, and A. Rapoport (1998). The limitations of the positional order effect: can it support silent threats and non-equilibrium behavior? *Journal of Economic Behavior and Organization 34*, 313–325.

Hackett, S. (1993). Incomplete contracting: A laboratory experimental analysis. *Economic Inquiry 31*, 274–297.

Hackett, S. (1994). Is relational exchange possible in the absence of reputations and repeated contract? *Journal of Law, Economics and Organization 10*, 360–389.

Klein, B., R. Crawford, and A. Alchian (1978). Vertical integration, appropriable rents, and the competitive contracting process. *Journal of Law and Economics 21*, 279–326.

Königstein, M. (2000). *Equity, efficiency and evolutionary stability in bargaining games with joint production.* Berlin Heidelberg: Springer Verlag.

Lau, S. (2002). Information and bargaining in the hold-up problem. Working paper, Washington University in St. Louis.

Leuven, E., H. Oosterbeek, R. Sloof, and C. van Klaveren (2005). Worker reciprocity and employer investment in training. *Economica 72*, 137–149.

McCabe, K., M. Rigdon, and V. Smith (2003). Positive reciprocity and intentions in trust games. *Journal of Economic Behavior and Organization 52*, 267–275.

Müller, W. (2001). The quality of the signal matters - a note on imperfect observability and the timing of moves. *Journal of Economic Behavior and Organization 45*, 99–106.

Oosterbeek, H., R. Sloof, and J. Sonnemans (1999). Promotion rules and skill acquistion: An experimental study. Scholar working paper series WP 07/09.

Oosterbeek, H., J. Sonnemans, and S. van Velzen (2003). The need for marriage contracts: An experimental study. *Journal of Population Economics 16*, 431–453.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review 83*, 1280–1302.

Rapoport, A. (1997). Order of play in strategically equivalent games in extensive form. *International Journal of Game Theory 26*, 113–126.

Riordan, M. (1990). What is vertical integratation? In M. Aoki, B. Gustafsson, and O. Williamson (Eds.), *The firm as a nexus of treaties*, Chapter 5, pp. 94–111. London: European Sage.

Rogerson, W. (1992). Contractual solutions to the hold-up problem. *Review of Economic Studies 59*, 777–793.

Schmidt, K. (1996). The costs and benefits of privatization: An incomplete contracts approach. *Journal of Law, Economics and Organization 12*, 1–24.

Sonnemans, J., H. Oosterbeek, and R. Sloof (2001). On the relation between asset ownership and specific investments. *Economic Journal 111*, 791–820.

Tirole, J. (1986). Procurement and renegotiation. *Journal of Political Economy 94*, 235–259.

Williamson, O. (1985). *The economic institutions of free capitalism.* Free Press.

Appendices

A.1 Proof of Theorem 1

We derive the subgame perfect equilibria for the two investment conditions. Theorem 1 directly follows from Propositions 1 and 2.

Proposition 1 (Observable investment) Suppose preferences are given by (1) and players know each others preferences. Then the unique subgame perfect equilibrium is given by:

(a) $\rho_S < \frac{1}{2}$ and $\sigma_B < \frac{C}{C+W}$: $q^*_{obs} = 0$, $P^*_0 = V$ and $P^*_1 = V + W$;

(b) $\rho_S < \frac{1}{2}$ and $\sigma_B > \frac{C}{C+W}$: $q^*_{obs} = 1$, $P^*_0 = V$ and $P^*_1 = V + W$;

(c) $\rho_S > \frac{1}{2}$: $q^*_{obs} = 1$, $P^*_0 = \frac{V}{2}$ and $P^*_1 = \frac{V+W-C}{2}$.

Proof. Owing to $\sigma_S < \frac{1}{2}$ the seller always demands at least half of the net surplus. Whether he demands for more depends on the value of ρ_S.

Parts (a) and (b). Let $\rho_S < \frac{1}{2}$. The seller then considers own payoffs more important also when he is ahead. Hence $P^*_0 = V$ and $P^*_1 = V + W$. For these prices, choosing $I = 0$ gives the buyer $\sigma_B \cdot V$ whereas $I = 1$ yields her $-C + \sigma_B \cdot (V + W + C)$. The former choice is (strictly) preferred when $\sigma_B < \frac{C}{C+W}$, the latter when $\sigma_B > \frac{C}{C+W}$.

Part (c). In case $\rho_S > \frac{1}{2}$ the seller always prefers an equal split of the net surplus and thus demands $P^*_0 = \frac{V}{2}$ and $P^*_1 = \frac{V+W-C}{2}$. No investment then yields the buyer $\frac{V}{2}$, investment gives her $\frac{V+W-C}{2}$. The buyer therefore invests for sure. *QED*

Proposition 2 (Unobservable investment) Suppose preferences are given by (1) and players know each others preferences. Let $h(\rho, \sigma) \equiv (\rho - \sigma)\cdot \max\{(W - C) - V, 0\}$. Then the unique equilibrium is given by:

(a) $\rho_S < \frac{1}{2}$ and $\sigma_B < \frac{C}{C+V+W}$: $q^*_{un} = \frac{(1-\rho_S)V}{(1-\rho_S)V+(1-2\rho_S)W+h(\rho_S,\sigma_S)}$, $P = V$ occurs with probability $p^* = \frac{C-\sigma_B\cdot(V+W+C)}{(C-\sigma_B\cdot(V+W+C))+(1-\sigma_B)(W-C)-h(\rho_B,\sigma_B)}$ and $P = V + W$ with probability $1 - p^*$;

(b) $\rho_S < \frac{1}{2}$ and $\sigma_B > \frac{C}{C+V+W}$: $q^*_{un} = 1$ and $P^* = V + W$;

(c) $\rho_S > \frac{1}{2}$: $q^*_{un} = 1$ and $P^* = \frac{V+W-C}{2}$.

Proof. Parts (a) and (b). Suppose $q_{un}^* = 0$. The seller with $\rho_S < \frac{1}{2}$ then chooses $P = V$ for sure. Given this price, $I = 0$ yields the buyer $\sigma_B \cdot V$ whereas $I = 1$ yields her at least $(1 - \sigma_B) \cdot (W - C) + \sigma_B \cdot V$, which is strictly more. Hence $I = 1$ is a best response, contradicting $q_{un}^* = 0$. Hence, necessarily $q_{un}^* > 0$.

Next suppose $q_{un}^* = 1$. Then the seller chooses $P = V + W$ for sure. Given this price, choosing $I = 1$ yields the buyer a utility of $u_B = -C + \sigma_B(V + W + C)$. A choice for $I = 0$ yields her $u_B = 0$. Hence $q_{un}^* = 1$ necessarily requires that $-C + \sigma_B(V + W + C)] \geq 0$, i.e. $\sigma_B \geq \frac{C}{V+W+C}$. This yields the equilibrium described in part (b).

