Strategic Shirking in Bilateral Trade

by

Christoph Lüllesmann*
Simon Fraser University

September 16, 2014

Summary

The paper investigates a bilateral trade relationship where the seller can undertake specific investments before the trade transaction takes place. In this setting, we identify a novel reason for hold up and contractual inefficiency. Specifically, an investing party may shirk for strategic reasons, that is, exert an effort so low that trade becomes inefficient and does not occur in equilibrium. Under a fixed-price contract (the second-best arrangement in the standard scenario where trade is always beneficial), strategic shirking can arise irrespective of the precontracted price. Moreover, if strategic shirking arises under a fixed-price contract, no general revelation mechanism exists which restores efficient trade. The shirking problem becomes more severe when the agents trade under certainty (value and costs of trade are known), relative to a situation where the agents trade before this information has become available. Finally, we establish that when the buyer can also undertake specific investments, shirking and non-shirking equilibria may coexist.

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

JEL-Classification: D23, H57, L51.

*Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada. Email: chuelles@sfu.ca. Phone: +1-778 782 8504. Financial support by SSHRC is gratefully acknowledged.
1 Introduction

Incomplete contracting models often study relationships where partners first invest into the success of their venture before performing the actual transaction (Grossman and Hart, 1986, 1990; Hart and Moore, 1988). Since both agents can revise any initial contract after investments have been made, Coasian bargaining ensures the efficiency of the final ‘trade’ outcome in this setting. On the other hand, finding a contractual arrangement that induces partners to exert optimal investments into the relationship can be challenging. As shown in the literature, the ensuing hold-up problem can be resolved when relationship-specific investments inflict no direct externality, for example, when the seller’s effort reduces her cost of producing the good (Edlin and Reichelstein, 1996).

Conversely, an efficient outcome may be elusive when investments by one agent inflict a direct externality on the other agent, for example, when a seller enhances the quality of the trading good. In this latter case, long-term contracts may even be detrimental so that second best investments require the absence of an initial trade agreement (Maskin and Moore, 1999; Che and Hausch, 1999).

The present paper suggests a novel reason why long-term contracting may be without merit in environments with relationship specific investments. Even stronger, we identify circumstances in which even though the relationship has value, agents should not start it.

---

1 This is always the case if only one agent invests, and when renegotiation follows a monotonic sharing rule. For both-sided investments, Chung (1991), Denski and Sappington (1991), Aghion et al. (1994), Hermalin and Katz (1993) and Nöldeke and Schmidt (1995) achieve the first best in setting where one agent holds the full bargaining power. Edlin and Reichelstein (1996) extend this efficiency result to monotonic sharing rules under the condition that cost and valuation functions exhibit certain separability properties.

2 This result provides a justification for models in the early hold up literature which rule out long-term contracts exogenously (Grout, 1984; Grossman and Hart, 1986; Williamson, 1985). Note that Che and Hausch also explore hybrid investments which affect both valuation and production costs. Settings with cooperative or hybrid investments have also been investigated in MacLeod and Malcomson (1993), Aghion and Tirole (1994), Schmitz (2002), Roider (2004) and Lülfesmann (2005), among others.
because it is destined to fail. To make these points, the paper studies a simple scenario in which only the seller invests. We show that she may deliberately shirk and eradicate all benefits from trade, by investing so little that trade yields no positive benefits and advancing the relationship becomes undesirable. A single key factor drives this result. In distinction to the workhorse model in the literature, we posit a positive investment threshold below which trade ceases to have a positive net value.\(^3\) Positive investment thresholds seem very natural in business transactions. A buyer may refrain from ordering a custom made item if she is not satisfied with the design proposed by the seller, or a joint project may be stopped prematurely if one party did not work hard enough to identify cost saving potentials during the planning stage.\(^4\) But similar issues often pertain to the private spheres of individuals, where specific investments often decide whether a personal relationship thrives or fails.

To examine whether the seller will actually shirk and invest below the threshold, consider first a scenario in which trade is always efficient regardless of investments. In this standard model, the optimal contractual arrangement induces the seller to expend investments at the second-best level \(e^*\), and trade is realized. Next, consider a situation in which investments above a threshold \(\bar{e}\) is needed to make final trade viable. With \(\bar{e} < e^*\), one may think that the threshold has no effect on the outcome of the bilateral relationship. In contrast to this belief, though, the investment threshold is shown to have a serious impact as it can lead the seller to invest strictly below \(\bar{e}\), rather than at level \(e^*\). Rational agents will then

\(^3\)The investment threshold makes the expected surplus from the relationship non-convex, in distinction to the standard models analyzed in the literature. See Section 2 below.

\(^4\)Problems of this type are well documented for many industries. In defence procurement, the development of an advanced weapon system requires a substantial amount of investments in form of human and physical capital (Kovacic, 1991; Rogerson, 1994). If the contractor’s investment is deemed insufficient, government buyers may stop the procurement cycle after the design phase. In biotechnology research, pharmaceutical companies terminate projects with a Biotech research firm after initial tests do not turn out sufficiently promising to warrant further steps towards completion (Lerner and Merges, 1998; Schmidt, 2003). Chakravarty and Mc Leod (2007) examine the use of standardized contracts in environments with staged and partially verifiable investments.
rescind their initial agreement and give up their trade intentions.

