On the Interplay of Hidden Action and Hidden Information in Simple Bilateral Trading Problems

Schmitz, Patrick W.

2002

On the Interplay of Hidden Action and Hidden Information in Simple Bilateral Trading Problems*

Patrick W. Schmitz

University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany.

January 2002

A buyer and a seller can exchange one unit of an indivisible good. While producing the good, the seller can exert unobservable effort (hidden action). Then the buyer realizes whether his valuation is high or low, which stochastically depends upon the seller's effort level (hidden information). The parties are risk neutral, they can rule out renegotiation and write complete contracts. It is shown that the first best cannot be achieved whenever the ex post efficient trade decision is trivial. The second-best contract is characterized and an application of the model to the choice of risky projects is briefly discussed.

JEL Classification Numbers: C72, C78, D23, D82

---

* I would like to thank Stefan Reichelstein for pointing out to me that this note, which is based on the third chapter of my Ph.D. dissertation, may be of some interest. I am also very grateful to an anonymous referee, Yeon-Koo Che, Anke Kessler, Georg Nöldeke, Stephanie Rosenkranz, Klaus Schmidt, Urs Schweizer, Dirk Sliwka and several seminar participants for helpful comments.

This is the working paper version of the following article:

Schmitz, P.W. (2002), "On the Interplay of Hidden Action and Hidden Information in Simple Bilateral Trading Problems", *Journal of Economic Theory,* Vol. 103 (2), 444-460.

1 Introduction

Consider the following stylized situation. There are two parties, a (potential) buyer who is interested in one unit of a specific good, and a (potential) seller who can produce this good for the buyer. While producing the good, the seller can exert more or less unobservable effort. When the buyer sees the finished product, he privately realizes whether his willingness-to-pay for the good is high or low, which stochastically depends upon the seller's effort level. Finally, the good can be exchanged and payments can be made. There are no third parties, no liability constraints, and both parties are risk-neutral. What kind of contract should they write? Although this seems to be one of the most basic and natural problems a contract theorist might think of, it has not yet been analyzed. The purpose of this paper is to fill this small but surprising gap in the literature.

It turns out that the simplest model capable of dealing with that sort of problem leads to some interesting insights. The first best cannot be achieved whenever trade is ex post efficient with certainty. This might be surprising for some readers, since none of the sources of inefficiencies most often discussed in the literature are present here. Namely, the parties are neither risk-averse nor wealth constrained, and there is no precontractual private information (i.e., adverse selection).[1] Furthermore, the parties can write complete contracts which are then enforced by the courts, so that no renegotiation-proofness constraints are imposed and money burning is not ruled out. The basic reason for the inefficiency result in this paper is the fact that the parties face a fundamental trade-off between inducing the *seller* to exert effort and inducing the *buyer* to reveal his valuation.

[1] Note that from a pure contract-theoretic point of view, the model developed in this note has an advantage over models which assume that there are rents due to precontractual private information or wealth constraints, because here the efficiency of the optimal contract does not depend upon the distribution of the initial bargaining powers.

The contract which is optimal in a second-best sense can be characterized as follows. If the seller's effort costs are smaller than a certain cut-off value, high effort will be induced by giving up ex post efficient trade sometimes. Specifically, if the buyer's valuation is low, the good is exchanged only with a probability smaller than one. However, if the buyer's valuation is high, trade always occurs, i.e., the well-known "no distortions at the top" property is rediscovered within the framework of this paper. This property has formerly been derived in quite different models involving adverse selection or risk aversion.[2] If the seller's effort costs are higher than the cut-off value, the second-best contract always leads to ex post efficient trade. In this case, however, the seller has no incentives to provide any effort.

This paper is related to the literature in several ways. First, it is well known from the literature on bilateral trading problems that if the parties are symmetrically informed at the contracting stage, hidden information does *not* lead to inefficiencies.[3] Second, it is also well-known from the literature on moral hazard in teams that two-sided hidden action can lead to inefficiencies, provided the parties cannot commit to burn money.[4] In this paper it is argued that inefficiency can be unavoidable if there is one-sided hidden action and one-sided hidden information, even if the parties can commit to burn money.

Moreover, Rogerson [27] showed that if the parties can write complete contracts and commit not to renegotiate, the first-best solution to the hold-up problem *can* be achieved, even if the parties' valuations are private in-

[2]See e.g. the surveys of Hart and Holmström [9] and Fudenberg and Tirole [7, Ch. 7], and the literature cited there.

[3]See d'Aspremont and Gérard-Varet [2]. Recall that Myerson and Satterthwaite's [24] famous impossibility result and the two-type version by Matsuo [20] require precontractual private information. See also Makowski and Mezzetti [16] and Schweizer [31] for references to the more recent literature.