Finally, let $0 < q_{un}^* < 1$. Clearly, the seller with $\rho_S < \frac{1}{2}$ then necessarily chooses between $P = V$ and $P = V + W$. Let $p \equiv \Pr(P = V)$. First assume that $W - C < V$. Buyer's indifference between no-investment and investment then requires $p \cdot (\sigma_B V) = p \cdot [(W - C) + \sigma_B \cdot (V - (W - C))] + (1 - p) \cdot [-C + \sigma_B \cdot (V + W + C)]$. This implies that $p^* = \frac{C - \sigma_B \cdot (V+W+C)}{(C - \sigma_B \cdot (V+W+C)) + (1-\sigma_B)(W-C)}$. To secure that $p^* \geq 0$ it is required that $\sigma_B \leq \frac{C}{V+W+C}$. The seller is prepared to mix only when he is indifferent, i.e. when $(1 - q)(1 - \rho_S)V + q \cdot [V + \rho_S \cdot ((W - C) - V)] = q \cdot [V + W + \rho_S \cdot (-C - (V + W))]$. This gives $q_{un}^* = \frac{(1-\rho_S)V}{(1-\rho_S)V + (1-2\rho_S)W}$.

Second, let $W - C > V$. For the buyer to be indifferent it is now required that $p \cdot (\sigma_B V) = p \cdot [(W - C) + \rho_B \cdot (V - (W - C))] + (1 - p) \cdot [-C + \sigma_B \cdot (V + W + C)]$. This gives $p^* = \frac{C - \sigma_B \cdot (V+W+C)}{(C - \sigma_B \cdot (V+W+C)) + (1-\rho_B)(W-C) + (\rho_B - \sigma_B)V}$. Again, $p^* \geq 0$ requires that $\sigma_B \leq \frac{C}{V+W+C}$. The seller is prepared to mix only when $(1 - q)(1 - \rho_S)V + q \cdot [V + \sigma_S \cdot ((W - C) - V)] = q \cdot [V + W + \rho_S \cdot (-C - (V + W)]$. This yields $q_{un}^* = \frac{(1-\rho_S)V}{(1-\sigma_S)V + (1-\sigma_S-\rho_S)W + (\rho_S - \sigma_S)C}$. Taking the two cases together and simplifying using $h(\rho, \sigma)$ yields part (a).

Part (c). Suppose $q_{un}^* = 0$. The seller with $\rho_S > \frac{1}{2}$ then chooses $P = \frac{V}{2}$ for sure. For this price $I = 0$ yields the buyer $\frac{V}{2}$ whereas $I = 1$ gives her $(1 - \rho_B) \cdot (W - C) + \frac{V}{2}$, which is strictly more. Hence, necessarily $q_{un}^* > 0$. Given this, any price $P > \frac{V+W-C}{2}$ is strictly dominated for the seller by $P = \frac{V+W-C}{2}$. This also applies for prices satisfying $V < P < \frac{V+W-C}{2}$. Finally, $\sigma_S < \frac{1}{2}$ implies that the seller never demands a price below $\frac{V}{2}$. Hence in equilibrium P is necessarily chosen from $[\frac{V}{2}, V] \cup \{\frac{V+W-C}{2}\}$.

First consider prices $\frac{V}{2} \leq P \leq V$. When the buyer chooses $I = 0$, she obtains $(1 - \sigma_B) \cdot (V - P) + \sigma_B \cdot P$. Investing keeps the seller's material payoffs at $\pi_S = P$ whereas π_B increases by $W - C$. Preferences in (1) entail that the buyer prefers more money to less, other things equal, so investment is strictly preferred. Next consider the remaining case where $P = \frac{V+W-C}{2} > V$. Then choosing $I = 0$ yields the buyer 0 while $I = 1$ gives her $\frac{V+W-C}{2} > 0$. Investment is thus always the best response, implying $q_{un}^* = 1$. Given this, the seller chooses $P^* = \frac{V+W-C}{2}$ for sure. *QED*

A.2 Proof of Theorem 2

We derive the sequential reciprocity equilibria (SRE) for the two investment conditions in two separate subsections. Theorem 2 directly follows from Propositions 3 and 4.

A.2.1 Observable investment

The following additional notation is used. The seller's pricing strategy is denoted (Δ_0, Δ_1). It consists of two probability distributions over $[0, V+W]$, one for each investment decision ($I = 0$ and $I = 1$) separately. In case the seller uses a pure pricing strategy, we use the more convenient notation (P_0, P_1). Expected price demands are denoted $(\overline{P_0}, \overline{P_1})$. Turning to beliefs, b_{ij} gives the (first order) belief of player i about the strategy of player j. For instance, b_{SB} denotes the seller's belief about the buyer's investment strategy $q_{obs} \equiv \Pr(I = 1)$. c_{iji} is used to denote the second order beliefs. It reflects the belief of player i about the belief of player j about the strategy of player i. For example, c_{SBS} denotes the seller's belief about the buyer's belief about the seller's pricing strategy (Δ_0, Δ_1). Because effectively only beliefs about *expected* prices are important, we use (c^0_{SBS}, c^1_{SBS}) to reflect the second order beliefs about $(\overline{P_0}, \overline{P_1})$.

The first and second order beliefs determine the factors κ_{ij} and λ_{iji} in the players' utility functions (2). In an SRE beliefs are necessarily correct: $b_{SB} = c_{BSB} = q^*_{obs}$ and $b_{BS} = c_{SBS} = (\Delta^*_0, \Delta^*_1)$. Observations 1 below characterizes the seller's equilibrium pricing strategy for any $Y_B \geq 0$, while Observation 2 does so for buyer's investment behavior. The results for $Y_B = 0$ are summarized in Proposition 3 at the end of this subsection.

Observation 1 In every SRE $P^*_0 = V$ and $P^*_1 = \min\{V + \frac{2}{Y_S}, V + W\}$.