The intuition behind this case of destructive shirking is most easily understood for a fixed-price contract where initially, buyer and seller agree to trade the good at the preset price. If the seller invests less than $\bar{e}$, both agents renegotiate this contract ex post to prevent undesirable trade. The associated bargaining surplus strictly decreases in the seller’s investments, because larger investments reduce the losses from trade. This negative investment effect plays out against the positive effect of investments on the seller’s default payoff, her payoff in absence of renegotiation. Taken together, these considerations yield a local investment optimum in which the seller invests a positive amount but strictly less than $\bar{e}$, the level required to make trade worthwhile. Overall, the seller then faces the alternatives to undertake second-best investment $e^*$ followed by trade, or to shirk and renegotiate to a no trade outcome. We find that when the seller’s investments affect the buyer’s valuation to some degree and when the seller’s bargaining power is not too small, she will often prefer the shirking option. In a parametric specification discussed in the main text, for example, shirking arises over almost a third of the feasible parameter space.

To generalize this result, we further demonstrate that if strategic shirking occurs under a fixed-price contract, there exists no general revelation mechanism (Maskin and Moore, 1999) under which trade can be an equilibrium outcome. This result extends findings in Che and Hausch (1999) to a new class of situations in which long-term contracting is useless.\(^5\)

Our results provide a novel argument of why rational agents are sometimes bound for a failure of their relationship. While in the existing literature, non-internalized investment externalities cause agents to underinvest, they never do so in order to make trade non-viable. In contrast, our model demonstrates that rational agents may shirk in order to

\(^{5}\)Note that in contrast to Che and Hausch, undesirability of contracting always renders trade impossible in the present setting.
rescind an existing trade contract, while extracting money in the renegotiation process. In other words, agents have an incentive to destroy the value of the relationship for personal gains.\textsuperscript{6}

The paper analyzes two variants of this basic setting, which are distinguished by the timing of trade. In a first variant the final trade decision is taken at a time where both agents face uncertainty about the buyer’s valuation, while they already received this information in the second variant. While strategic shirking occurs in both scenarios, the scope of the shirking problem turns out to be more severe in the latter case. This result has an intuitive explanation: when the parties trade in a situation of uncertainty, they have to rely on expected gains; hence, shirking prevents trade even in states where it would actually be beneficial to both agents. The presence of these states reduces the negotiation gain from shutting down trade, and diminishes the seller’s incentive to shirk in the first place. In economic environments in which the actual date of trade is a choice variable, the trade partners may hence want to commit themselves to ‘trade too early’, in a situation of uncertainty, in order to mitigate strategic shirking.

While the basic model only allows the seller to invest, we briefly extend the analysis to the more general case of two-sided investments. Specifically, we focus on a scenario where the buyer has no independent motive to shirk: if the seller exerts the proper investment, the buyer will always do the same.\textsuperscript{7} Shirking equilibria are shown to exist in this extended setting as well. In particular, and perhaps surprisingly, the buyer will never overinvest in order to ‘save the day’ and to lessen the seller’s shirking incentive. We also find that

\textsuperscript{6}An important difference to the literature on sabotage is that agents in the present paper do not undertake effort to reduce the observable performance of other agents in a contest. For an overview of this literature, see Konrad (2000). Schmitz (2002) investigates a similar indivisible trade model in which the seller invests into the buyer’s valuation, which is this agent’s private information. The interplay of moral hazard and hidden information renders the first best infeasible even if the trading problem is trivial, i.e., if ex-post trade is efficient regardless of investments.

\textsuperscript{7}This is the case if the buyer’s investments are ‘purely selfish’, that is, raise his valuation of the trading good while not affecting the seller’s costs.
shirking and non-shirking equilibria now co-exist. An agent shirks when she expects the other agent to shirk, and vice versa. These beliefs are self-enforcing and lead to multiple equilibria with opposing outcomes.

The paper proceeds as follows. Section 2 sets up the basic model. Section 3 discusses a simple example, while Section 4 provides a general analysis of the baseline scenario where the parties trade before learning the buyer’s valuation. Section 5 explores an alternative scenario where trade occurs after the buyer’s valuation has been realized. Section 6 studies two-sided investments of both agents. Section 7 briefly concludes.

2 The Model

Two risk-neutral agents, the seller $S$ and the buyer $B$, can trade one unit of an indivisible good. Before trade occurs, the seller expends relationship-specific investments $e \in \mathbb{R}_0^+$, henceforth called effort. Effort decreases the seller’s production costs $c(e)$, and increases the buyer’s valuation $v(e, s)$. The valuation is subject to a stochastic shock $s \in [\underline{s}, \bar{s}]$ which is distributed according to a continuous distribution function $F(s)$ and realized after effort was expended.

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

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

where the function $\psi(e)$ represents the seller’s costs of undertaking effort. Denote by $V(e) = E_s v(e, s)$ the buyer’s expected valuation for given investments. We impose standard assumptions on $v(e, s), c(e)$ and $\psi(e)$. These functions are twice continuously differentiable in $e$, $v(e, s)$ is increasing in $s$ and $e$ and concave in investments, $c(e)$ is decreasing and concave, and $\psi(e)$ is increasing and convex in investments.\(^8\) We also assume that

\(^8\)To ensure interior solutions, we additionally require that either $v(\cdot)$ or $-c(\cdot)$ are strictly concave
$v(e, s)$ is separable between $e$ and $s$, so $v_{es} = 0$. The effort level $e$ and the state of the world $s$ are perfectly observed by either party, but non-verifiable.