[4]See Holmström [12], and cf. Kim and Wang [13] and Strausz [32] for more recent literature.

formation.[5] However, in the standard formulation of the hold-up problem, parties' investments only influence their own valuations from trade (they are 'selfish'). In contrast, here the seller's investment (i.e., effort) influences the buyer's type. In a complete information framework Maskin and Moore [17] showed that with such 'cooperative' investment the first best may not be achievable if parties cannot commit not to renegotiate.[6] Here it is argued that with hidden information the first best may not be implementable even if renegotiation can be ruled out.[7] Farrell and Gibbons [6] also analyze a trade-off between information revelation and investment incentives. However, their model is driven by adverse selection, while here the parties are *ex ante* symmetrically informed. Finally, Gul [8] also models a hold-up problem with one-sided investment and one-sided private information. He considers 'selfish' investment which deterministically influences the buyer's valuation and assumes that no contracts can be written. Using arguments related to the literature on the Coase conjecture, he shows that when the seller can make repeated offers, the first best is achieved if the time between offers vanishes.

The remainder of the paper is organized as follows. Section 2 describes the basic model. Section 3 demonstrates that no contract can achieve the first best whenever trade is always ex post efficient. Contracts which are optimal in the second-best sense are derived in Section 4. An interesting

[5] See also Konakayama, Mitsui, and Watanabe [15] and proposition 1 (ii) in Hermalin and Katz [11].

[6] Cf. also Che and Hausch's [3] generalization of their result. Note that with 'selfish' investments the first best may be achievable with renegotiation (see Aghion, Dewatripont, and Rey [1], Nöldeke and Schmidt [25], and Edlin and Reichelstein [4] for the case of observable valuations and Schmitz [30] for the case of unobservable valuations).

[7] Che and Hausch [3] argue that under complete information the first best *can* be achieved if renegotiation can be ruled out and if the class of contracts is not restricted. See the working paper version of their article for a discussion of such restrictions. I would like to thank Yeon-Koo Che for pointing out to me this relation to the present note.

application of the basic model is briefly considered in Section 5. Finally, some concluding remarks with regard to renegotiation, the inclusion of third parties, and risk-aversion follow in Section 6.

2 The model

Consider a buyer (B) and a seller (S) both of whom are risk neutral. At some initial date 0, they can write a contract specifying the production and the terms of trade of an indivisible good.[8] At date 1, the seller chooses how much effort to exert while producing the good. Suppose that she can either be lazy ($e = e_l$) or provide a high level of effort, $e = e_h$, where $0 \leq e_l < e_h < 1$. The monetary equivalent of her disutility of effort is represented by $c(e)$, which satisfies $c(e_l) = 0$ and $c(e_h) = c > 0$.[9] The buyer's valuation for the good is determined by the seller's effort level and the state of the world, which is realized at date 2. Specifically, assume that the buyer's valuation is high ($v = v_h$) with probability e and low ($v = v_l$) with probability $1 - e$, where $v_h > v_l > 0$. Finally, trade can occur and payments can be made at date 3. In addition to her effort costs at date 1, the seller may have to incur certain commonly known costs k if trade occurs at date 3.[10]

[8] In an alternative interpretation the good is divisible, but there is a capacity constraint of 1 and the buyer's utility is linear in quantity. In what follows, the term 'probability of trade' can then be replaced by 'level of trade'. Notice that most papers on bilateral trading problems such as Myerson and Satterthwaite [24] assume indivisibility. The contributions of Maskin and Riley [18], McAfee [21], and McKelvey and Page [22] suggest that the case of strictly diminishing marginal utility might be a fruitful topic for further research.

[9] The model can easily be extended to the case $0 < c(e_l) < c$. The assumption $c(e_l) = 0$ does not lead to the loss of any economic insights and is made for expositional purposes only. Note also that additional production costs which are verifiable could be dealt with in a straightforward way.

[10] Since the idea is that the good is valuable for the buyer only, the date-3-costs k may well be zero. They could be positive if there are delivery costs. Alternatively, the seller

Let $q \in [0, 1]$ be the probability of trade and let t^B and t^S denote the payments which the buyer and the seller receive, respectively. Suppose that there is no third party which could act as a budget breaker, i.e., the ex post budget constraint $t^B + t^S \leq 0$ must be met.[11] Then the (expected) utilities of the buyer and the seller after date 2 are given by

$$\begin{aligned} u^B &= t^B + qv \\ u^S &= t^S - qk - c(e). \end{aligned}$$

The reservation utilities of the parties are normalized to zero. Total welfare is measured by the social surplus $u^B + u^S$. Note that if there were no incentive problems, the parties would agree on welfare-maximizing investment, trade and transfer decisions which would yield the first-best expected surplus

$$\Psi^{FB} = \max_{e \in \{e_l, e_h\}} \left(e \left[v_h - k \right]^+ + (1 - e) \left[v_l - k \right]^+ - c(e) \right), \tag{1}$$

where $[\cdot]^+ = \max\{0, \cdot\}$. I say that "the first best can be achieved" if a contract can be written which induces rational parties to act in a way such that the expected surplus is Ψ^{FB}.

Lemma 1 *In order to achieve the first best it is necessary to implement the high effort level,* $e^{FB} = e_h$, *if and only if*

$$c < (e_h - e_l) \left([v_h - k]^+ - [v_l - k]^+ \right).$$

may have the possibility to sell the good for its scrap value if trade does not occur, so that k are the seller's opportunity costs. In both cases it makes sense to assume that k is a fixed number. Note that negative values of k are formally covered by the model (the scrap value may be negative if there are disposal costs).