Proof. We first derive the kindness terms κ_{SB} for any possible price demand $P \in [0, V + W]$. When the buyer chooses $I = 1$ the seller can give her at least 0 and at most $V + W$ (the buyer's investment costs are sunk at this stage and thus do not affect the kindness term). The equitable payoff for the buyer thus equals $\pi^e_B(I = 0) = \frac{1}{2} \cdot [0 + (V + W)] = \frac{V+W}{2}$. This implies $\kappa_{SB}(P, I = 1) = (V + W - P) - \pi^e_B(I = 1) = (\frac{V+W}{2} - P)$. Similarly, after $I = 0$ the equitable payoff for the buyer equals $\pi^e_B(I = 0) = \frac{V}{2}$. Hence $\kappa_{SB}(P, I = 0) = \max\{V - P, 0\} - \pi^e_B(I = 0) = \max\left\{(\frac{V}{2} - P), -\frac{V}{2}\right\}$. For $I = 0$ the reciprocity payoffs of the seller are thus the same for any $P \geq V$. Monetary payoffs π_S equal P for $P \leq V$ and 0 for $P > V$. Hence after $I = 0$ the seller strictly prefers $P = V$ above any $P > V$, so the latter are never chosen in a SRE. Hence necessarily $\overline{P^*_0} \leq V$.

Next we turn to the beliefs about intended kindness. The seller's belief about how much the buyer intends to give him by choosing $I = 1$ equals c^1_{SBS}. For $I = 0$ this is c^0_{SBS}. (Here we in fact already use that in a SRE beliefs have to be correct, so beliefs about Δ_0 can not put any weight on prices $P > V$. That is, trade is always correctly believed to take place.) Hence $\lambda_{SBS}(I = 1, c_{SBS}) = \frac{1}{2}(c^1_{SBS} - c^0_{SBS})$ and $\lambda_{SBS}(I = 0, c_{SBS}) = \frac{1}{2}(c^0_{SBS} - c^1_{SBS})$.

In a SRE beliefs are correct, so that $c^1_{SBS} = \overline{P^*_1}$ and $c^0_{SBS} = \overline{P^*_0}$. Now suppose $\overline{P^*_0} > \overline{P^*_1}$. In that case $\lambda_{SBS}(I = 1, c_{SBS}) < 0$. The seller's reciprocity payoff $Y_S \cdot \kappa_{SB} \cdot \lambda_{SBS}$ is then increasing in P (for $Y_S > 0$), just like his monetary payoff π_S. Hence the seller chooses $P = V + W$ for sure. This implies $\overline{P^*_1} = V + W$ and contradicts the supposition that $\overline{P^*_0} > \overline{P^*_1}$. Hence necessarily $\overline{P^*_0} \leq \overline{P^*_1}$.

From $\overline{P^*_0} \leq \overline{P^*_1}$ it follows that $\lambda_{SBS}(I = 0, c_{SBS}) \leq 0$ in a SRE. In words, no investment is never seen as kind. Seller's utility after $I = 0$ equals $u_S = P + Y_S \cdot (\frac{V}{2} - P) \cdot \frac{1}{2}(c^0_{SBS} - c^1_{SBS})$ for $P \leq V$. This is clearly maximized for $P^*_0 = V$.

Given $P^*_0 = V$, $\lambda_{SBS}(I = 1, c_{SBS}) = \frac{1}{2}(c^1_{SBS} - V)$. Seller's overall utility after $I = 1$ thus equals $u_S = P + Y_S \cdot (\frac{V+W}{2} - P) \cdot \frac{1}{2}(c^1_{SBS} - V)$ for $0 \leq P \leq V + W$. We obtain $\frac{\partial u_S}{\partial P} = 1 - \frac{Y_S}{2}(c^1_{SBS} - V)$. For $c^1_{SBS} < V + \frac{2}{Y_S}$ this is strictly positive, hence $P < \min\{V + \frac{2}{Y_S}, V + W\}$ cannot occur. Similarly, for $c^1_{SBS} > V + \frac{2}{Y_S}$ the derivative is negative, hence $P > \min\{V + \frac{2}{Y_S}, V + W\}$ cannot occur. Necessarily then $P^*_1 = \min\{V + \frac{2}{Y_S}, V + W\}$. *QED*

Observation 2 Equilibrium investment behavior is characterized by the following three implications:

(a) if $Y_S < \frac{2+V\cdot Y_B}{W-C}$, then there exists a SRE with $q^*_{obs} = 0$;

(b) if either (i) $\frac{2+V\cdot Y_B}{W-C} < Y_S < \frac{2}{W-C} \cdot \left(1 + Y_B \cdot \frac{V}{2} - Y_B \cdot \left(\frac{W}{2} - \frac{2}{Y_S}\right)\right)$ or (ii) $\left(1 + Y_B \cdot \frac{V}{2} - Y_B \cdot \left(\frac{W}{2} - \frac{2}{Y_S}\right)\right) < Y_S < \frac{2+V\cdot Y_B}{W-C}$, then there exists a SRE with $q^*_{obs} = \frac{Y_S}{Y_B} \cdot \frac{Y_S(W-C)-2-VY_B}{(4-W\cdot Y_S)}$;

(c) if $Y_S > \frac{2}{W-C}\cdot\left(1 + Y_B \cdot \frac{V}{2} - Y_B \cdot \left(\frac{W}{2} - \frac{2}{Y_S}\right)\right)$, then there exists a SRE with $q^*_{obs} = 1$.

Proof. We first derive the buyer's utility $\pi_B + Y_B \cdot \kappa_{BS} \cdot \lambda_{BSB}$ at the root of the game tree, i.e. before the investment is made. Given correct beliefs about the seller's pricing strategy, monetary payoffs of investment strategy q equal $\pi_B = q \cdot (W - \min\{\frac{2}{Y_S}, W\} - C)$. Turning to the kindness term κ_{BS}, under correct beliefs the buyer can give the seller a material payoff of at least $P_0^* = V$ by choosing $I = 0$ and at most $P_1^* = V + \min\{\frac{2}{Y_S}, W\}$ by choosing $I = 1$. The equitable payoff for the seller thus equals $\pi_S^e = V + \frac{1}{2} \cdot \min\{\frac{2}{Y_S}, W\}$. The kindness of strategy q under correct beliefs $b_{BS} = (P_0^*, P_1^*)$ is then $\kappa_{BS}(q, (P_0^*, P_1^*)) = (q - \frac{1}{2}) \cdot \min\{\frac{2}{Y_S}, W\}$. Finally, consider beliefs about intended kindness λ_{BSB}. The buyer's belief about how much the seller intends to give her by choosing (P_0^*, P_1^*) equals $c_{BSB} \cdot (V + W - C - P_1^*)$. Her belief about what the seller believes is the maximum he can give to the buyer equals $c_{BSB} \cdot (V + W) - C) + (1 - c_{BSB}) \cdot V = V + c_{BSB} \cdot (W - C)$. For the minimum this is $c_{BSB} \cdot (-C)$. The buyer thus believes that the seller believes that the buyer's equitable payoffs are $\frac{V}{2} + c_{BSB} \cdot \left(\frac{W}{2} - C\right)$. Buyer's beliefs about intended kindness then equal $\lambda_{BSB}((P_0^*, P_1^*), c_{BSB}) = c_{BSB} \cdot (V + W - C - P_1^*) - \left[\frac{V}{2} + c_{BSB} \cdot \left(\frac{W}{2} - C\right)\right]$. This reduces to $c_{BSB} \cdot \left(\frac{W}{2} - \min\{\frac{2}{Y_S}, W\}\right) - \frac{V}{2}$.

