Bonn Econ Discussion Papers

Discussion Paper 4/2001

Strategic Defection in Bilateral Trade

by

Christoph Lülfesmann

February 2001

Bonn Graduate School of Economics
Department of Economics
University of Bonn
Adenauerallee 24 - 42
D-53113 Bonn

The Bonn Graduate School of Economics is
sponsored by the

Deutsche Post World Net

MAIL EXPRESS LOGISTICS FINANCE

Strategic Defection in Bilateral Trade

by

Christoph Lülfesmann *
University of Bonn

February 2001

Summary

The paper studies a generic bilateral trade model with relationship-specific investments. Only the seller invests, and subsequent trade becomes inefficient when his effort is too low. We show that the seller may defect strategically under a fixed-price contract even though he attains any arbitrary surplus when expending the (second-best) efficient investment. In this case, no general mechanism facilitates trade and the parties should not start their relationship. Also, the defection problem may be more severe when the parties trade after the buyer's valuation has been drawn by nature, as compared to a situation where the parties have to complete trade under uncertainty.

Keywords: Bilateral Trade, Hold-Up, Specific Investments.

JEL-Classification: D23, H57, L51.

---

* *Address*: Department of Economics, University of Bonn, 53113 Bonn, Germany. *email*: clmann@wiwi.uni-bonn.de; *phone*: +49-(228)-737939. This research was started while the author visited UC Berkeley. I would like to thank Thomas Gaube and Anke S. Kessler for helpful discussions, and the Walter A. Haas School of Business for its hospitality. Financial support by DAAD is gratefully acknowledged.

1 Introduction

The recent literature on bilateral trade has found that the implementability of non-contractible investments crucially depends on whether those are 'selfish' or 'cooperative' in nature. Edlin and Reichelstein (1996) show that simple specific-performance contracts often lead to efficiency when investments are selfish, i.e., when they cause no direct externality on the other party.[1] This positive implementation result, though, does not carry over to the case of cooperative investments which directly benefit the trading partner.[2] For the extreme case where investments are purely cooperative and have no selfish element, Che and Hausch (1999) find that the parties should refrain from signing any long-term contract, and instead rely on ex-post bargaining after investments have been expended. They also consider the intermediate case of hybrid investments which affect both value and production costs. There, contracting often remains valuable but in general the first best cannot be attained.

The present note extends these results by identifying a new reason of why contracting may be useless. We study a simplified indivisible-good version of the bilateral trade model where only the seller invests, and where his investments are primarily (but not exclusively) selfish. In the Che and Hausch framework, contracting is then valuable and leads the seller to undertake a certain second-best investment, say e^* which (as we show) can be implemented through a simple fixed-price contract. We then assume that the joint surplus from trade is positive is positive only if the seller undertakes a certain minimum investment, say $\bar{e}$. Under this reasonable assumption, the expected surplus from the relationship is often not concave in investments even if the valuation

[1]This result always holds if only one party expends relationship-specific investments, and if bargaining proceeds according to some linear sharing rule (for example, the Nash-bargaining solution). Hart and Moore (1988), Chung (1991), Aghion, Dewatripont and Rey (1994), Hermalin and Katz (1993) and Nöldeke and Schmidt (1995) have shown that option contracts or specific-performance contracts attain the first best even for two-sided investments if one party can be assigned the entire bargaining power. For monotone sharing rules, Edlin and Reichelstein (1996) show that both parties can be induced to invest efficiently if the cost and valuation functions exhibit certain separability properties.

[2]For example, rather than reducing his costs of production the seller's investments may increase the buyer's value of the good.

and cost functions are well behaved.[3] Since we consider a threshold investment $\bar{e}$ strictly smaller than e^*, one would at first glance presume that e^* can be implemented whether or not $\bar{e}$ is positive. Yet, we find that under a fixed-price contract, the seller may now stategically 'defect' in the sense that he chooses an investment level strictly smaller than $\bar{e}$ to voluntarily render subsequent trade inefficient. If this happens, we further find that even a general mechanism cannot prevent the seller from defection. In plausible situations where trade pays off only if the seller's investment is large enough, the parties may thus be unable to make subsequent trade valuable and should therefore not start their relationship.

To understand this result intuitively, consider for the moment a deterministic setting where the surplus from trade solely depends on the seller's investment decision. Also suppose the parties sign a contract that specifies a fixed price at which trade is to occur. If the surplus from trade is positive even if the seller does not invest at all, the results in Che and Hausch imply that such a contract implements maximal investments as long as those are primarily selfish. Now assume that a positive effort level is required to trigger subsequent trade. If the seller defects so that the surplus from trade becomes negative, the initial contract will have to be renegotiated to ensure that the good is not traded in equilibrium. The resulting renegotiation gain is positive and, moreover, strictly *decreasing* in the seller's investments. As we demonstrate below, the seller's payoff when exerting an investment smaller than $\bar{e}$ can exceed the payoff he obtains when expending e^* (or any smaller investment which induces subsequent trade) under any fixed-price contract with arbitrary trade price. Then, defection is optimal, i.e., the good is not traded in equilibrium and the relationship has no value.