[11] Actually, the optimal contract will never prescribe any money burning, so that one could make the usual assumption that the seller receives exactly the amount of money which the buyer pays, i.e., $t^B + t^S \equiv 0$. However, if the parties can commit to ex post inefficient choices of q, it is not obvious why commitment to money burning should be ruled out *a priori*.

Proof. This is obvious. ■

Notice that if $k < v_l$, the first best requires that trade always occurs. In this case, it is efficient to exert high instead of low effort if the costs of doing so, c, are smaller than the increase in probability that the buyer's valuation will be high, $e_h - e_l$, times the additional gains from trade if $v = v_h$, which are given by $v_h - v_l$. If $v_l \leq k < v_h$, it is ex post efficient that trade only occurs whenever the buyer's valuation is high. In this case it is efficient to choose $e = e_h$, if the effort costs c are smaller than the increase in probability that the buyer's valuation is high, $e_h - e_l$, times the additional gains from trade, which are then given by $v_h - k$.

3 The basic trade-off

Following the traditional complete contracting (or mechanism design) approach, I assume throughout that the courts enforce contracts specifying trade and payment rules $\left(q, t^B, t^S\right)$ which are based on verifiable variables (possibly including announcements of the parties). It is straightforward to see that the first best could easily be achieved if it were possible to contract upon the level of effort or upon the realization of the buyer's valuation. However, throughout the remainder of this note I assume that *neither* the seller's effort choice e, *nor* the buyer's valuation v is observable.

Due to the revelation principle (cf. Myerson [23]), the analysis can be confined to direct revelation mechanisms which prescribe a probability of trade $q(\tilde{v})$ and transfers $t^B(\tilde{v})$, $t^S(\tilde{v})$ contingent upon the buyer's announcement of his type, $\tilde{v}$. For notational simplicity, let $\left(q_l, t_l^B, t_l^S\right)$ and $\left(q_h, t_h^B, t_h^S\right)$ represent the alternatives between which the buyer can choose by claiming to be of type v_l and v_h, respectively. The incentive compatibility conditions which make truth-telling an optimal strategy for the buyer are

$$\begin{aligned} q_h v_h + t_h^B &\geq q_l v_h + t_l^B && (2) \\ q_l v_l + t_l^B &\geq q_h v_l + t_h^B. && (3) \end{aligned}$$

Notice that only the buyer is required to reveal his private information. It is not necessary to ask the seller which effort level she has chosen, since this is already known in equilibrium. The seller has an incentive to choose the high level of effort whenever the following condition is satisfied:

$$e_h \left(t_h^S - q_h k\right) + (1 - e_h)\left(t_l^S - q_l k\right) - c \geq e_l \left(t_h^S - q_h k\right) + (1 - e_l)\left(t_l^S - q_l k\right)$$

This condition is obviously equivalent to

$$(e_h - e_l)\left(t_h^S - t_l^S - (q_h - q_l)k\right) \geq c. \tag{4}$$

Given that the effort choice e is to be implemented, the following participation constraints for the buyer and the seller, respectively, must hold:

$$\begin{aligned} e\left(q_h v_h + t_h^B\right) + (1 - e)\left(q_l v_l + t_l^B\right) &\geq 0 && (5) \\ e\left(t_h^S - q_h k\right) + (1 - e)\left(t_l^S - q_l k\right) - c(e) &\geq 0 && (6) \end{aligned}$$

The following proposition says that in the (seemingly) trivial case, in which trade is always efficient, the first best cannot be achieved. However, in the non-trivial case, in which it depends upon the realization of the buyer's valuation v whether or not trade is efficient, the first best can be achieved.

Proposition 1 *Assume that $e^{FB} = e_h$.*

a) If $k < v_l$, there is no contract which achieves the first best.

b) If $k \geq v_l$, the first best can be achieved with an appropriate option contract.

Proof. See the appendix. ■

At first sight, one could guess that it should be easier to achieve the first best in the case of a trivial trade decision, but somewhat surprisingly this is not true. The intuition for the result is as follows. If trade is *always* ex post efficient and the first best is to be implemented, then the parties' utilities are entirely determined by the payments they get. Therefore, the buyer will not reveal his type truthfully unless he always gets the same (negative) payment, independent of his report. Yet, if no money may be wasted, this implies that the seller always receives the same amount of money. As a consequence, she has no incentives to provide effort. Note that part a) of Proposition 1 can easily be generalized to the case of a continuous effort choice and general distributions of v. Whenever the ex post efficient trade decision is trivial, so that the first best requires $q(v) \equiv 1$, truth-telling implies that the buyer's payment must be constant, and hence the seller cannot be induced to exert costly effort.