Taking all terms together we obtain:

$$u_B = q \cdot \left[\left(W - \min\{\frac{2}{Y_S}, W\}\right) - C\right] + \\ Y_B \cdot \left(q - \frac{1}{2}\right) \cdot \min\{\frac{2}{Y_S}, W\} \cdot \left[c_{BSB} \cdot \left(\frac{W}{2} - \min\{\frac{2}{Y_S}, W\}\right) - \frac{V}{2}\right]$$

It follows that when $\min\{\frac{2}{Y_S}, W\} = W$ utility u_B is strictly decreasing in q. Hence $q_{obs}^* = 0$ for $Y_S < \frac{2}{W}$. Note that $\frac{2}{W} < \frac{2 + V \cdot Y_B}{W - C}$.

Next consider the case $Y_S > \frac{2}{W}$. Then $\min\{\frac{2}{Y_S}, W\} = \frac{2}{Y_S}$ and we get $\frac{\partial u_B}{\partial q} = \left((W - C) - \frac{2}{Y_S}\right) + Y_B \cdot \frac{2}{Y_S} \cdot \left[c_{BSB} \cdot \left(\frac{W}{2} - \frac{2}{Y_S}\right) - \frac{V}{2}\right]$. First suppose $q_{obs}^* = 0$. Then $\frac{\partial u_B}{\partial q} \le 0$ is required for $c_{BSB} = 0$. This reduces to $\left((W - C) - \frac{2}{Y_S}\right) \le Y_B \cdot \frac{V}{Y_S}$, i.e. $Y_S \le \frac{2 + Y_B \cdot V}{W - C}$, and yields part (a). Second, suppose that $q_{obs}^* = 1$. Then we must have $\frac{\partial u_B}{\partial q} \ge 0$ at $c_{BSB} = 1$. This requires $W - C - \frac{2}{Y_S} \ge -Y_B \cdot \frac{2}{Y_S} \left(\frac{W}{2} - \frac{2}{Y_S} - \frac{V}{2}\right)$, i.e. $W - C \ge \frac{2}{Y_S} \cdot \left(1 - Y_B \cdot \left(\frac{W}{2} - \frac{2}{Y_S} - \frac{V}{2}\right)\right)$. This gives part (c).

Finally, in order to have $0 < q_{obs}^* < 1$, necessarily $\frac{\partial u_B}{\partial q} = 0$ at $c_{BSB} = q_{obs}^*$. This implies $\left((W - C) - \frac{2}{Y_S}\right) = -Y_B \cdot \frac{2}{Y_S} \cdot \left[q_{obs}^* \cdot \left(\frac{W}{2} - \frac{2}{Y_S}\right) - \frac{V}{2}\right]$. When $Y_B = 0$ this can only hold for the degenerate case $Y_S = \frac{2}{W - C}$. In the sequel we do not consider such knife-edge cases. In case $Y_B > 0$ we obtain:

$$q_{obs}^* = \frac{Y_S}{Y_B} \cdot \frac{Y_S(W - C) - 2 - V \cdot Y_B}{(4 - W \cdot Y_S)}$$

First, suppose $Y_S < \frac{4}{W}$. Then $0 < q_{obs}^* < 1$ requires $\frac{2 + V \cdot Y_B}{W - C} < Y_S < \frac{2}{W - C} \cdot \left(1 + Y_B \cdot \frac{V}{2} - Y_B \cdot \left(\frac{W}{2} - \frac{2}{Y_S}\right)\right)$. The latter can only hold when $-Y_B \cdot$

$\left(\frac{W}{2}-\frac{2}{Y_S}\right)>0$, i.e. $Y_S<\frac{4}{W}$. This yields part (bi). Second, assume $Y_S>\frac{4}{W}$. In that case $0<q^*_{obs}<1$ requires $\frac{2}{W-C}\cdot\left(1+Y_B\cdot\frac{V}{2}-Y_B\cdot\left(\frac{W}{2}-\frac{2}{Y_S}\right)\right)<Y_S<\frac{2+V\cdot Y_B}{W-C}$. This can only hold when $-Y_B\cdot\left(\frac{W}{2}-\frac{2}{Y_S}\right)<0$, i.e. $Y_S>\frac{4}{W}$. This gives part (bii). *QED*

Observation 2 reveals that the probability of investment q^*_{obs} is increasing in Y_S. In case Y_S is low the buyer never invests. When Y_S is sufficiently large, the buyer always invests. Hence for investment to occur, parameter Y_S is decisive. Yet Y_B also plays a minor role. This follows because in a SRE the seller always chooses $P^*_0=V$, and this is seen by the buyer as an unkind act. She therefore may want to punish the seller, and can do so by choosing $q_{obs}=0$. Hence negative reciprocity may make no investment more attractive to the buyer. However, because investment is sufficiently rewarded for Y_S high enough ($P^*_1=V+\frac{2}{Y_S}$ is then low), the buyer always prefers to invest when the seller is sufficiently reciprocal. Assuming $Y_B=0$ therefore yields qualitatively the same results.[21] Proposition 3 below immediately follows from Observations 1 and 2.

Proposition 3 (Observable investment) Suppose preferences are given by (2) with $Y_B=0$. Then the unique SRE is given by:

(a) $Y_S<\frac{2}{W-C}$: $q^*_{obs}=0$, $P^*_0=V$ and $P^*_1=\min\{V+\frac{2}{Y_S},V+W\}$;

(b) $Y_S>\frac{2}{W-C}$: $q^*_{obs}=1$, $P^*_0=V$ and $P^*_1=V+\frac{2}{Y_S}$.

A.2.2 Unobservable investment

The seller's pricing strategy is now given by an unconditional probability distribution over $[0,V+W]$. First and second order beliefs b_{BS} and c_{SBS} are now defined with respect to this strategy. The first observation in this subsection gives a condition on the second order beliefs which structures the subsequent equilibrium analysis.

Observation 3 In any SRE necessarily $\lambda_{SBS}\leq\frac{1}{Y_S}$. Moreover, it holds that:

(a) $\lambda_{SBS}=\frac{1}{Y_S}\Longleftrightarrow q^*_{un}=1$;

(b) $\lambda_{SBS}<\frac{1}{Y_S}\Longleftrightarrow 0<q^*_{un}<1$.