We analyze two variants. We first investigate a situation where the trade decision is taken under uncertainty, i.e., the buyer learns his valuation of the good only after trade has already taken place.[4] Next, we consider the alternative case where the trade decision is taken under certainty, i.e., the state of the world has already been

[3] Thus, Assumption 4 in Che and Hausch (1999) is in general violated.
[4] This scenario is technically equivalent to one where the surplus from trade is deterministic.

realized when the parties proceed to the trade stage.[5] Defection can occur in both scenarios. Under certainty, though, the problem is more severe in that defection arises over a broader parameter range. This implies that the attainable surplus from the relationship may be strictly higher if the parties trade before they learn the buyer's valuation.

The remainder proceeds as follows. Section 2 sets up the model. Section 3 studies the setting where takes place under uncertainty. Section 4 assumes that trade is conducted under certainty and compares both scenarios. Section 5 concludes.

2 The Model

Consider two risk-neutral parties, the seller S and the buyer B, who can trade one unit of an indivisible good. Prior to production, the seller expends relationship-specific investments $e \in \mathbb{R}_0^+$. These investments have two effects: they decrease the seller's production costs $c(e)$, and raise the buyer's valuation $v(e,s)$ of the commodity. The stochastic variable $s \in [\underline{s}, \bar{s}]$ represents a state of the world that is distributed according to a continuous distribution function $F(s)$ and realized after the seller expended his investment (effort). Let t be a monetary transfer from B to S and $q \in \{0,1\}$ be an indicator variable for the trade decision. The parties' utilities are

$$U^S = t - qc(e) - \psi(e) \qquad \text{and} \qquad U^B = qv(e,s) - t, \tag{1}$$

where the function $\psi(e)$ represents the seller's costs of exerting effort. Denote by $V(e) = E_s v(e,s)$ the buyer's expected valuation for given investments. We assume that $v(e,s)$, $c(e)$ and $\psi(e)$ are twice continuously differentiable in e. Also, $v(e,s)$ is increasing in s and and concave in investments, and $c(e)$ and $\psi(e)$ are convex in investments.[6] In line with the literature, the investment level e and the state of the world s are perfectly observed by either party, but non-verifiable. An enforcing party can, however, observe

[5] This setting is commonly investigated in the literature on the hold-up problem. See, e.g., Edlin and Reichelstein (1996) and Che and Hausch (1999).

[6] To ensure interior solutions, we additionally require that either $v(\cdot)$ and $-c(\cdot)$ are strictly concave in e, or that $\psi(\cdot)$ is strictly convex in e.

whether trade occurs and whether S delivers the good (this assumption renders fixed-price contracts and general mechanisms feasible).

In what follows, we will explore two different scenarios. In Section 3, we assume that the parties have to take their trade decision *before* the state of the world s is revealed (scenario UC). This corresponds to a situation where the commodity is an 'experience good' whose actual valuation is learned only after the purchase.[7] Then, the maximal joint surplus is defined by

$$\phi^{uc}(e) = \max_{q \in \{0,1\}} q[V(e) - c(e)]. \tag{2}$$

Since trade is ex-post efficient whenever $V(e) \geq c(e)$, we can define

$$\bar{e} \equiv \inf\{ e \mid V(e) - c(e) \geq 0\} \tag{3}$$

as the (unique) threshold investment below which the joint surplus from trade (gross of investment costs) is negative. Similarly, define $S^{uc}(e) = \phi^{uc}(e) - \psi(e)$ as the expected net surplus. Under our previous assumptions, $S^{uc}(e)$ is strictly concave in e if $\bar{e} = 0$, i.e., if trade is ex-post efficient even though the seller does not invest. If $\bar{e} > 0$, however, efficiency of trade requires some positive minimum investment. In this case, $S^{uc}(e)$ cannot be globally concave as long as it is positive for some investment level: for any $e < \bar{e}$, $\phi^{uc}(\cdot) = 0$ so that $S^{uc}(\cdot)$ is decreasing in e, while $S^{uc}(\cdot) > 0$ for a nonempty range of investments $e \geq \bar{e}$.[8]

Thereafter, in Section 4, we study a setting where the state of the world s has already been revealed when the parties proceed to the trade stage (scenario C). In this case, the maximal joint surplus for given investments e and a given state of the world s is again unique and defined as

$$\phi^{c}(e, s) = \max_{q \in \{0,1\}} q[v(e, s) - c(e)]. \tag{4}$$

[7]The model is also compatible with a scenario where B's valuation is realized only after he exerted an additional investment to increase the good's market value (if it is an intermediary good), or to increase B's intrinsic valuation (when he plans to consume it himself).