Consider now the non-trivial case in which trade is ex post efficient if and only if the buyer has a high valuation. Then it is possible to achieve ex post efficiency by writing a contract which gives the buyer the option to buy the good at a price that lies in the interval $[v_l, v_h]$. The seller then gets a payment if and only if the buyer's valuation is high. In particular, the price can be set equal to v_h, which makes the seller residual claimant, so that she has an incentive to exert effort. Note that part b) of Proposition 1 also holds if the effort choice is continuous. Since the seller can be made residual claimant, it is still possible to induce the efficient effort decision in this case. However, part b) of Proposition 1 does depend on the assumption that $v \in \{v_l, v_h\}$. If there are more than two possible realizations of the buyer's valuation, truth-telling implies that the buyer's payment must be constant over all valuations for which trade is efficient, and it also must be constant over all valuations for which trade is inefficient. Hence, it is in general no longer possible to make the seller residual claimant, so that the first best may not be achievable.[12]

Remark 1 *If $e^{FB} = e_l$, the first best can always be achieved.*

Proof. See the appendix. ■

4 The second-best contract

Consider now the case in which according to Proposition 1 the first best cannot be achieved. The following proposition characterizes a second-best contract, i.e., a contract which yields the expected surplus that can maximally be achieved under the relevant incentive constraints.

Proposition 2 *Consider the case $e^{FB} = e_h$ and $k < v_l$. Define*

$$\kappa = \frac{(e_h - e_l)^2 (v_h - v_l)(v_h - k)}{(e_h - e_l)(v_h - k) + (1 - e_h)(v_l - k)}.$$

- *If $c \leq \kappa$, the parties implement high effort by writing a contract with the following properties:*
  - *No inefficiencies in the good state:*

$$q_h = 1, \; t_h^B + t_h^S = 0$$

  - *Inefficiently low probability of trade, but no money burning in the bad state:*

$$q_l = 1 - \frac{c}{(e_h - e_l)(v_h - k)}, \; t_l^B + t_l^S = 0$$

[12] It is straightforward to construct examples with three possible realizations of v, such that the first best cannot be achieved even if the efficient trade decision is non-trivial. If v is continuously distributed, the truth-telling constraints imply that the buyer's payment is uniquely determined up to an integration constant by the decision rule (see Myerson and Satterthwaite [24]), so that if the trade decision has to be ex post efficient, no costly effort can ever be induced without inefficient money burning. Details are available from the author upon request.

Furthermore, the transfer payments are determined by $t_h^S = t_l^S + (1 - q_l)\, v_h$ *and* $t_l^S \in [q_l k - (1 - q_l) e_l (v_h - k),\ q_l\,(e_h v_h + (1 - e_h) v_l)]$, *depending upon the distribution of the initial bargaining powers.*

- *If* $c > \kappa$, *the parties implement no effort and trade always takes place. An optimal contract is given by* $q_h = q_l = 1$ *and* $t_h^S = t_l^S = -t_h^B = -t_l^B \in [k,\ e_l v_h + (1 - e_l) v_l]$.

Proof. See the appendix. ■

Proposition 2 says that two cases have to be distinguished. First, if the seller's effort costs are lower than a certain cut-off value κ, the parties want to implement high effort. However, by Proposition 1 this is not possible as long as trade always occurs and no money is burnt. The proof of Proposition 2 shows that giving up trade in some instances is the cheapest way to implement high effort. Specifically, trade always takes place in the good state of the world. This result is reminiscent of the "no distortions at the top" property which has formerly been derived in models that are based on other sources of inefficiencies (precontractual private information or risk aversion). In the bad state of the world, trade only takes place with a probability smaller than one.[13] Second, if the seller's effort costs are higher than the cut-off level κ, it is too expensive to implement high effort. In this case, the parties prefer to achieve ex post efficiency, i.e., trade always takes place.

Remark 2 *The cut-off value* κ *lies in the interval* $(0, (e_h - e_l)\,[v_h - v_l])$. *It is increasing in* v_h, e_h, *and* k, *and decreasing in* v_l *and* e_l.

Proof. The proof is straightforward. ■

[13] Note that one could interpret this mechanism as a modified option contract. The buyer has the option to buy the good for sure at the price t_h^S. If the buyer does not exercise this option, he has to pay t_l^S and the allocation of the good is determined by an appropriate lottery.

The first part of the remark shows that both cases described above can actually occur. The second part confirms the intuition that, *ceteris paribus,* implementing high instead of low effort becomes relatively more attractive compared to achieving ex post efficiency (i.e., always trade) if v_h, e_h or k are increased, while the opposite is true if v_l or e_l are increased.

5 An application: Endogenous preferences for risky projects

Suppose that ex ante the parties can choose between two different technologies (or projects). If technology $i \in \{A, B\}$ is used, the buyer's valuation is v_h^i in the good state and v_l^i in the bad state of the world. For simplicity, assume that $k = 0$ and that there are no further differences between the two technologies. The choice of a technology is assumed to be verifiable.

Definition 1 *Technology A is called* strictly superior *with respect to technology B, if*

$$e_l v_h^A + (1 - e_l)\, v_l^A > e_l v_h^B + (1 - e_l)\, v_l^B \text{ and}$$
$$e_h v_h^A + (1 - e_h)\, v_l^A > e_h v_h^B + (1 - e_h)\, v_l^B.$$

Thus, technology A is strictly superior if it *always* leads to a higher expected valuation, independent of whether the seller exerts low or high effort. At first sight, one might guess that the parties always prefer a strictly superior technology. However, the following proposition shows that this is actually not the case.