The sequence of events is as follows. Both agents sign an initial contract at date 0. While mostly focusing on non-contingent fixed-price contracts, we will also explore general mechanisms (Maskin, 1999; Maskin and Moore, 1999) which allow truthful revelation of information received by buyer and seller over the course of the game. Subsequently at date 1, the seller undertakes her investment $e$ before, in a revelation mechanism, the parties send messages to an external enforcer at date 2. They are then free to renegotiate the terms of contract at date 3. For concreteness, the outcome of renegotiation is described by the generalized Nash-bargaining solution, which parameterizes the seller’s bargaining power as $\gamma \in [0, 1]$. Finally, buyer and seller can carry out physical trade at date 4 before the state of the world $s$ is revealed. This means the commodity is an ‘experience good’ whose actual valuation is learned only after the purchase. The alternative scenario in which the agents trade under certainty, that is, after learning $s$ will be analyzed in Section 5 below.

Figure 1 below illustrates the timing.

---

9 This assumption is inconsequential but allows to simplify the exposition considerably.

10 In a general mechanism, both agents submit messages to an enforcer after the seller chose her effort level. These messages can be statements about an agreement to trade. For example, under a buyer’s option, the buyer could announce whether he exercises his option or not.

11 The model can also be interpreted as one in which B’s valuation materializes only after she transformed the good after trade. For example, B might use the good as in input in her own production process, or he might be a retailer who undertakes some marketing effort with stochastic returns.
A first best outcome entails efficient trade (ex-post efficiency) $q^{FB}$ and efficient relationship specific investments (ex-ante efficiency) $e^{FB}$. Efficient trade requires the trade decision $q \in [0, 1]$ to maximize

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

(2)

Note that $q^{FB} = 1$ when $V(e) \geq c(e)$, and $q^{FB} = 0$ otherwise. Hence, we can define

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

(3)

as the unique threshold effort below which final trade should not be completed.

Define $S(e) = \phi(e) - \psi(e)$ as the expected total surplus for given $e$. Our assumptions on functional forms ensure $S(e)$ to be strictly concave if $\bar{e} = 0$. In contrast, $S(e)$ cannot be globally concave in a setting with positive threshold investment $\bar{e} > 0$: for any $e < \bar{e}$, $\phi(\cdot) = 0$ so that $S(\cdot)$ is decreasing in $e$, while $S(\cdot) > 0$ in a nonempty range of investments $e \geq \bar{e}$. We should note that in this latter case, the global-concavity assumption in Che and Hausch (1999) is violated (Assumption 4 in their paper).

When positive, the first-best investments $e^{FB}$ are implicitly defined by the first-order condition (subscripts denote derivatives)

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

(4)

We will say that $S$’s investments are purely selfish whenever $c_e(\cdot) < 0$ and $v_e(\cdot) = 0$, whereas investments are purely cooperative when $c_e = 0$ and $v_e(\cdot) > 0$ for all $e$. It is also useful to introduce

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

7
as a measure for the effect of investments. The measure \( \hat{\gamma} \) rises when investments become less cooperative. It takes on a value of one when effort only affects costs (purely selfish investments), and it is zero when the effort only affects valuation (purely cooperative investments). Investments will be said to be primarily selfish if the seller’s exogenous bargaining strength is some \( \gamma \leq \hat{\gamma} \), while they are primarily cooperative if \( \gamma > \hat{\gamma} \). Notice that if \( \gamma = \frac{1}{2} \) so that both agents have the same bargaining strength, investments are primarily selfish if \( -c_e > V_e \), while otherwise they are primarily cooperative.

To relate setup and results to those in the literature, look at the benchmark case \( \bar{e} = 0 \) in which trade is always ex-post efficient. The following Proposition is in the spirit of well-known results in Edlin and Reichelstein (1996) and Che and Hausch (1999):

**Proposition 1.** Let \( \bar{e} = 0 \). A first-best is attained with purely selfish investments, or if \( \gamma = 1 \). Otherwise, trade prevails but the seller underinvests. More specifically,

1. If investments are primarily selfish (\( \gamma \leq \hat{\gamma} \)), the second-best effort \( e^* (< e^{FB}) \) solves
   \[
   -c_e(e^*) = \psi_e(e^*),
   \]
   and is implemented by a fixed-price contract \( (t, q = 1) \).

2. If investments are primarily cooperative (\( \gamma > \hat{\gamma} \)), second-best investments \( e^{**} (< e^{FB}) \) solve
   \[
   \gamma [V_e(e^{**}) - c_e(e^{**})] = \psi_e(e^{**}),
   \]
   and the optimal contract is a no-trade contract \( (t, q = 0) \).

**Proof:** Follows from results in the literature. Available upon request.

---

12The parameter \( \hat{\gamma} \) corresponds to the measure \( \alpha \) in Che and Hausch (1999).
13Note that, for a given tuple \( (e, s) \), \( \gamma \nu_e(e, s) + (1 - \gamma)c_e(e) \) is an increasing function of \( \gamma \).
14While these papers consider variable quantities, we focus on binary trade. The only significant difference in results arises for primarily selfish investments where the first best might be achievable in the variable-quantity framework while it is unachievable in the current setting. Intuitively, with a variable trade volume a fixed-price contract can prescribe a quantity in excess of the efficient level, which stimulates investment incentives.
When investments are primarily selfish, the optimal ex-ante trade arrangement is a simple fixed-price contract which obliges the seller to deliver the good in exchange for a monetary transfer \( t \). Under such a forcing contract, the seller’s utility function becomes \( U^S = t - c(e) - \psi(e) \). \( S \) reaps the full marginal return from her investments in cost reduction, while \( B \) appropriates the cross effect of investments which increase his valuation. Because of this externality, equilibrium investments reach the first-best level \( e^{FB} \) when they are purely selfish, while otherwise the seller will underinvest, and more so the larger the cross effect. Conversely, when investments are primarily cooperative, equilibrium effort peaks when the parties do not sign any initial agreement, and instead rely on spot contracting after effort has been expended. The seller then obtains a fraction \( \gamma \) of the overall return from the relationship, and invests a positive amount \( e^{**} \) that is increasing in her bargaining strength \( \gamma \).