[21] Assuming $Y_B=0$ excludes multiple equilibria, which may exist for $Y_B>0$. To illustrate, suppose $Y_B=Y_S=Y$ and let $W=80$, $V=50$ and $C=20$. Then for $\frac{1}{15}<Y<\frac{1}{5}$ there exist three SRE at the same time: $q^*_{obs}=0$, $q^*_{obs}=\frac{2-10Y}{80-4Y}$ and $q^*_{obs}=1$.

Proof. We first derive the seller's utility $\pi_S + Y_S \cdot \kappa_{SB} \cdot \lambda_{SBS}$. The seller believes that the buyer uses investment strategy b_{SB}. Given these beliefs, monetary payoffs equal $\pi_S = P$ when $P \leq V$ and $\pi_S = b_{SB} \cdot P$ for $P > V$. The seller thinks he can give the buyer at least $b_{SB} \cdot (-C)$ and at most $b_{SB} \cdot (V + W - C) + (1 - b_{SB}) \cdot V$. Hence the equitable payoff for the buyer equals $\pi_B^e(b_{SB}) = \frac{V}{2} + b_{SB} \cdot \left(\frac{W}{2} - C\right)$. By choosing $P \leq V$ the seller intends to give to the buyer $b_{SB} \cdot (V + W - C - P) + (1 - b_{SB}) \cdot (V - P)$. Hence the kindness of such a choice equals $\kappa_{SB}(P \leq V, b_{SB}) = (\frac{V}{2} - P) + b_{SB} \cdot \frac{W}{2}$. Similarly, by choosing $P > V$ the seller intends to give the buyer a payoff of $b_{SB} \cdot (V + W - C - P)$. In that case $\kappa_{SB}(P > V, b_{SB}) = b_{SB} \cdot (V - P) - \frac{V}{2} + b_{SB} \cdot \frac{W}{2}$. The overall utility of the seller equals:

$$\begin{aligned} u_S &= P + Y_S \cdot \lambda_{SBS} \cdot \left[\left(\frac{V}{2} - P\right) + b_{SB} \cdot \frac{W}{2}\right] \text{ when } P \leq V \\ &= b_{SB} \cdot P + Y_S \cdot \lambda_{SBS} \cdot \left[\left(b_{SB} - \frac{1}{2}\right) \cdot V - b_{SB} \cdot P + b_{SB} \cdot \frac{W}{2}\right] \text{ for } P > V \end{aligned} \tag{3}$$

First suppose $\lambda_{SBS} > \frac{1}{Y_S}$. Then from the above expression $\frac{\partial u_S}{\partial P} < 0$ and the seller strictly prefers $P = 0$. But when $P = 0$ for sure, the buyer cannot be kind or unkind to the seller with her investment decision. This implies $\lambda_{SBS} = 0$, a contradiction. Hence necessarily $\lambda_{SBS} \leq \frac{1}{Y_S}$.

Next consider the case $\lambda_{SBS} = \frac{1}{Y_S}$. Then $u_S = \frac{V}{2} + b_{SB} \cdot \frac{W}{2}$ when $P \leq V$ and $u_S = (b_{SB} - \frac{1}{2}) \cdot V + b_{SB} \cdot \frac{W}{2}$ for $P > V$. Suppose $b_{SB} < 1$. Then the seller always prefers $P \leq V$ over $P > V$. Knowing that $P \leq V$, the buyer chooses $q_{un} = 1$ (for $P \leq V$ the buyer cannot be kind or unkind to the seller). The latter contradicts $b_{SB} < 1$ under correct equilibrium beliefs $b_{SB} = q_{un}^*$. Hence necessarily $b_{SB} = q_{un}^* = 1$. In sum, we have the implication $\lambda_{SBS} = \frac{1}{Y_S} \Longrightarrow q_{un}^* = 1$.

Finally, consider the case where $\lambda_{SBS} < \frac{1}{Y_S}$. Suppose $q_{un}^* = 1$. Then under correct beliefs $b_{SB} = q_{un}^*$ we have $u_S = P + Y_S \cdot \lambda_{SBS} \cdot \left[(\frac{V}{2} - P) + \frac{W}{2}\right]$ for all P. From this we obtain $\frac{\partial u_S}{\partial P} = 1 - Y_S \cdot \lambda_{SBS} > 0$. The seller thus wants to choose $P = V + W$ for sure. On the basis of monetary payoffs the buyer then strictly prefers $q_{un} = 0$. Because $P = V + W$ is seen as unkind, the buyer also prefers $q_{un} = 0$ on the basis of the reciprocity payoffs. This contradicts $q_{un}^* = 1$. Next, suppose $q_{un}^* = 0$. Then under correct beliefs $b_{SB} = 0$ it holds that u_S is maximized for $P = V$. But when $P = V$, the buyer strictly prefers $q_{un} = 1$. Again we obtain a contradiction. We obtain the implication $\lambda_{SBS} < \frac{1}{Y_S} \Longrightarrow 0 < q_{un}^* < 1$.

The two derived implications, together with $\lambda_{SBS} \leq \frac{1}{Y_S}$ necessarily, immediately yield the implications in the opposite direction. *QED*

Observation 4 In any SRE with $q^*_{un} = 1$ the seller is indifferent between all $P \in [0, V+W]$. Let $P_l \equiv E[P \mid P \leq V]$, $P_h \equiv E[P \mid P > V]$ and $\tilde{p} \equiv \Pr(P \leq V)$ for the equilibrium strategy Δ^* used by the seller. A SRE with $q^*_{un} = 1$ necessarily requires $(1-\tilde{p}) \cdot P_h = \frac{2}{Y_S}$ and exists if:

$$Y_S \geq \max\left\{\frac{2}{P_h}, \frac{2}{W-C} \cdot \left[\frac{P_h - V}{P_h} - Y_B \cdot \left(\frac{V+W-2P_l}{2} - \frac{2(P_h - P_l)}{Y_S \cdot P_h}\right)\right]\right\}$$

Proof. When $q^*_{un} = 1$ we have from Observation 3 that necessarily $\lambda_{SBS} = \frac{1}{Y_S}$ and thus $u_S = \frac{V+W}{2}$ under correct beliefs $b_{SB} = q^*_{un} = 1$ (cf. expression (3)). The seller's utility is thus independent of his pricing strategy and any price is a best response.