Proposition 3 *Let $e^{FB} = e_h$. Suppose that the parties can contractually specify the usage of a technology $i \in \{A, B\}$. They prefer technology B whenever the expected surplus resulting from the second-best contract,*

$$\max\left\{ e_h v_h^i + (1 - e_h) v_l^i - c\left(1 + \frac{(1 - e_h) v_l^i}{(e_h - e_l)\, v_h^i}\right), e_l v_h^i + (1 - e_l) v_l^i \right\},$$

is maximized by $i = B$. This can happen even if technology A is strictly superior.

Proof. The first claim follows immediately from the proof of Proposition 2. In order to prove the second claim, it is sufficient to give an example. One such example is illustrated in Figure 1.[14] ■

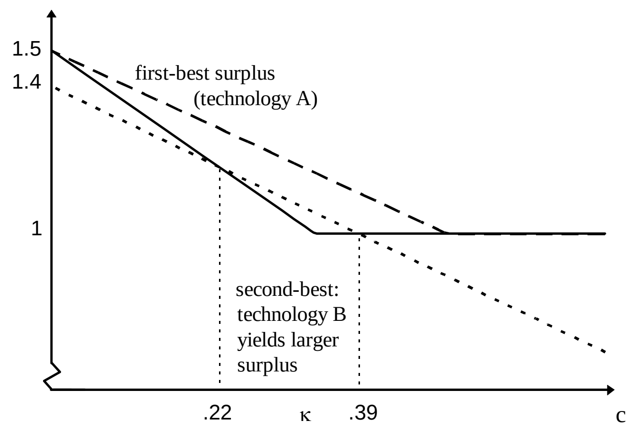

Figure 1

In Figure 1, the solid line represents the expected surplus generated by the second-best contract if technology A is chosen, depending upon the seller's effort costs c. If c is smaller than κ, high effort is induced. Otherwise, the surplus is given by $e_l v_h^A + (1 - e_l) v_l^A$, which is independent of c. As a benchmark, the expected surplus which can be generated in a first-best world is illustrated by the dashed line. Of course, for high levels of c it is optimal to exert no effort even in a first-best world. Note also that for $c = 0$ the first

[14]In this example, $e_h = \frac{1}{2}$, $e_l = 0$, $v_l^A = 1$, $v_h^A = 2$, $v_l^B = .1$, and $v_h^B = 2.7$. Obviously, technology A is strictly superior. However, straightforward calculations show that the expected surplus generated by the second-best contract is higher for technology B, if c is larger than .216 and smaller than $\frac{27}{70} \approx .386$.

best can be achieved. Finally, the dotted line shows the expected surplus which can be generated by a second-best contract if technology B is used. If c is large enough, technology A will certainly be preferred, since by assumption $e_l v_h^A + (1 - e_l) v_l^A > e_l v_h^B + (1 - e_l) v_l^B$. Also, if c is (close to) zero, the superior technology A must by definition (and due to continuity) yield a higher expected surplus than technology B. However, the example shows that there can be an intermediate range of values of c for which technology B generates a higher expected surplus than technology A.

The intuitive feeling that the parties should always prefer a strictly superior technology is justified only in a first-best world. In a second-best world, a superior technology can be suboptimal because it can make the introduction of larger distortions necessary. In other words, even if technology A is superior, it can be easier to induce high effort when technology B is used.

It is interesting to note that although the parties are risk neutral, they may hence prefer a technology (or a project) which is more risky, even if its expected value is smaller. In order to see this, note that B can be preferred over A in the above discussion, if $v_l^B < v_l^A$ and $v_h^B > v_h^A$ (which makes the loss due to the distortion in case B smaller than in A), so that the spread $v_h^B - v_l^B$ is larger than $v_h^A - v_l^A$.

6 Discussion

A. Renegotiation

It has been assumed throughout that the parties can write complete contracts that rule out renegotiation. While this is a common assumption in traditional mechanism design theory,[15] recent papers on incomplete contracts

[15] Cf. Myerson and Satterthwaite [24], Samuelson [29], Matsuo [20], or Klibanoff and Morduch [14], who base their explanation of ex post inefficiencies in bilateral trade relationships on precontractual private information. In these papers it is also assumed that

have stressed the effects of renegotiation. For this purpose, consider again the case $k < v_l$, i.e., the case in which the first-best cannot be achieved. Since it is common knowledge that trade is efficient, perfect renegotiation ex post should always lead to $q = 1$, so that by the proof of Proposition 1 it were impossible to give the seller any incentives to provide effort. Thus, the optimal contract in this case would be what Hart and Moore [10] call the "quintessentially incomplete contract", namely no contract at all.[16] However, as has been pointed out by Maskin and Tirole [19], the assumption that renegotiation cannot be prevented is motivated in the literature by considerations that lie outside the existing models. It is unclear why a court should not enforce a non-renegotiation clause if the parties register their contract publicly.[17] Moreover, as has been pointed out by Rogerson [27], even if one believes that renegotiation cannot always be ruled out in reality, it is important to have a benchmark in order to assess the resulting welfare losses.