Note that in this standard framework with zero investment threshold, strategic shirking does not arise because even the lowest effort does not allow the seller to block equilibrium trade. With a positive investment threshold \( \bar{e} > 0 \), strategic shirking remains of no concern when the parties refrain from an initial trade arrangement, and rely on spot contracting instead. If the seller shirks by investing less than \( \bar{e} \), the buyer simply refuses to trade, and renegotiation does not arise. Clearly, the seller cannot benefit from such a strategy. Since spot contracting is the best arrangement when investments are primarily cooperative, strategic shirking requires the seller’s effort to be primarily selfish, \( \gamma \leq \hat{\gamma} \), the scenario on which we focus from now on. Throughout our discussion, we assume the relationship to be viable when the seller chooses the second best effort \( e^* \). Hence, \( S(e^*) > 0 \) which immediately implies \( e^* > \bar{e} \).
3 Example

Before presenting the general analysis, it may be useful to solve a simple linear-quadratic example. Suppose the buyer’s expected valuation is $V(e) = v + e$ and the seller’s production costs are $C(e) = f - (3/2)e$. With investment costs $\psi(e) = e^2/2$, first best investments become $e^{FB} = V_e - C_e = 5/2$. Note that for any $\gamma < \hat{\gamma} = 3/5$, investments are primarily selfish. In the standard model ($v \geq f$ so that $\bar{e} = 0$), any fixed price trade contract $t$ implements the second best effort, computed as $e^* = -C_e = 3/2$, by Proposition 1.

Next, assume now $v < f$, which implies an investment threshold $\bar{e} > 0$ below which final trade becomes inefficient.\footnote{To be precise, $\bar{e}$ satisfies $V(\bar{e}) - C(\bar{e}) = 0$ so that in the current example, $\bar{e} = (2/5)(f - v)$.} In this case, the relationship remains viable (i.e., $S(e^*) > 0$) at second-best investments $e^*$ whenever

$$f - v < (5/2)e^* - (e^*)^2/2 = 21/8.$$ Even if this condition applies, the seller may not invest $e^*$ under a fixed price contract $t$ but instead, defect and choose a shirking effort $e_{sh}$ below $\bar{e}$. To see why, note that after shirking occurred and was observed by the buyer at date 1, both parties have an interest to renegotiate and rescind their trade agreement at date 3. With Nash bargaining, these ex-post negotiations allow the seller to appropriate a share $\gamma$ of the associated bargaining surplus, which is $0 - [V(e) - C(e)] = C(e) - V(e) = f - v - (5/2)e > 0$. One can now verify that the seller prefers shirking when for some $e < \bar{e}$, her associated utility

$$t - [f - 3/2e] + \gamma[f - v - 5/2e] - 1/2 e^2$$

exceeds $t - [f - (3/2)e^*] - (e^*)^2/2$, her payoff when she exerts the second best effort $e^*$ and trades the good. With an optimal shirking effort $e_{sh} = 3/2 - (5/2)\gamma$ (this effort level maximizes (5) for any $\gamma \leq \hat{\gamma}$), a payoff comparison reveals that the seller will adopt a
shirking strategy iff
\[ f - v > \frac{9}{8\gamma} - \frac{1}{2} \left( \frac{3}{2} - \frac{5}{2}\gamma \right)^2. \]
The shirking range of parameters is non-empty for any \( \gamma > 2/5 \). If \( \gamma = 1/2 \), the seller shirks when \( v - f < -35/16 \) and hence, over 16 per cent of the admissible parameter range \( v - f \in [-21/8, 0] \). With \( \gamma \) close to \( 3/5 \), shirking occurs if \( v - f < -45/24 \) and hence, for almost 30 per cent of admissible parameters. Strategic shirking thus poses a significant threat for the success of the bilateral relationship.

4 Minimum Effort and Strategic Shirking

This Section generalizes the previous result to arbitrary cost and value functions, and it allows for general contractual arrangements. In contrast to the workhorse setting analyzed in the literature and discussed in Proposition 1, we now allow for a positive threshold investment that is needed to make trade viable. Let \( \bar{e} \geq 0 \), suppose that investments are primarily selfish, and first reconsider a date-0 fixed-price contract \((q = 1, t)\) which can be renegotiated at date 3. For any \( e < \bar{e}(< e^*) \) the joint continuation surplus \( \phi(e) \) is zero because the agents will renegotiate any non-viable trade agreement. Specifically, the seller’s date-1 utility is described by the piecewise defined function
\[
U^S(e) = \begin{cases} 
  t - c(e) - \psi(e) & \text{for } e \geq \bar{e} \\
  t - c(e) + \gamma[0 - (V(e) - c(e))] - \psi(e) & \text{for } e < \bar{e}.
\end{cases}
\]
Note that if \( S \) chooses an effort \( e < \bar{e} \) at stage 1, the parties will at date 3 rescind the initial contract. In this bargaining process, the seller receives a fraction \( \gamma \) of the renegotiation gain. This renegotiation gain is the difference between the (zero) joint surplus when no trade occurs, and the (negative) joint surplus when the original contract was to be executed.