We next derive the buyer's utility $\pi_B + Y_B \cdot \kappa_{BS} \cdot \lambda_{BSB}$ under correct (i.e. equilibrium) first and second order beliefs $b_{BS} = c_{SBS} = \Delta^*$ about the seller's pricing strategy. Monetary payoffs of investment strategy q equal $\pi_B = \tilde{p} \cdot (V - P_l) + q \cdot [(1-\tilde{p})V + W - (1-\tilde{p})P_h - C]$. Turning to κ_{BS}, by choosing $I = 0$ the buyer (correctly) believes to give the seller a monetary payoff of $\tilde{p} \cdot P_l$, while for $I = 1$ this amounts to $\tilde{p} \cdot P_l + (1-\tilde{p}) \cdot P_h$. The equitable payoff for the seller is thus $\pi^e_S = \tilde{p} \cdot P_l + \frac{1}{2}(1-\tilde{p})P_h$. In turn we obtain $\kappa_{BS}(q) = (1-q) \cdot \tilde{p} \cdot P_l + q \cdot [\tilde{p} \cdot P_l + (1-\tilde{p}) \cdot P_h] - \pi^e_S = \left(q - \frac{1}{2}\right)(1-\tilde{p})\,P_h$. Finally, for the beliefs about intended kindness λ_{BSB} we obtain that $\lambda_{BSB} = \tilde{p} \cdot (V - P_l) + c_{BSB} \cdot [(1-\tilde{p})V + W - (1-\tilde{p})P_h - C] - \frac{1}{2}[V + c_{BSB} \cdot (W - C) + c_{BSB} \cdot (-C)]$. This reduces to $\tilde{p} \cdot (V - P_l) + c_{BSB} \cdot [(1-\tilde{p})V + \frac{W}{2} - (1-\tilde{p})P_h] - \frac{V}{2}$. Taken together, buyer's utility equals:

$$\begin{aligned} u_B = {} & \tilde{p} \cdot (V - P_l) + q \cdot [(1-\tilde{p})V + W - (1-\tilde{p})P_h - C] \\ & + Y_B \cdot \left[\left(q - \frac{1}{2}\right)(1-\tilde{p})\,P_h\right] \cdot \\ & \left(\tilde{p} \cdot (V - P_l) + c_{BSB} \cdot [(1-\tilde{p})V + \frac{W}{2} - (1-\tilde{p})P_h] - \frac{V}{2}\right) \end{aligned} \tag{4}$$

For $q^*_{un} = 1$ to be an equilibrium we need $\frac{\partial u_B}{\partial q} \geq 0$ at the correct second order beliefs $c_{BSB} = 1$, i.e. $(1-\tilde{p})V + W - (1-\tilde{p})P_h - C \geq -Y_B \cdot (1-\tilde{p})\,P_h \cdot \left(\frac{V+W}{2} - \tilde{p} \cdot P_l - (1-\tilde{p}) \cdot P_h\right)$. Moreover, the seller's belief about the buyer's kindness should be correct; $\lambda_{SBS} = \kappa_{BS}(1) = \frac{1}{2}(1-\tilde{p})\,P_h$. From Observation 3 we know that $q^*_{un} = 1 \iff \lambda_{SBS} = \frac{1}{Y_S}$. Hence necessarily $(1-\tilde{p}) \cdot P_h = \frac{2}{Y_S}$. To secure $\tilde{p} \geq 0$ this requires $Y_S \geq \frac{2}{P_h}$ and yields the first part of the stated condition. Substituting $(1-\tilde{p}) \cdot P_h = \frac{2}{Y_S}$ in the above inequality $\frac{\partial u_B}{\partial q} \geq 0$ and rewriting yields the second part. If the condition is satisfied, $q^*_{un} = 1$ together with Δ^* satisfying $(1-\tilde{p}) \cdot P_h = \frac{2}{Y_S}$ constitutes a SRE. *QED*

Observation 5 In any SRE with $0 < q^*_{un} < 1$ the seller necessarily *strictly* mixes between $P = V$ and $P = V + W$ only.

Proof. Let $0 < q^*_{un} < 1$. Observation 3 implies that necessarily $\lambda_{SBS} < \frac{1}{Y_S}$. From (3) we then obtain $\frac{\partial u_S}{\partial P} > 0$ for all $P \neq V$. The seller therefore chooses between $P = V$ and $P = V + W$ only. Now suppose $P = V + W$ is not chosen. Under correct beliefs about the seller's strategy, buyer's utility directly follows from (4) when we plug in the values $\tilde{p} = 1$, $P_l = V$ and $P_h = V + W$. The reciprocity term vanishes and we immediately obtain $u_B = q \cdot (W - C)$. This is strictly increasing in q, contradicting $q^*_{un} < 1$. Next suppose $P = V$ is not chosen. Utility u_B then follows from plugging in $\tilde{p} = 0$, $P_l = V$ and $P_h = V + W$ in (4). We obtain $\frac{\partial u_B}{\partial q} = -C + Y_B \cdot (V + W) \cdot \left(-\frac{1}{2}\left[c_{BSB} \cdot W\right] - \frac{V}{2}\right) < 0$. Hence the unique best response is $q_{un} = 0$, contradicting $q^*_{un} > 0$. *QED*

Observation 6 Consider the SRE with $0 < q^*_{un} < 1$. Mixing probabilities q^*_{un} and $p^* \equiv \Pr(P = V)$ are characterized by the solutions (q_{un}, p) to the following two equations:

$$p \cdot W + Y_B \cdot (1-p)\,(V+W) \cdot \left[\left(p - \frac{1}{2}\right) \cdot q_{un} \cdot W - \frac{V}{2}\right] = C \qquad (5)$$

$$q_{un} \cdot (V+W) - Y_S \cdot q_{un} \cdot W \cdot \left[\left(q_{un} - \frac{1}{2}\right)(1-p)\,(V+W)\right] = V \qquad (6)$$

Proof. The buyer's utility given correct beliefs about the seller's strategy follows from plugging in $\tilde{p} = p^*$, $P_l = V$ and $P_h = V + W$ in (4). This yields:

$$\begin{aligned} u_B &= q\,(p^* \cdot W - C) \\ &\quad + Y_B \cdot \left(q - \frac{1}{2}\right)\left[(1-p^*)\,(V+W)\right] \cdot \left[\left(p^* - \frac{1}{2}\right)\left[c_{BSB} \cdot W\right] - \frac{V}{2}\right] \end{aligned}$$

Now $0 < q^*_{un} < 1$ implies that necessarily $\frac{\partial u_B}{\partial q} = 0$ at q^*_{un}. This comes down to $p^* \cdot W - C + Y_B \cdot \left[(1-p^*)\,(V+W)\right] \cdot \left[\left(p^* - \frac{1}{2}\right)\left[c_{BSB} \cdot W\right] - \frac{V}{2}\right] = 0$. In equilibrium second order beliefs are correct, so $c_{BSB} = q^*_{un}$. Plugging this in yields equality (5).