B. Inclusion of third parties

Throughout it has been assumed that there are only two parties. The results would change if it were possible to include a risk-neutral third party who could act as a financial wedge between the buyer and the seller.[18] Consider once again the case $k < v_l$ and suppose that the following contract is written. The good is always exchanged and the buyer has to pay t_l to the seller. In addition, independent of whether trade occurs, the buyer always has to pay a fixed premium $t_h - t_l$, either to the seller, or to the third party. Hence, it

renegotiation can be ruled out.

[16] Cf. also Proposition 1 in Rubinstein and Wolinsky [28], who analyze a related problem in a complete information framework.

[17] See also Tirole [33] for a more extensive discussion of this point.

[18] The assumption that there is no budget-breaker is also essential in Holmström's [12] "moral hazard in teams" problem. Note that the inefficiency result obtained here is in a certain sense stronger, since in Holmström's model the first best *could* be achieved if commitment to money burning were allowed.

is an optimal strategy for the buyer to pay the premium to the seller if and only if his valuation is high. It is then possible to give the seller incentives to exert high effort by choosing $t_h - t_l \geq \frac{c}{e_h - e_l}$. Therefore, the first-best trade and effort levels can be implemented without inefficient money burning. The gains from the inclusion of a third party could be divided by an up-front payment.

However, it is not clear whether one could actually find a third party willing to participate in such a contractual agreement. The reason is that a third party would not be willing to make an up-front payment unless it really believed to receive the premium whenever the buyer's valuation is low.[19] But once the buyer's valuation has been realized, he loses nothing if he pays the premium to the seller instead of to the third party. In particular, if there is some possibility of collusion between the buyer and the seller, the potential efficiency gains from the inclusion of a third party may well be prevented from actually being realized.[20]

C. Risk aversion

Following most papers on bilateral trading and hold-up problems, so far it has been assumed that the parties' utility functions are linear in money. It will now be shown that risk aversion can make high effort implementable without money burning or the presence of a third party, even if trade occurs with certainty.[21] Assume that the seller is still risk-neutral, but the buyer is risk-averse in income. Let the buyer's utility be additively separable in money and consumption and let his disutility of having to make a transfer t be given by the strictly concave function $u(-t)$. Suppose that the parties

[19] Note that without an up-front payment from the third party, inclusion of the third party is (from the point of view of the buyer and the seller) the same as money burning. It has already been shown that the buyer and the seller never want to make use of this possibility.

[20] See also Eswaran and Kotwal's [5] discussion of Holmström's [12] model.

[21] I would like to thank an anonymous referee for suggesting the following argument.

write the following contract. The good is always exchanged. When the buyer announces v_i, $i \in \{l, h\}$, his payment to the seller is $\bar{t}_i$ with probability x_i and $\underline{t}_i$ with probability $1-x_i$. Since trade always occurs, the buyer's truth-telling constraints (2) and (3) now require

$$x_h u(-\bar{t}_h) + (1 - x_h)u(-\underline{t}_h) = x_l u(-\bar{t}_l) + (1 - x_l)u(-\underline{t}_l).$$

Adapting (4), the seller chooses $e = e_h$ if

$$(e_h - e_l)\left[x_h\bar{t}_h + (1 - x_h)\underline{t}_h - x_l\bar{t}_l - (1 - x_l)\underline{t}_l\right] \geq c.$$

Due to Jensen's inequality it is possible to satisfy both conditions by designing the two lotteries such that the expected payment if v_h is announced sufficiently exceeds the expected payment if v_l is announced, but the lottery triggered by v_l is riskier such that the risk-averse buyer is indifferent between both lotteries.

Note, however, that this solution does not achieve the first best in the modified model, since in equilibrium the risk-averse buyer is exposed to a stochastic transfer scheme.[22] This is to be contrasted with Rasmusen [26], who has pointed out that Holmström's [12] famous "moral hazard in teams" problem can be solved if the agents are risk-averse. In his model a lottery is used to punish risk-averse agents out of equilibrium only.[23]

[22] This is similar to the standard principal-agent problem, in which the first-best effort level could well be implemented by exposing the agent to risk, but in the first-best solution the agent's wage must be constant (see e.g. Hart and Holmström [9]).

[23] Cf. also Maskin and Moore [17], who demonstrate that their well known result that under symmetric information the hold-up problem cannot be resolved if investments are 'cooperative' and renegotiation cannot be ruled out only holds under risk-neutrality, but not under risk-aversion.