This payoff function has two local optima. The second-best effort \( e^* \) remains the seller’s optimal choice among all trade-inducing effort levels \( e \geq \bar{e} \). Within the complementary
subset \( e < \bar{e} \), her local optimum is some shirking effort \( e_{sh} \in [0, \bar{e}) \). Overall, the seller’s global optimum under a fixed-price contract remains \( e^* \) if and only if

\[
t - c(e^*) - \psi(e^*) \geq t - (1 - \gamma)c(e_{sh}) - \gamma V(e_{sh}) - \psi(e_{sh}).
\]

(7)

Note first that condition (7) is insensitive with respect to the preset trade price. Hence, strategic shirking cannot be avoided by modifying the terms of trade. A full analysis shows

**Proposition 2.** Consider a fixed-price contract, and let \( \bar{e} > 0 \).

a) If investments are purely selfish, the seller invests \( e^* = e^{FB} \).

b) Let investments be primarily selfish, and define

\[
\gamma_{sh} := \frac{[c(e_{sh}) - c(e^*)] - [\psi(e^*) - \psi(e_{sh})]}{c(e_{sh}) - V(e_{sh})} (> 0).
\]

(8)

The seller shirks and chooses \( e_{sh} < \bar{e} \) if her bargaining parameter is some \( \gamma \in [\gamma_{sh}, \hat{\gamma}) \) with \( \gamma_{sh} < \hat{\gamma} \). Agents then renegotiate and do not trade in equilibrium. For \( \gamma \leq \gamma_{sh} \), the seller exerts the second best effort level \( e^* \).

Proof: First-best investment \( e^{FB} \) are implementable if investments are purely selfish, i.e., \( V(e) \equiv V \forall e \). To see this, notice that (7) holds iff

\[
c(e^*) + \psi(e^*) < (1 - \gamma)c(e) + \gamma V(e) + \psi(e) \quad \forall e < \bar{e}.
\]

(9)

Since \( S(e^*) > 0 \), we have \( c(e^*) + \psi(e^*) < V < c(e) \) for all \( e < \bar{e} \) and the result follows. For primarily selfish investments, check that (7) is violated iff \( \gamma > \gamma_{sh} [> 0] \). Accordingly

\[16\]The shirking effort maximizes \( t - (1 - \gamma)c(e) - \gamma V(e) - \psi(e) \). Notice that \( e_{sh} \) is always positive (for primarily selfish investments), falls in \( \gamma \), and converges to zero for \( \gamma \to \hat{\gamma} \).

\[17\]When (7) is violated so that the seller prefers shirking, the associated optimal shirking effort \( e_{sh} = \arg\max_e \ t - c(e) - \gamma[c(e) - V(e)] - \psi(e) \) always satisfies \( e_{sh} < \bar{e} \) as required. To see why, notice that by definition of \( \bar{e} \), the renegotiation gain \( c(e) - V(e) \) is positive only if \( e < \bar{e} \). For \( e \geq \bar{e} \), the seller’s (now hypothetical) shirking payoff would therefore be bounded by \( t - c(e) - \psi(e) \), and hence, be smaller than his payoff from not shirking and choosing \( e^* \) instead.

\[18\]The threshold parameter \( \gamma_{sh} \) is strictly positive because for \( \gamma \to 0 \), the seller’s bargaining surplus converges to zero and shirking is dominated. Formally, \( c(e^*) + \psi(e^*) < c(e_{sh}) + \psi(e_{sh}) \) by definition of \( e^* \), and \( c(e) - V(e) > 0 \) for all \( e < \bar{e} \).
shirking $e < \bar{e}$ under a fixed-price contract arises when the seller’s bargaining parameter \( \gamma \) is (1) larger than \( \gamma_{sh} \), and (2) smaller than \( \hat{\gamma} \). \( \square \)

Shirking arises when defective behavior is in the seller’s best interest. On the one hand, this requires her bargaining share in renegotiations to be sufficiently large. At the same time, the seller’s bargaining share cannot be arbitrarily big because investments also need to be primarily selfish – with primarily cooperative investments, the agents should not write a long term contract (Proposition 1). Shirking thus arises over some intermediate range of bargaining parameters, as characterized in the Proposition. It is also worth emphasizing that shirking requires some element of investment externalities, for the following reason. With second-best investments, the buyer’s gross valuation always exceeds the sum of production and investment costs (otherwise, the relationship would not be viable). If she exerts these investments, the seller’s payoff is \( t \) minus the sum of these two costs. Conversely, if she shirks, her payoff is always smaller than \( t \) minus the buyer’s gross valuation evaluated at the shirking effort. If investments are purely selfish so that the buyer’s valuation does not vary in \( e \), the latter payoff can never dominate. In contrast, when investments have a cooperative effect, effort raises the buyer’s valuation. His gross valuation evaluated at the shirking effort can then be smaller than the sum of costs evaluated at efficient investments, and shirking becomes more attractive.

The next step to ask is whether some more refined contractual arrangement is able to implement \( e^* \), or at least an effort level large enough to support subsequent trade. For a general revelation mechanism (Maskin-Moore, 1999), the answer is given in

**Proposition 3.** Let investments be primarily selfish, i.e. \( \gamma < \hat{\gamma} \). If a fixed-price contract cannot implement \( e^* \), there exists no general mechanism which renders trade feasible.