To obtain the seller's utility under correct beliefs about q^*_{un}, note that monetary payoffs equal $pV + (1-p) \cdot q^*_{un} \cdot (V + W)$. From the proof of Observation 3 we have $\pi^e_B(q^*_{un}) = \frac{V}{2} + q^*_{un} \cdot \left(\frac{W}{2} - C\right)$. Hence $\kappa_{SB}(p, q^*_{un}) = p \cdot \left[q^*_{un} \cdot (W - C)\right] + (1-p) \cdot \left[q^*_{un} \cdot (-C)\right] - \pi^e_B(q^*_{un}) = \left(p - \frac{1}{2}\right) \cdot q^*_{un} \cdot W - \frac{V}{2}$. The seller's belief about the intended kindness of the buyer λ_{SBS} is most easily obtained from the kindness function κ_{BS} (derived in the proof of Observation 4) by moving one level up in the belief hierarchy. This yields $\lambda_{SBS} = \left(q^*_{un} - \frac{1}{2}\right)(1 - c_{SBS})(V + W)$, where c_{SBS} is the seller's belief about what the buyer believes about p. Taking all terms together we get:

$$
\begin{aligned}
u_S &= pV + (1-p)\cdot q^*_{un}\cdot (V+W) \\
&\quad + Y_S \cdot \left[\left(p - \frac{1}{2}\right)\cdot q^*_{un}\cdot W - \frac{V}{2}\right]\cdot\left(q^*_{un} - \frac{1}{2}\right)(1 - c_{SBS})(V+W)
\end{aligned}
$$

From Observation 5 we know that necessarily $0 < p^* < 1$. This implies that $\frac{\partial u_S}{\partial p} = 0$ at p^* and reduces to $V - q^*_{un}\cdot (V+W) + Y_S \cdot q^*_{un}\cdot W \cdot \left(q^*_{un} - \frac{1}{2}\right)(1 - c_{SBS})(V+W) = 0$. In equilibrium second order beliefs are correct, so $c_{SBS} = p^*$. This gives equality (6). *QED*

Observations 3 through 6 characterize all possible SRE for the unobservable investment case. Necessarily $q^*_{un} > 0$ (Observation 3), but $q^*_{un} = 1$ only when the seller is sufficiently reciprocal (i.e. Y_S is large enough, cf. Observation 4). The latter equilibria can typically be supported by a continuum of pricing strategies. In any SRE with $0 < q^*_{un} < 1$ the seller necessarily mixes between the two prices $P = V$ and $P = V + W$ though. Because for $q^*_{un} = 1$ to occur parameter Y_S is decisive, assuming that the buyer is not reciprocal at all ($Y_B = 0$) leads to qualitatively the same results. Proposition 4 below summarizes the results for this simpler case and also assumes that $V < W$ (i.e. the case considered in the experiment).

Proposition 4 (Unobservable investment) Suppose preferences are given by (2) with $Y_B = 0$ and consider the case $V < W$. Define $g(q) = q \cdot \left[1 - Y_S \cdot (q - \frac{1}{2})(W - C)\right]$ and let q_l (q_h) be the smallest (largest) solution to the equation $g(q) = \frac{V}{V+W}$. Equilibrium behavior is then characterized by:

(a) $Y_S < \frac{2}{V+W-C}$: there exists a unique SRE with $0 < q^*_{un} = q_l \leq \frac{V}{V+W}$ (a strict inequality for $Y_S > 0$) and $p^* \equiv \Pr(P = V) = \frac{C}{W}$;

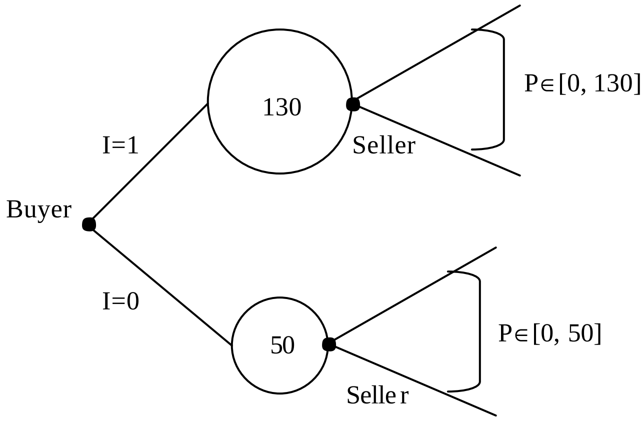
(b) $\frac{2}{V+W-C} < Y_S < \frac{2}{W-C}\cdot\frac{W}{V+W}$: besides the SRE described in part (a), there exists a (continuum of) SRE with $q^*_{un} = 1$. In the latter equilibria equilibrium pricing behavior necessarily satisfies $(1 - \tilde{p})\cdot P_h = \frac{2}{Y_S}$, where $P_h \equiv E[P \mid P > V]$ and $\tilde{p} \equiv \Pr(P \leq V)$;

(c) if $Y_S > \frac{2}{W-C}\cdot\frac{W}{V+W}$: besides all SRE of parts (a) and (b), there exists a SRE with $\frac{1}{2} < q^*_{un} = q_h < 1$ and $p^* \equiv \Pr(P = V) = \frac{C}{W}$.

Proof. Let $Y_B = 0$. First, consider SRE with $0 < q^*_{un} < 1$. For $Y_B = 0$ equality (5) in Observation 6 yields that necessarily $p^* = \frac{C}{W}$. Plugging in this value into equation (6) yields that q^*_{un} follows from the solutions to $g(q) = \frac{V}{V+W}$. Now $g(q)$ is continuous and strictly concave in q, with $g(0) = 0$ and $g(\frac{1}{2}) = \frac{1}{2}$. Moreover, $\frac{\partial g}{\partial q} \geq 1$ at $q = 0$. For $V < W$ we have that $\frac{V}{V+W} < \frac{1}{2}$. From $g(0) = 0$, $g(\frac{1}{2}) = \frac{1}{2}$ and the intermediate value theorem it then follows that necessarily $0 < q_l < \frac{1}{2}$ and $q_h > \frac{1}{2}$. Hence an equilibrium with $q^*_{un} = q_l$ and $p^* = \frac{C}{W}$ exists for any value of Y_S. For $q < \frac{1}{2}$ term $\left[1 - Y_S \cdot (q - \frac{1}{2})(W - C)\right]$ in $g(q)$ weakly exceeds 1, and strictly so when $Y_S > 0$. Hence necessarily $q_l \leq \frac{V}{V+W}$ and a strict inequality for $Y_S > 0$. This yields the equilibrium described in part (a). For $q_{un} = q_h$ to be an equilibrium it is required that $q_h \leq 1$. Given the concavity of $g(\cdot)$ a necessary and sufficient condition for this is that $g(1) \leq \frac{V}{V+W}$. Rewriting this yields that $Y_S \geq \frac{2}{W-C} \cdot \frac{W}{V+W}$ is required. This gives the additional equilibrium described in part (c).