Appendix

A. Proof of Proposition 1.

a) First, consider the case $k < v_l$. The first-best surplus $e_h(v_h - k) + (1 - e_h)(v_l - k) - c$ can only be generated if $q_l = q_h = 1$ and $e = e_h$ are implemented. The incentive compatibility conditions (2), (3), and (4) then read

$$t_h^B - t_l^B \geq 0 \tag{7}$$

$$t_h^B - t_l^B \leq 0 \tag{8}$$

$$t_h^S - t_l^S \geq \frac{c}{e_h - e_l}. \tag{9}$$

The inequalities (7) and (8) imply that the (possibly negative) transfer to the buyer must not depend on the realization of his valuation, i.e., $t_h^B = t_l^B$ must hold. This implies that (9) cannot be satisfied without inefficient money burning. Thus, the first best cannot be achieved.

b) Consider the case $v_l \leq k$. Now the first-best surplus is given by $e_h(v_h - k) - c$, which can be generated by the decisions $q_h = 1$, $q_l = 0$, and $e = e_h$. In this case, the incentive compatibility conditions (2), (3), and (4) become

$$t_h^B - t_l^B \geq -v_h \tag{10}$$

$$t_h^B - t_l^B \leq -v_l \tag{11}$$

$$t_h^S - t_l^S \geq \frac{c}{e_h - e_l} + k. \tag{12}$$

If no money is burnt, $t_h^B = -t_h^S$ and $t_l^B = -t_l^S$, so that (10) and (11) read

$$v_l \leq t_h^S - t_l^S \leq v_h. \tag{13}$$

Note that $v_l < \frac{c}{e_h - e_l} + k < v_h$ (see Lemma 1). It is hence possible to satisfy (12) and (13) simultaneously, e.g. by choosing $t_h^S - t_l^S = v_h$. Thus, the first-best solution can be implemented. Any division of the expected surplus can then be achieved by an appropriate choice of t_l^S, which must lie in the interval $[c - e_h (v_h - k), 0]$, due to the participation constraints (5) and (6). Note that the outlined direct mechanism corresponds to a simple option contract. The buyer always pays t_l^S to the seller. He then has the option to buy the good at a price $t_h^S - t_l^S$. ■

B. Proof of Remark 1.

Consider the following simple contract. The buyer can decide whether or not to trade at the fixed price k. Then, ex post efficiency is achieved and the buyer receives the total gains from trade, so that the seller has no incentive to invest. If the expected surplus is positive, it can be divided between the parties by an additional up-front payment. ■

C. Proof of Proposition 2.

First suppose that the parties want to implement $e = e_h$. The problem of the buyer and the seller is to design a contract $\left(q_h, q_l, t_h^B, t_l^B, t_h^S, t_l^S\right)$ in order to maximize the expected surplus

$$e_h \left(q_h [v_h - k] + t_h^B + t_h^S\right) + (1 - e_h) \left(q_l [v_l - k] + t_l^B + t_l^S\right) - c$$

subject to the incentive compatibility constraints $(2) - (4)$, the participation constraints $(5) - (6)$ with $e = e_h$, the budget constraints

$$\begin{aligned} t_h^B + t_h^S &\leq 0 \\ t_l^B + t_l^S &\leq 0, \end{aligned}$$

and the constraints that q_h and q_l must lie in the unit interval.

First note that (2) and (3) imply $q_h \geq q_l$. Furthermore, note that $t_h^B + t_h^S < 0$ cannot hold in the optimal solution, since else it would be possible to increase the expected surplus by increasing t_h^S.

In order to prove the remaining claims, I first ignore the constraints (3), (5), (6), and $0 \leq q_l \leq 1$. The simplified problem is thus to maximize

$$e_h (q_h [v_h - k]) + (1 - e_h) \left(q_l [v_l - k] + t_l^B + t_l^S\right) - c$$

under the constraints

$$
\begin{aligned}
-v_h(q_h - q_l) + t_l^B &\leq t_h^B && (14)\\
-k(q_h - q_l) - \frac{c}{e_h - e_l} - t_l^S &\geq t_h^B && (15)\\
t_l^B + t_l^S &\leq 0 && (16)
\end{aligned}
$$

and $0 \leq q_h \leq 1$. Suppose that $q_h < 1$. This cannot be true in the optimal solution, since else it would be possible to increase q_h and q_l by the same marginal amount while keeping all other side constraints satisfied. Thus, $q_h = 1$. Conditions (14) and (15) then imply that

$$q_l \leq 1 - \frac{t_l^B + t_l^S + \frac{c}{e_h - e_l}}{v_h - k}$$

must hold. Ignoring (14) and (15), this constraint must be binding, since otherwise q_l could be increased. Now one can easily see that no money must be burnt. For this purpose, assume that (16) holds with strict inequality. To be sure, $q_l = 1$ could then be achieved. However, it would be possible to increase $t_l^B + t_l^S$ by decreasing q_l. Indeed, increasing $t_l^B + t_l^S$ by one marginal unit makes it necessary to decrease q_l by $\frac{1}{v_h - k}$ units. But this means that the objective function's net change would be positive, since $\frac{v_l - k}{v_h - k} < 1$. Finally, it is straightforward to check that the claimed solution satisfies the omitted constraints.