[8]Hence, Assumption 4 in Che and Hausch (1999) is violated for $\bar{e} > 0$.

Accordingly, trade is ex-post efficient whenever $v(e,s) \geq c(e)$, while the parties should refrain from trade otherwise. Let

$$\dot{e} \equiv \inf\{\ e \mid v(e,\underline{s}) - c(e) \geq 0\} \tag{5}$$

be the threshold investment above which trade is ex-post efficient even in the worst state of the world, $\underline{s}$. Notice that $\dot{e} \geq \bar{e}$ because, at an investment level $\bar{e}$ where the expected surplus from trade becomes zero, the surplus is still negative in the least favorable state of the world. The expected net surplus in scenario C is indicated as $S^c(e) = \int_s \phi^c(e,s)dF(s) - \psi(e)$ and again strictly concave if $\dot{e} = 0$. Conversely, for $\dot{e} > 0$, small investments render trade inefficient in some unfavorable states of the world. Since higher investments increase the likelyhood of efficient trade for any $e < \dot{e}$, global concavity of $S^c(e)$ can then not be ensured even though the functions $v(\cdot), c(\cdot)$ and $\psi(\cdot)$ are well behaved.

To ease the comparison of both scenarios, we will suppose that trade in scenario C is efficient in every state s if the seller expends first-best investments.[9] As a consequence, first-best investments e^{FB} coincide in both scenarios, and are implicitly determined by the first-order condition (subscripts denote derivatives)

$$\phi_e^{uc}(e^{FB}) = \phi_e^c(e^{FB}) = V_e(e^{FB}) - c_e(e^{FB}) = \psi_e(e^{FB}). \tag{6}$$

The sequence of events is as follows: at date 0, the parties can write an initial contract. Although we mostly focus on non-contingent fixed-price contracts, we will also investigate the outcome of general mechanisms. Subsequently, S can invest at date 1. Thereafter, the parties can trade at date 2 either before (scenario UC) or after (scenario C) the state of the world s is revealed. In either case, B and S can revise their initial contract before the final trade decision is taken. For concreteness, the outcome of renegotiation is described by the generalized Nash-bargaining solution, where the seller's bargaining power is parameterized as $\gamma \in [0,1]$. If the parties finally agree on trade, S produces the good and delivers it to the buyer; payoffs are then realized at date 3.

[9] In other words, it is not optimal to sacrifice trade in a subset of states of the world in order to economize on effort costs. Technically, this assumption is equivalent to $e^{FB} > \dot{e}$.

We will say that S's investments are *purely selfish* whenever $c_e(.) < 0$ and $v_e(.) = 0$, while they are *purely cooperative* when $c_e = 0$ and $v_e(.) > 0$ for all e. Also, define

$$\tilde{\gamma} \equiv \sup\{\, \gamma \,|\gamma v_e(e,s) + (1-\gamma)c_e(e) \leq 0 \,\forall\, e,s\}$$

and, following Che and Hausch (1999),[10]

$$\hat{\gamma} \equiv \inf\{\, \gamma \,|\gamma v_e(e,s) + (1-\gamma)c_e(e) \geq 0 \,\forall\, e,s\}$$

as parameters which indicate the cooperativeness of investments. Both measures are increasing in the degree of selfishness, and are equal to unity if investments are purely selfish. Likewise, they decrease if investments become more cooperative, and are zero for purely cooperative investments. Note that, for a given tuple (e,s), $\gamma v_e(e,s) + (1-\gamma)c_e(e)$ is an increasing function of γ. Since $\tilde{\gamma}$ (respectively, $\hat{\gamma}$) represents the largest (smallest) bargaining parameter for which this expression is non-positive (non-negative) for arbitrary (e,s), we must have $\hat{\gamma} \geq \tilde{\gamma}$.[11] In what follows, we will say that investments are *primarily selfish* if $\gamma \leq \tilde{\gamma}$, while they are *primarily cooperative* if $\gamma \geq \hat{\gamma}$.

As stated earlier, we will focus on situations where the seller can trigger a no-trade outcome through expending sufficiently small effort. For comparison with the literature, however, consider first a setting where $\bar{e} = \dot{e} = 0$. Then, trade is always ex-post efficient irrespective of e (and s in scenario C), and the surplus functions $S^{uc}(e)$ and $S^{c}(e)$ are strictly concave. We can now state the following results which go back to Edlin and Reichelstein (1996) and Che and Hausch (1999):

Proposition 1. *Suppose $\bar{e} = \dot{e} = 0$. Then, in either scenario UC and C, the first-best effort level e^{FB} can be implemented if investments are purely selfish or if $\gamma = 1$. Otherwise, the seller underinvests under the optimal initial contract. More specifically,*

(1) if investments are primarily selfish $(\gamma \leq \tilde{\gamma})$, a fixed-price contract t implements the maximal attainable effort level e^*, which is implicitly determined by $-c_e(e^*) = \psi_e(e^*)$;

[10]The parameter $\hat{\gamma}$ corresponds to the measure $\underline{\alpha}$ in their paper.