Second, consider SRE with $q^*_{un} = 1$. From Observation 4 it follows that $Y_S \geq \max\{\frac{2}{P_h}, \frac{2}{W-C} \cdot \frac{P_h - V}{P_h}\}$ is necessary requirement. The first argument in the max-term is decreasing in P_h, the second increasing. They are equal for $P_h = V + W - C$. Hence $Y_S \geq \frac{2}{V+W-C}$ is the minimum requirement. A pricing strategy with $P_h \equiv E[P \mid P > V] = V + W - C$, $\tilde{p} \equiv \Pr(P \leq V) = 1 - \frac{2}{Y_S(V+W-C)}$ and $P_l \equiv E[P \mid P \leq V]$ arbitrary then supports $q^*_{un} = 1$. Observation 4 gives the general restrictions on equilibrium pricing behavior. This gives the additional equilibrium of part (b). *QED*

Figure 1a. The experimental game in the obervable investment case

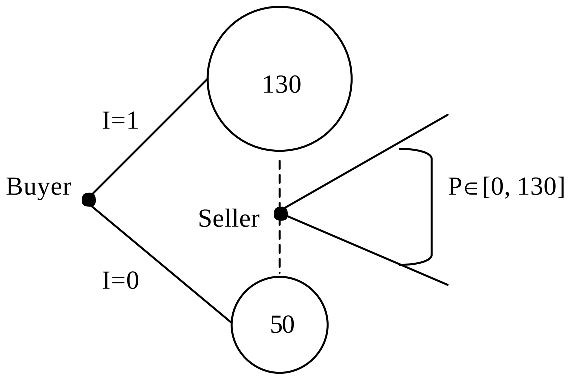


Payoff buyer: $50 + I\cdot 80 - P - I\cdot C$

Payoff seller: P

$C \in \{20,40,60\}$

Figure 1b. The experimental game in the unobervable investment case



Payoff buyer: $\max\{0, 50 + I\cdot 80 - P\} - I\cdot C$

Payoff seller: P if $P \leq 50 + I\cdot 80$
0 otherwise

$C \in \{20,40,60\}$

Figure 2: Frequency distribution of demands in the observable investment case (by costs level)

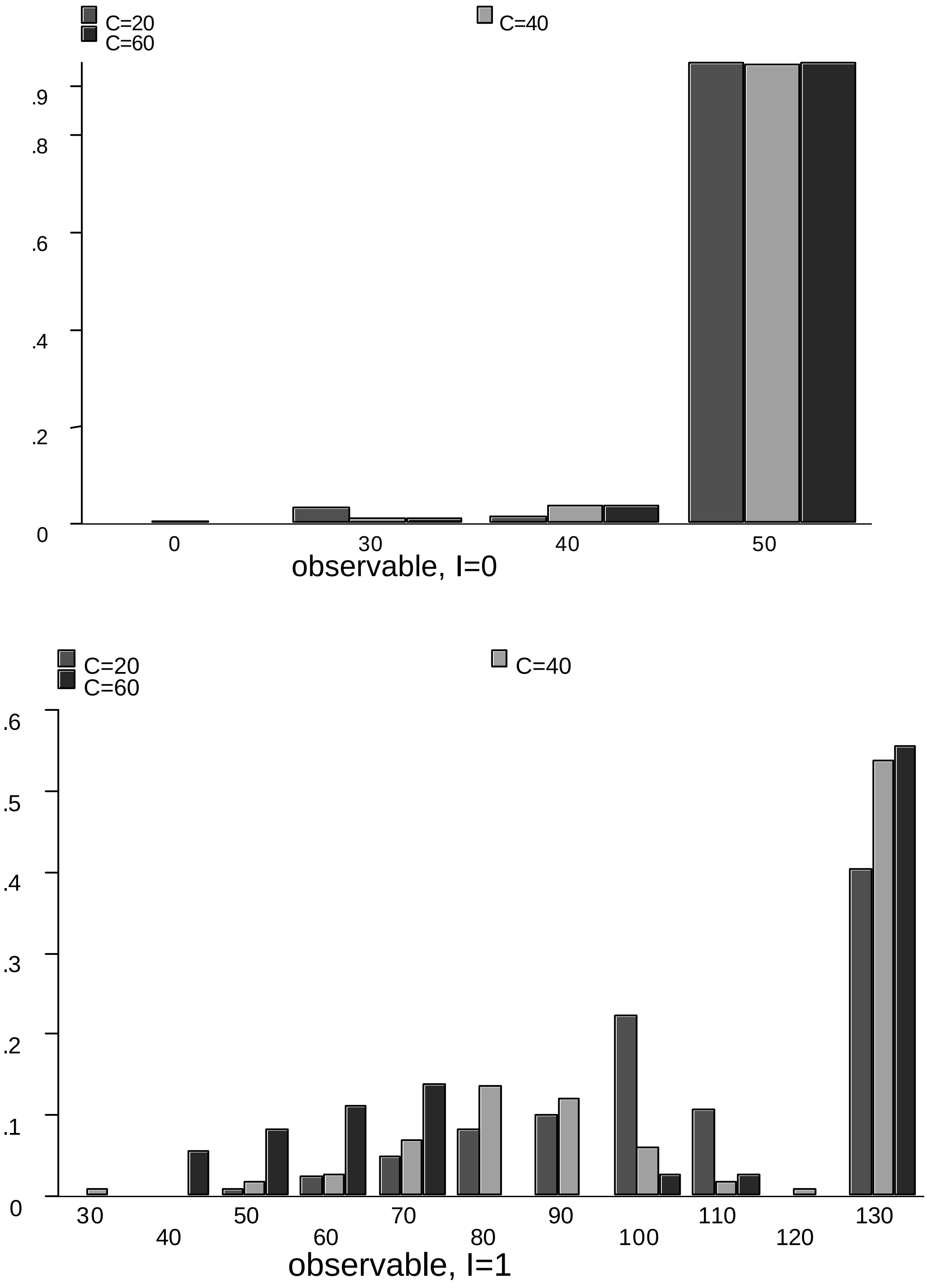


Figure 3: Frequency distribution of demands in the unobservable investment case (by costs level)