**Proof:** By the revelation principle, we can restrict attention to a direct revelation mecha-
nism 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
\[
 u^S(e_2) \geq t(e_2, e_1) - \delta(e_2, e_1)[c(e_2) + \gamma(V(e_2) - c(e_2))]. \tag{10}
\]
Next, assume that \( S \) expended \( e_1 \). Then, \( B \) announces truthfully iff
\[
 u^B(e_1) \geq -t(e_2, e_1) + \delta(e_2, e_1)V(e_1) + (1 - \delta(e_2, e_1))(1 - \gamma)[V(e_1) - c(e_1)]. \tag{11}
\]
Since the possibility of renegotiation ensures that \( u^S(e_1) + u^B(e_1) = V(e_1) - c(e_1) \), condition (11) translates into
\[
 u^S(e_1) \leq t(e_2, e_1) + \gamma(1 - \delta(e_2, e_1))V(e_1) - [\delta(e_2, e_1) + (1 - \delta(e_2, e_1)\gamma)c(e_1)]. \tag{12}
\]
Combining (10) and (11) yields
\[
 u^S(e_1) - u^S(e_2) \leq \gamma(1 - \delta(e_2, e_1))V(e_1) - [\delta(e_2, e_1) + (1 - \delta(e_2, e_1)\gamma)c(e_1)
+ \delta(e_2, e_1)[(1 - \gamma)c(e_2) + \gamma(V(e_2))] \equiv \xi(e_1, e_2, \delta(\cdot)). \tag{13}
\]
Fix \((e_1, e_2)\) and consider the derivative of \( \xi(\cdot) \) with respect to \( \delta \),
\[
 \frac{d\xi(\cdot)}{d\delta(\cdot)} = (1 - \gamma)[c(e_2) - c(e_1)] - \gamma[V(e_1) - V(e_2)]. \tag{14}
\]
This derivative is positive iff investments are primarily selfish, i.e., \( \gamma < \hat{\gamma} \).

Accordingly, \( \delta(\cdot) = 1 \) implements the seller’s maximum payoff difference \( \Delta u^S(e_1, e_2, \delta) \equiv \xi(e_1, e_2, \delta) - [\psi(e_1) - \psi(e_2)] \) between any two investment levels characterized by \( e_1 > \bar{e} > e_2 \).
Notice that \( e^* \) can be implemented only if \( \Delta u^S(e^*, e_2, 1) \geq 0 \) for all \( e_2 < \bar{e} \), and consider \( e_2 = e_{sh} < \bar{e} \). Verify that the seller’s payoff difference \( \Delta u^S(e^*, e_{sh}, 1) \) between \( e^* \) and \( e_{sh} \) is then equivalent to the corresponding payoff difference under a fixed-price contract.\(^{19}\) Consequently, a revelation mechanism can implement \( e^* \) if and only if a fixed-price arrangement does. Finally, consider a situation where a fixed-price contract leads to defection, i.e., \( \Delta u^S(e^*, e_{sh}, 1) < 0 \). Since \( e_1 = e^* \) maximizes \( \Delta u^S(e_1, e_2, 1) = [c(e_2) - c(e_1)] + \gamma[V(e_2) - c(e_2)] - [\psi(e_1) - \psi(e_2)] \) for arbitrary \( e_2 \), we also have \( 0 > \Delta u^S(e^*, e_{sh}, 1) \geq \Delta u^S(e_1, e_{sh}, \delta) \) for all \( e_1 \geq \bar{e}, \delta \in [0, 1] \), which completes the proof. \( \square \)

Proposition 3 delivers a strong message. If a fixed price contract leads to strategic shirking, the same holds true with any general contract. At an intuitive level, fixed price contracts maximize the seller’s payoff difference between any ‘shirking’ and any ‘non-shirking’ effort. If this payoff difference is negative in a fixed price contract, it cannot be positive under any alternative contractual arrangement, including the no-contract. Hence, fully rational agents who are aware of the shirking problem will find it advisable not to start their relationship from the outset.

5 Trade under Certainty

We will now show shirking to become an even more serious threat when the buyer’s valuation is already known at the date of trade, i.e., when the state \( s \) is realized prior to date 4.

In this case, the maximum joint surplus for given \((e, s)\) is \( \phi^r(e, s) = \max_{q \in \{0,1\}} q[v(e, s) - c(e)] \phi^r(e, s) \)

\(^{19}\)Suppose instead that investments are primarily cooperative in which case the derivative in (14) is negative, and \( \delta() = 0 \) generates maximum effort incentives. In this case, spot contracts are optimal whenever \( \Delta u^S(e_1 = e^{**}, e_2 = 0, 0) = \gamma[V(e^{**}) - c(e^{**})] - \psi(e^{**}) > 0 \). Otherwise, even the largest possible payoff difference is too small to induce equilibrium trade, and the relationship has no value. In reverse, the outcome implies that if (investments are primarily selfish and) a spot contract implements \( e^{**} \), strategic shirking does not occur under a fixed price contract. In other words, shirking does not arise if \( \gamma[V(e^{**}) - c(e^{**})] - \psi(e^{**}) \geq 0 \).
To analyze this alternative scenario, it is useful to introduce some further definitions. Let 
\[ \hat{s}(e) = \sup \{ s \mid v(e, s) - c(e) \leq 0 \} \]
be a threshold state so that trade becomes ex-post efficient for all \( s \geq \hat{s}(e). \) Also, denote \( q(e) = 1 - F(\hat{s}(e)) \) as the (equilibrium) trade probability for given investment level. Let
\[ \tilde{e} \equiv \inf \{ e \mid v(e, \underline{s}) - c(e) \geq 0 \} \tag{15} \]
be the threshold investment above which trade is ex-post efficient even in the least favorable state of the world, \( \underline{s}. \) Notice that \( \tilde{e} \geq \bar{e} \) as defined in (3) because, at an effort level \( \bar{e} \) where the expected gross surplus breaks even, the actual surplus is still negative in state \( \underline{s}. \) The expected net surplus in this setting is \( S^c(e) = \int_{s} \phi^c(e, s) dF(s) - \psi(e) \) and again strictly concave if \( \bar{e} = 0. \) In contrast, for \( \tilde{e} > 0, \) trade becomes inefficient in low states for small investments, with the consequence \( q(0) < 1. \) This latter case will be the focus of the subsequent analysis.