Thus, the expected surplus given that the parties implement $e = e_h$ is

$$e_h v_h + (1 - e_h)\, v_l - k - c - \frac{1 - e_h}{e_h - e_l} \frac{v_l - k}{v_h - k} c, \tag{17}$$

where the last term represents the loss due to the distortion in the bad state of the world. This expression is larger than the surplus in the low-effort case, $e_l v_l + (1 - e_l) v_l - k$, if and only if $c \leq \kappa$. It is obvious to see that $e = e_l$ can be implemented as claimed in the last part of the proposition. ■

References

1. P. Aghion, M. Dewatripont, and P. Rey, Renegotiation design with unverifiable information, *Econometrica* **62** (1994), 257–282.

2. C. d'Aspremont and L.-A. Gérard-Varet, Incentives and incomplete information, *J. Public Econ.* **11** (1979), 25–45.

3. Y.-K. Che and D.B. Hausch, Cooperative investments and the value of contracting, *Amer. Econ. Rev.* **89** (1999), 125–147.

4. A.S. Edlin and S. Reichelstein, Holdups, standard breach remedies and optimal investment, *Amer. Econ. Rev.* **86** (1996), 478–501.

5. M. Eswaran and A. Kotwal, The moral hazard of budget-breaking, *RAND J. Econ.* **15** (1984), 578–581.

6. J. Farrell and R. Gibbons, Cheap talk about specific investments, *J. Law Econ. Organ.* **11** (1995), 313–334.

7. D. Fudenberg and J. Tirole, "Game Theory," MIT Press, Cambridge, MA, 1991.

8. F. Gul, Unobservable investment and the hold-up problem, mimeo Princeton University (1998).

9. O. Hart and B. Holmström, The theory of contracts, *in:* "Advances in Economic Theory" (T. Bewley, Ed.), pp. 71–155, Cambridge University Press, Cambridge, 1987.

10. O. Hart and J. Moore, Foundations of incomplete contracts, *Rev. Econ. Stud.* **66** (1999), 115–138.

11. B. Hermalin and M. Katz, Judicial modifications of contracts between sophisticated parties: A more complete view of incomplete contracts and their breach, *J. Law Econ. Organ.* **9** (1993), 230–255.

12. B. Holmström, Moral hazard in teams, *Bell J. Econ.,* **13** (1982), 324–340.

13. S.K. Kim and S. Wang, Linear contracts and the double moral-hazard, *J. Econ. Theory* **82** (1998), 342–378.

14. P. Klibanoff and J. Morduch, Decentralization, externalities and efficiency, *Rev. Econ. Stud.* **62** (1995), 223–247.

15. A. Konakayama, T. Mitsui, and S. Watanabe, Efficient contracting with reliance and a damage measure, *RAND J. Econ.* **17** (1986), 450–457.

16. L. Makowski and C. Mezzetti, Bayesian and weakly robust first best mechanisms: Characterizations, *J. Econ. Theory* **64** (1994), 500–519.

17. E. Maskin and J. Moore, Implementation and renegotiation, *Rev. Econ. Stud.* **66** (1999), 39–56.

18. E. Maskin and J. Riley, Monopoly with incomplete information, *RAND J. Econ.* **15** (1984), 171–196.

19. E. Maskin and J. Tirole, Unforeseen contingencies, property rights, and incomplete contracts, *Rev. Econ. Stud.* **66** (1999), 83–114.

20. T. Matsuo, Efficient mechanisms for bilateral trading, *J. Econ. Theory* **49** (1989), 189–94.

21. R.P. McAfee, Efficient allocation with continuous quantities, *J. Econ. Theory* **53** (1991), 51–74.

22. R.D. McKelvey and T. Page, Status quo bias in bargaining: An extension of the Myerson Satterthwaite theorem with an application to the Coase theorem, Caltech Working Paper (1998).

23. R.B. Myerson, Optimal coordination mechanisms in generalized principal-agent problems, *J. Math. Econ.* **10** (1982), 67–81.

24. R.B. Myerson and M.A. Satterthwaite, Efficient mechanisms for bilateral trading, *J. Econ. Theory* **29** (1983), 265–281.

25. G. Nöldeke and K.M. Schmidt, Option contracts and renegotiation: A solution to the hold-up problem, *RAND J. Econ.* **26** (1995), 163–179.

26. E. Rasmusen, Moral hazard in risk-averse teams, *J. Econ. Theory* **18** (1987), 428–435.

27. W.P. Rogerson, Contractual solutions to the hold-up problem, *Rev. Econ. Stud.* **59** (1992), 774–794.

28. A. Rubinstein and A. Wolinsky, Renegotiation-proof implementation and time preferences, *Amer. Econ. Rev.* **82** (1992), 600–614.

29. W. Samuelson, Bargaining under asymmetric information, *Econometrica* **52** (1984), 995–1005.

30. P.W. Schmitz, Simple contracts, renegotiation under asymmetric information, and the hold-up problem, *Europ. Econ. Rev.* **46** (2002), 169–188.

31. U. Schweizer, Robust possibility and impossibility results, SFB 303 Discussion Paper No. A-590, University of Bonn (1998).

32. R. Strausz, Efficiency in sequential partnerships, *J. Econ.Theory* **85** (1999), 140–156.

33. J. Tirole, Incomplete contracts: Where do we stand, *Econometrica,* **67** (1999), 741–781.