[11]Generically, $\tilde{\gamma} = \hat{\gamma}$ only if valuation and cost function are linear in investments and the valuation function is separable in e and s.

(2) *if investments are primarily cooperative ($\gamma \geq \hat{\gamma}$), maximal investments prevail when the parties sign no initial contract, and bargain over trade after investments have been expended. The second-best effort level e^{**} is then implicitly defined by $\gamma[V_e(e^{**}) - c_e(e^{**})] = \psi_e(e^{**})$.*

Proof: see the Appendix.

If investments are primarily selfish, the optimal contract is a simple fixed-price arrangement which obliges the seller to deliver the item in exchange for a monetary transfer t. A fixed-price contract gives rise to a payoff function $U^S = t - c(e) - \psi(e)$. Hence, S reaps the full marginal return from her investments in cost reduction, while the increase in the valuation of the good accrues to B. As a consequence, equilibrium investments coincide with the first-best level e^{FB} if they are purely selfish, while underinvestment cannot be avoided otherwise.[12] Conversely, when investments are primarily cooperative, equilibrium effort can be raised when the parties refrain from signing any initial agreement, and solely rely on bargaining after effort has been expended. Then, S obtains a fraction γ of the overall return from the relationship, which leads him to invest even if investments are purely cooperative.[13]

In the remainder of this paper, we will focus on the case where investments are primarily selfish, $\gamma \leq \tilde{\gamma}$.[14] Hence, when strategic defection is not taken into account, a fixed-price contract t is optimal and implements the effort level e^* which equates the marginal reduction in production costs, and the marginal costs of effort. Throughout Sections 3 and 4, we suppose e^* to be large enough (or, equivalently, investments to be sufficiently selfish) so that $S^i(e^*) > 0$, $i \in \{uc, c\}$. Hence, the relationship should be started from

[12] In the more general framework analyzed in Che and Hausch (1999), first-best investment can (for some non-empty interval of bargaining parameters) be implemented even if investments are not purely selfish. To understand the difference to our model, notice that a fixed-price contract in their variable-quantity setting can prescribe a trade level in excess of the expected efficient quantity [see also Edlin and Reichelstein (1996)]. While renegotiation ensures that the ex-post efficient quantity is traded in equilibrium, a large default quantity raises investment incentives and can induce overinvestments as long as investments are not too cooperative.

[13] In an intermediate interval $\gamma \in (\tilde{\gamma}, \hat{\gamma})$, more general contracts may be second-best optimal. See Che and Hausch (1999).

[14] When investments are primarily cooperative, the best initial contract is no contract and defection has no value for the seller.

an ex-ante point of view provided e^* can be implemented.

3 Trade Decision under Uncertainty

Imagine that the state of the world s is realized after trade has taken place [scenario UC]. Then, the parties can base their trade decision at date 2 only on the observed investment level e. In contrast to the scenario analyzed in Proposition 1, we now suppose that $\bar{e} > 0$, i.e., the parties will not trade for sufficiently small investments. We will show two main results: first, a positive $\bar{e}$ may make it impossible to implement e^*. Second, in this event, no contract can implement an investment large enough to trigger subsequent trade, and the parties will not start their relationship.

To this end, consider again a deterministic fixed-price contract t. At first glance, one may think that the maximand of the seller's optimization program remains e^*, because for any $e < \bar{e}$ the joint surplus $\phi^{uc}(e)$ is zero. Nevertheless, the seller may now have an interest to 'defect' and to expend a small effort level which prevents subsequent trade. To see why this can happen for any trade price t, observe that the seller's utility is now given by the piecewise defined function

$$U^S(e) = \begin{cases} t - c(e) - \psi(e) & \text{for } e \geq \bar{e} \\ t - c(e) + \gamma[0 - (V(e) - c(e))] - \psi(e) & \text{for } e < \bar{e}. \end{cases} \tag{7}$$

If S expends $e < \bar{e}$ at stage 1, the parties will at date 2 renegotiate the initial contract to ensure a now efficient no-trade outcome. In this bargaining process, the seller receives a fraction γ of the renegotiation gain, which is the difference between the (zero) joint surplus when no trade occurs, and the (negative) joint surplus when the initial contract is executed.