Finally, to make the setting comparable to the analysis in the base setting, suppose that \( S^c(e^*) > S^c(e) \) for all \( e < e^*, \) and \( q(e^*) = 1 \) so that trade is efficient in every state if \( e^* \) can be implemented. In absence of a strategic shirking motive, a fixed-price contract \( t \) would then again induce the second-best efficient effort \( e^* \) 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 value of the good for given 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
\[ U^S(e) = t - c(e) + [1 - q(e)] \gamma_c(e) - V^-(e) - \psi(e), \tag{16} \]

\[ \text{Since higher investments increase the likelihood of efficient trade for any } e < \hat{e}, \text{ global concavity of } S^c(e) \text{ is then not guaranteed even if the functions } v(\cdot), c(\cdot) \text{ and } \psi(\cdot) \text{ are well behaved.} \]
\[ \text{Note that for } q(e^*) < 1, \text{ the second-best effort level would be smaller than } e^* \text{ and implicitly determined by } -q(e)c_e(e) = \psi(e). \]
where the expression in brackets represents the expected gain from renegotiation if trade is ex-post inefficient, and $[1 - q(e)]$ denotes the likelihood of this event for given $e$. Since $q(0) < 1$, this payoff function is not necessarily concave.\footnote{Observe that $e^*$ is the local maximum in the range $e \geq \tilde{e}$ and also constitutes the global optimum if $U^S(e)$ is strictly concave.} In particular, the seller chooses $e^*$ if and only if

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

We now show that condition (17) holds less often than the corresponding condition (7) in our baseline scenario. In other words, the seller shirks and induces a breakdown of trade more frequently than she does in a setting where trade occurs under uncertainty of the state of the world. To validate this claim, one can confine attention to a subset $e < \tilde{e}$ of possible deviations: since $\tilde{e} \geq \bar{e}$, the range of possible defections is weakly larger in the present scenario, as compared to the baseline setting. Note that seller’s expected utility for $e = e^*$ coincides in either setting for given $t$. Now consider a deviation $e \leq \tilde{e}$. Then, the seller’s payoff from shirking when trade occurs under certainty exceeds her shirking payoff in the baseline scenario if, by (6) and (16),

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

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 $e < \tilde{e}$ and $q(e) > 0$ by the definition of $V^+(\cdot)$. In terms of bargaining parameters and recalling that $e_{sh}$ indicates $S$’s optimal shirking effort in the uncertainty case, the seller will now shirk at least if

$$\gamma > \gamma^c =: \frac{[c(e_{sh}) - c(e^*)] - [\psi(e_{sh}) - \psi(e^*)]}{c(e_{sh}) - V(e_{sh}) + q(e_{sh})[V^+(e_{sh}) - c(e_{sh})]}, \quad (19)$$

A comparison with (8) reveals that $\gamma^c \leq \gamma_{sh}$, and that the strict inequality $\gamma^c < \gamma_{sh}$ applies whenever $q(e_{sh}) > 0$. Along the lines of the proof of Proposition 3, one can also
show that no other more general mechanism can improve upon the outcome of fixed-price contracting. Accordingly, we can state

**Proposition 4.** Let investments be primarily selfish, and suppose the agents trade after learning state $s$. If trade is always efficient for the second-best investment $e^*$, implementing $e^*$ is impossible if the seller’s bargaining parameter is some $\gamma \in (\gamma_c, \hat{\gamma})$. Since $\gamma_c \leq \gamma_{sh}$ (with strict inequality at least if $q(e_{sh}) > 0$), strategic shirking occurs over a wider range of parameters in the certainty scenario.

To understand the economic reason behind these findings, consider a situation where the seller shirks when the parties trade under uncertainty. They then renegotiate in a position of uncertainty about the true state $s$ and a no-trade outcome prevails in every state, including those states where trading would actually have been beneficial. Hence, the existence of these states reduces the available bargaining surplus in renegotiations, which diminishes the seller’s incentives to shirk in the first place. Clearly, no similar disincentive arises in a scenario where trade and renegotiation take place under certainty. As a consequence, shirking becomes more attractive because the additional information in the latter regime avoids the loss of renegotiation surplus, and encourages strategic shirking that prevents trade in a subset of states.

To illustrate this using the example in Section 3, suppose that the buyer’s valuation (with expectation $v$) materializes as $\bar{v}$ or $v$ with equal probability. Let $\bar{v} > f > v$ so that $q(0) = 1/2 > 0$. Also, let $f - (5/2)e^* > v < f - (5/2)e_{sh}$ for $e^* = 3/2$ and the shirking effort $e_{sh} = 3/2 - (5/2)\gamma$ from the certainty case, respectively. Trade in the low state $v$ is then inefficient if the seller exerts $e_{sh}$, while it is efficient if she exerts $e^*$ instead. Suppose the seller chooses $e_{sh}$. With a fixed price contract, both agents then renegotiate after observing the low state, and the seller’s shirking utility becomes $t - f + (3/2)e_{sh} + (1/2)\gamma [f - v - (5/2)e_{sh}] - e_{sh}^2/2$. This shirking payoff strictly exceeds the
corresponding payoff (5) in the certainty case (note that $v = \bar{v}/2 + \underline{v}/2$). Recalling that the seller’s payoff from the second best effort $e^*$ coincides in both settings, this means she will shirk over a wider range of parameters when trade occurs after the state of the world has been revealed.