Notice that e^* is the seller's optimal choice from the subset of effort levels $e \geq \bar{e}$ which trigger trade. For the complementary set $e < \bar{e}$, define a defection effort $e^D \in [0, \bar{e})$ as the local maximizer. Hence, e^* is globally optimal under a fixed-price contract whenever

$$t - c(e^*) - \psi(e^*) \geq t - (1-\gamma)c(e^D) - \gamma V(e^D) - \psi(e^D). \tag{8}$$

Analyzing this condition, we can state the following result:

Proposition 2. *Suppose investments are primarily selfish and $\bar{e} > 0$. Under a fixed-price contract, e^* can generically be implemented and trade occurs if investments are purely selfish. Conversely, if investments are primarily selfish, it is impossible to implement any $e \geq \bar{e}$ whenever (a) $\gamma_{uc} < \tilde{\gamma}$ where*

$$\gamma_{uc} =: \frac{[c(e^D) - c(e^*)] - [\psi(e^*) - \psi(e^D)]}{c(e^D) - V(e^D)} > 0, \tag{9}$$

and (b) the seller's bargaining parameter is some $\gamma \in (\gamma_{uc}, \tilde{\gamma})$.

Proof: We first show that the first-best investment $e^{FB} = e^*$ can generically be implemented if investments are purely selfish, i.e., $V(e) \equiv V \; \forall \; e$. To see this, notice that (8) holds iff

$$c(e^*) + \psi(e^*) < (1 - \gamma)c(e) + \gamma V(e) + \psi(e) \quad \forall \; e < \bar{e}. \tag{10}$$

Since $S^{uc}(e^*) > 0$ by assumption, we have $c(e^*) + \psi(e^*) < V < c(e)$ for all $e < \bar{e}$ and the result follows. Next, if investments are hybrid, (8) is violated if $\gamma > \gamma_{uc}[> 0]$.[15] Hence, defection under a fixed-price contract arises if the seller's bargaining parameter γ is (a) larger than γ_{uc}, and (b) smaller than $\tilde{\gamma}$ which is required for primarily selfish investments. In the appendix, we provide an example where this condition applies so that a fixed-price contract cannot implement any $e \geq \bar{e}$ for a non-empty range of bargaining parameters $(\gamma_{uc}, \tilde{\gamma})$. $\square$

The next proposition shows that this negative result cannot be overcome when general contracts (which involve a direct revelation mechanism) are admitted.

Proposition 3. *Suppose that investments are primarily selfish, i.e. $\gamma < \tilde{\gamma}$. Then, if a fixed-price contract does not implement e^*, there exists no general mechanism which renders trade feasible.*

[15] The threshold parameter γ_{uc} is strictly positive because $c(e^*)+\psi(e^*) < c(e^D)+\psi(e^D)$ by definition of e^*, and $c(e) - V(e) > 0$ for all $e < \bar{e}$.

Proof: By the revelation principle, we can restrict attention to a direct revelation mechanism which prescribes the (pre-renegotiation) contract terms as a function of both parties' announcements on the investment level e. Define these announcements as e_B and e_S, respectively. Then, the mechanism $\{\delta(e_B, e_S), t(e_B, e_S)\}$ specifies the probability of trade $\delta \in [0,1]$ and the monetary transfer t to the seller in dependence of the announcement profile. Let e_1 and e_2 be two different investment levels with $e_1 > \bar{e} > e_2$, and define the parties' post-renegotiation utilities (gross of the seller's investment costs) as $u^B(e)$ and $u^S(e)$, respectively. Due to the efficiency of renegotiation and since trade is realized only if $e \geq \bar{e}$, we have $u^S(e_1) + u^B(e_1) = V(e_1) - c(e_1)$ and $u^S(e_2) + u^B(e_2) = 0$. Suppose first that S expended e_2. Then, incentive-compatibility of the direct mechanism requires that

$$u^S(e_2) \geq t(e_2, e_1) - \delta(e_2, e_1)[c(e_2) + \gamma(V(e_2) - c(e_2)]. \tag{11}$$

Next, assume that S expended e_1. Then, B announces truthfully iff

$$u^B(e_1) \geq -t(e_2, e_1) + \delta(e_2, e_1)V(e_1) + (1 - \delta(e_2, e_1))(1-\gamma)[V(e_1) - c(e_1)]. \tag{12}$$

Since the possibility of renegotiation ensures that $u^S(e_1) + u^B(e_1) = V(e_1) - c(e_1)$, condition (12) translates into

$$u^S(e_1) \leq t(e_2, e_1) + \gamma(1 - \delta(e_2, e_1))V(e_1) - [\delta(e_2, e_1) + (1 - \delta(e_2, e_1)\gamma]c(e_1). \tag{13}$$

Combining (11) and (12), we obtain

$$\begin{aligned} u^S(e_1) - u^S(e_2) \quad &\leq \quad \gamma(1 - \delta(e_2, e_1))V(e_1) - [\delta(e_2, e_1) + (1 - \delta(e_2, e_1)\gamma]c(e_1) \\ &+ \quad \delta(e_2, e_1)[(1-\gamma)c(e_2) + \gamma(V(e_2)]. \end{aligned} \tag{14}$$

The derivative

$$\frac{du^S(e_1) - u^S(e_2)}{d\delta(\cdot)} = [c(e_2) - c(e_1)] - \gamma[V(e_1) - V(e_2) + c(e_2) - c(e_1)] \tag{15}$$

is positive if $\gamma < \tilde{\gamma}$. Hence, $\delta(\cdot) = 1$ implements the maximal payoff difference between any two investment levels characterized by $e_1 > \bar{e} > e_2$, so that $u^S(e_1) - u^S(e_2) \leq [c_2 - c_1] + \gamma[V(e_2) - c(e_2)]$. Since this difference is (weakly) smaller than the seller's payoff difference under a fixed-price contract, the result follows. □

4 Trade under Certainty

We will now show that, if the state of the world s is revealed prior to trade [scenario C], it may be even more difficult to implement e^* and to facilitate trade if investments are primarily selfish. Define $\hat{s}(e)$ as a threshold state where, for given investments e, trade is ex-post inefficient whenever $s < \hat{s}(e)$. Also, indicate $q(e) = 1 - F(\hat{s}(e))$ as the (equilibrium) trade probability for given investment level and assume that $q(0) > 0$. Let $\dot{e} > 0 \Leftrightarrow q(0) < 1$ so that trade is ex-post efficient with less than full probability if the seller expends zero investments. To ease comparison with scenario UC, suppose that $S^c(e^*) > S^c(e)$ for all $e < e^*$, and that trade is always efficient if the second-best effort level e^* is implemented. Hence, e^* generates a larger joint surplus than any smaller investment, and $q(e^*) = 1$ so that trade is efficient in every state s for investment e^*.[16]

Replicating our previous arguments from scenario UC, one can verify that a fixed-price contract t again implements the maximum attainable effort level if investments are primarily selfish. For subsequent reference, define $V^+(e) = \int_{s \geq \hat{s}(e)} v(e,s)dF(s)/q(e)$ as the buyer's expected valuation of the commodity for given ex-post optimal trade decision and given e, and let $V^-(e) = \int_{s<\hat{s}(e)} v(e,s)dF(s)/(1-q(e))$ be his expected valuation over the complementary set of states. Under a fixed-price contract, S then chooses e to maximize

$$U^S(e) = t - c(e) + [1 - q(e)]\gamma[c(e) - V^-(e)] - \psi(e), \tag{16}$$

where the expression in brackets again represents the gain from renegotiation if trade is ex-post inefficient, and $[1 - q(e)]$ denotes the likelyhood of this event for given e. Since $q(0) < 1$, this payoff function is not necessarily concave.[17] The feasibility of e^*

[16] For $q(e^*) < 1$, the second-best effort level would be smaller than e^* and implicitly determined by $-q(e)c_e(e) = \psi(e)$.

[17] Observe that e^* is the local maximum in the range $e \geq \dot{e}$ and also constitutes the global optimum if $U^S(e)$ is strictly concave.

requires

$$U^S(e^*) = t - c(e^*) - \psi(e^*) \geq t - [1 - q(e)]\gamma[V^-(e) - c(e)] - \psi(e) = U^S(e) \quad \forall \; e < \dot{e}. \tag{17}$$

We will now show that condition (17) can be violated. In addition, we prove that S shirks and triggers no trade even in situations where defection does not arise in scenario UC. To prove these claims, notice first that we can confine attention to a subset of possible deviations $e < \bar{e}$ because $\dot{e} \geq \bar{e}$ implies that the range of possible defections is (weakly) larger in C than in UC. The seller's expected utility for $e = e^*$ coincides in scenarios UC and C for given t. Now consider a deviation $e < \bar{e}$. Then, the seller's payoff in scenario C is larger than in scenario UC if [by (7) and (16)]

$$c(e) - V(e) < [1 - q(e)][c(e) - V^-(e)]. \tag{18}$$

Recalling that $V(e) = (1 - q(e))V^-(e) + q(e)V^+(e)$, we can rewrite (18) as $q(e)V^+(e) > q(e)c(e)$, which holds for any $e < \bar{e}$ by the definition of $V^+(\cdot)$. In terms of bargaining parameters (and recalling that e^D indicates S's optimal defective investment in UC), defection in C thus always occurs for

$$\gamma > \gamma_c =: \frac{[c(e^D) - c(e^*)] - [\psi(e^D) - \psi(e^*)]}{c(e^D) - V(e^D) + q(e^D)[V^+(e^D) - c(e^D)]}. \tag{19}$$

Note that $\gamma_c > \gamma_{uc}$ because $q(e) > 0$. Therefore, we can state

Proposition 4. *Suppose the parties trade after observing s, that investments are primarily selfish, and that e^* is the second-best optimal effort level. Then, e^* can be implemented unless $\gamma_c < \tilde{\gamma}$ and the seller's bargaining parameter is some $\gamma \in (\gamma_c, \tilde{\gamma})$. Since $\gamma_c < \gamma_{uc}$, the range of bargaining parameters that lead to defection is larger in scenario C as compared to scenario UC.*

5 Conclusion

This paper investigates a bilateral trade relationship where the seller can undertake investments that have a cost-decreasing and a value-enhancing effect. We have focused on a natural scenario where trade becomes inefficient (or inefficient with positive probability) when the seller's investments are too low. Perhaps surprisingly, we found that it may be impossible to implement the second-best effort level which prevails when the ex-post efficiency of trade requires no minimum investment. In particular, the seller may have an incentive to defect strategically and to choose an investment strategy for which subsequent trade will not be realized. If strategic defection occurs under a simple fixed-price contract, there exists not even a general mechanism which facilitates successful trade, so that the relationship should not be started. Finally, defection may arise for a broader set of parameters when the agents trade after learning the state of the world, than otherwise. If technologically feasible, the parties may thus find it useful to complete the trade transaction before the buyer's actual valuation is realized.

6 Appendix

Proof of Proposition 1

By the revelation principle, we can restrict attention to a direct revelation mechanism which prescribes the (pre-renegotiation) contract terms as a function of both parties' announcements on the state θ where $\theta = (e)$ in scenario UC, and $\theta = (e, s)$ in scenario C. Define these announcements as θ_B and θ_S, respectively. Then, the mechanism $\{\delta(\theta_B, \theta_S), t(\theta_B, \theta_S)\}$ specifies the probability of trade $\delta(\cdot) \in [0, 1]$ and the monetary transfer $t(\cdot)$ to the seller in dependence of the announcement profile. Define the parties' post-renegotiation utilities (gross of the seller's investment costs) as $u^B(\theta)$ and $u^S(\theta)$, respectively. Then, by the constant-sum nature of the game, we have $u^S(\theta) + u^B(\theta) = B(\theta) - c(\theta)$. Finally, define $\phi(\theta) = B(\theta) - c(e)$, where $B(\theta) = V(\theta)$ in scenario UC, and $B(\theta) = v(\theta)$ in scenario C. Consider now two states θ, θ', and suppose first that the true state is θ'. Then, incentive-compatibility of the direct mechanism requires that

$$u^S(\theta') \geq t(\theta', \theta) - \delta(\theta', \theta)c(\theta') + (1 - \delta(\theta', \theta))\gamma\phi(\theta'). \tag{20}$$

Next, suppose the true state is θ. Then, B announces truthfully iff

$$u^B(\theta) \geq -t(\theta', \theta) + \delta(\theta', \theta)B(\theta) + (1 - \delta(\theta', \theta))(1 - \gamma)\phi(\theta). \tag{21}$$

Since the possibility of renegotiation ensures that $u^S(\theta) = B(\theta) - c(\theta) - u^B(\theta)$, conditions (20) and (21) translate into

$$u^S(\theta) - u^S(\theta') \leq \gamma(1 - \delta(\cdot))[\phi(\theta) - \phi(\theta')] - \delta(\cdot)[c(\theta) - c(\theta')]. \tag{22}$$

Fixing $\theta' = (e', s)$ in C (and $\theta' = e'$ in UC), this inequality implies that

$$\begin{aligned} \frac{\partial u^S(\theta)}{\partial e} &= \limsup_{e' \to e} \frac{du^S(\theta) - u^S(\theta')}{e - e'} \\ &\leq \gamma\phi_e(\theta') - \liminf_{e' \to e} \delta(\cdot)[\gamma v_e(\theta) + (1 - \gamma)c_e(\theta)] \begin{cases} \leq & \gamma\phi_e(\theta) \;\text{ if } \gamma \geq \hat{\gamma} \\ < & \phi_e(\theta) \;\text{ if } \gamma < \hat{\gamma} \\ \leq & -c_e(e) \;\text{ if } \gamma \leq \tilde{\gamma}. \end{cases} \end{aligned} \tag{23}$$

Noting that $U^S_e(\theta) \leq u^S_e(\theta) - \psi_e(e)$ in UC and $U^S_e(\theta) \leq \int_s u^S_e(\theta)dF(s) - \psi_e(e)$ in scenario C, first-best investments are unfeasible unless investments are purely selfish or $\gamma = 1$. For $\gamma \geq \hat{\gamma}$, the expression in brackets in (23) is non-negative so that $\delta(\cdot) = 0$ implements the maximal marginal utility difference between two neighboring states. Accordingly, it is second-best optimal to sign no initial contract. Finally, for $\gamma < \tilde{\gamma}$, the expression in brackets in (23) is non-positive. Thus, $\delta(\cdot) = 1$ implements maximal investments which coincide with those under a fixed-price contract. □

Example

Suppose the buyer's valuation function is $v(e,s) = s + me$, the seller's production costs are $c(e) = f - ke$, and his investment costs are $\psi(e) = e^2/2$. Let m, f, k be non-negative parameters, and $s \in \{\underline{s}, \bar{s}\}$ be a state of the world. State $\bar{s}$ is realized with probability $q \in (0,1)$, and $\bar{s} > f > \underline{s}$ implying that $\dot{e} = [f - \underline{s}]/(m+k) > 0$. Let $s =: q\bar{s} + (1-q)\underline{s} < f$ so that $\bar{e} = [f - s]/(m+k) > 0$, and notice that $\dot{e} > \bar{e}$ and $V(e) = s + me$. In this framework, $\tilde{\gamma} = \hat{\gamma} = k/(m+k)$. To simplify the analysis, we suppose that $\gamma = 1/2$ and, in order to focus on primarily selfish investments, assume $k > m$. If first-best and second-best investments are positive, these levels are $e^{FB} = m + k$ and $e^* = -c_e(e^*) = k$, respectively. To guarantee positive second-best investments, the following condition must hold:

$$S^{uc}(e^*) > 0 \quad \Longleftrightarrow \quad f - s < \frac{1}{2}k^2 + mk. \qquad (C1)$$

In UC, S defects under a fixed-price contract t if condition (8) is violated for some $e \in [0, \bar{e}]$. In case of defection, he chooses $e^D = k - [m+k]/2$ (for an interior solution; otherwise, defection cannot be optimal). Inserting e^D and rearranging terms, defection thus arises if

$$f - s > k^2 - [e^D]^2 = \frac{3}{4}k^2 + \frac{1}{2}mk - \frac{1}{4}m^2. \qquad (C2)$$

Combining conditions (C1) and (C2), we find that defection cannot be avoided and (by Propositon 3) the bilateral relationship has no value whenever[18]

$$\frac{3}{4}k^2 + \frac{1}{2}mk - \frac{1}{4}m^2 < f - s < \frac{1}{2}k^2 + mk. \qquad (24)$$

[18] In particular, we can check that the parties cannot facilitate trade by signing no initial contract. Then, S does not invest if his payoff for the investment level $e = (m+k)/2$ is negative, i.e., if $f - s > (m+k)^2/4$. By conditions (C1) and (C2), S will therefore not invest for the range of parameters where defection arises under a fixed-price contract.

Note that these conditions hold for a generic set of parameter values. In particular, as long as $k < m(1+\sqrt{2})$, there exists a nonempty set of parameter differences $(f-s)$ for which (24) is satisfied.[19]

Finally, we check that defection in UC (and thus in C as well) can arise if e^* is the second-best optimal effort in scenario C. If $V(e^*)-c(e^*)-\psi(e^*) > q[v(\hat{e},\bar{s})-c(\hat{e})]-\psi(\hat{e})$ for $\hat{e}$ implicitly defined by $-qc_e(\hat{e}) = \psi_e(\hat{e})$, e^* is indeed second-best optimal in C. In our example, $\hat{e} = qk$ and we can rewrite the above condition as

$$f - s < [\frac{1}{2}k^2 + mk](1-q^2) - q[\bar{s} - f]. \qquad (C1')$$

As expected, (C1') is more restrictive than condition (C1) for any $q \in (0,1]$. For $q \to 0$, (C1') is identical to (C1), while (C1') cannot hold for $q \to 1$. Since the right-hand side of (C1') is monotonically decreasing in q, defection in UC is thus compatible with (C1') for q being sufficiently small.

[19] Conversely, (24) cannot hold if investments are purely selfish, i.e., $m = 0$.

References

Aghion, P., Dewatripont, M. and P. Rey (1994): Renegotiation Design with Unverifiable Information, *Econometrica*, 62: 257-282.

Che, Y.-K. and D.B. Hausch (1999): Cooperative Investments and the Value of Contracting, *American Economic Review*, 89: 125-147.

Chung, T.-Y. (1991): Incomplete Contracts, Specific Investments and Risk Sharing, *Review of Economic Studies*, 58: 1031-1042.

Edlin, A. and S. Reichelstein (1996): Holdups, Standard Breach Remedies, and Optimal Investment, *American Economic Review*, 86: 478-501.

Hart, O. and J. Moore (1988): Incomplete Contracts and Renegotiation, *Econometrica*, 56: 755-785.

Hermalin, B.E. and M.L. Katz (1993): Judicial Modifications of Contracts Between Sophisticated Parties: A More Complete View of Incomplete Contracts and Their Breach, *Journal of Law, Economics and Organization*, 9: 230-255.

Nöldeke, G. and K.M. Schmidt (1995): Option Contracts and Renegotiation: A Solution to the Hold-up Problem, *RAND Journal of Economics*, 26: 163-179.