Agents in a trade relationship may thus benefit from finalizing their trade transaction in a state of uncertainty about the buyer’s valuation for the good. While trading under uncertainty comes at the cost that some undesirable trades are carried out in low valuation states, it makes the relationship less vulnerable to the possibility of strategic shirking. Notice that in many environments, trading ‘early’ or ‘late’ is not at the discretion of the agents, but rather a technological characteristic of the item to be traded. Specifically, the buyer’s valuation of an experience good is unveiled only after the purchase, which according to our results increases the risk of seller shirking relative to the case of a ‘search good’ whose value is already learned before the trade transaction takes place.

6 Two-sided Investments

This Section extends our setting to allow not only the seller, but also the buyer to invest into the joint relationship. For concreteness, we will concentrate on a situation where the buyer’s investment is completely selfish, which means, affects only his own benefit from trade. This case not only deserves attention from an empirical perspective, but provides

---

23 Remember that her best shirking effort under uncertainty will exceed $e_{sh}$, so that considering $e_{sh}$ underestimates the true shirking potential.

24 In some circumstances, the timing of the transaction before or after the resolution of uncertainty may be a choice variable. While our results suggest that agents then may want to commit to ‘early’ trade, it is an open question whether such an ex-ante commitment can be made credible ex post. After all, agents would have an incentive to renegotiate the timing of trade after investments have been exerted.

25 When the buyer’s investments are partially cooperative (they affect not only his valuation but also the seller’s costs), his preferences mirror those of the seller, so that he may have an independent interest to shirk.
a useful benchmark for studying some novel theoretical aspects: as shown before for one-sided investments, with purely selfish investments an agent lacks a shirking motive. In the present context with two-sided investments, one might even conjecture that the buyer sometimes overinvests, in order to offset a shirking seller and to prevent a breakdown of trade.

To fix ideas and ensure comparability, we define seller’s investments in the same way as in the previous Sections. Suppose that the buyer can now also undertake an idiosyncratic investment \( b \geq 0 \) to raise his expected valuation \( V(e, b) \) from trade. Let \( V(\cdot) \) be increasing and concave in \( b \), and denote the buyer’s personal and convex investment costs as \( \phi(b) \). To simplify the exposition, we impose the otherwise inconsequential assumption \( V_{eb} = 0 \) so that the marginal investment returns of both agents are independent. Define production costs \( c(e) \) as before and again, assume that the seller’s investments are primarily selfish (\( \gamma \leq \hat{\gamma} \)) so that fixed-price trade contracts are second-best efficient. Then, the buyer’s optimal investment (if trade has value) \( b^* \) is described by the first order condition \( V_b(e, b^*) = \phi_b(b^*) \), and the seller’s second best optimal investment \( e^* \) by the condition \( \psi_e(e^*) = -c_e(e^*) + V_e(e^*, \cdot) \).

In what follows, we again assume that the relationship has value at investment levels \( (e, b) \), that is, \( S(e, b) > 0 \). Moreover, the analysis focuses on fixed price contracts.

As already said, the buyer has no independent shirking motive for shirking in this setting. Since his investments are purely selfish, \( B \)’s best response to \( e^* \) is to exert the optimal investment \( b^* \) himself.\(^{27}\) Therefore, situations of interest are again those in which seller \( S \) might shirk. For concreteness, define \( \bar{e}(b) \) as the investment threshold \( \bar{e} \) below which

\(^{26}\)Note that since \( V_{eb} = 0 \), \( e^* \) and \( \hat{\gamma} \) are defined as in the baseline model of Section 3. Note also that because \( B \)’s investments are purely selfish his optimal investment coincides with the first best level, \( b^* = b^{FB} \).

\(^{27}\)The argument is completely symmetric to the one laid out in the proof of Proposition 2. Intuitively, when \( B \)’s investments are purely selfish, underinvestment does not cause a negative externality for agent \( S \), so that shirking cannot be beneficial.
the continuation value of trade is negative. This investment level is implicitly defined by 
\[ V(\bar{e}(b),b) = 0, \]
and decreases in the buyer’s investment. Define \( \bar{e} \equiv \bar{e}(b^*) \) and assume \( \bar{e} > 0 \). Similarly, let \( \bar{b}(e) \) as implicitly defined by 
\[ V(e, \bar{b}(e)) = 0 \]
be the buyer’s minimum investment needed to ensure a positive continuation value of trade.\(^{28}\)

We start the analysis stating a useful preliminary result. As already said, one may be tempted to believe that the buyer may sometimes *overinvest* to help trade proceed. This belief is refuted in

**Lemma 1** The buyer’s best response to the seller’s investment choice \( e \geq 0 \) is some \( b \leq b^* \).

Specifically, he will choose a defective investment level \( b_{sh}(< b^* < \bar{b}(e)) \) when \( e < \bar{e} \).

**Proof:** Consider first some \( e \geq \bar{e} \). Since \( b = b^* \) triggers trade and the buyer’s investment is purely selfish, his best response is \( b^* \) in analogy to the arguments in the proof of Proposition 2. Second, consider seller investments in the range \( e < \bar{e} \). Notice that by definition of \( \bar{e} \), a buyer investment of size \( b \) (or less) then triggers renegotiation and no trade. Specifically, some \( b \geq \bar{b}(e)(> b^* \text{ for } e < \bar{e}) \) is needed to make subsequent trade viable. By choosing an investment from the range \( b \geq \bar{b}(e) \), \( B \) can accrue a payoff of 
\[ V(b,e) - t - \phi(b). \]
The maximizer of this payoff function \( b^* \) (recall \( V_{eb} = 0 \)) lies outside the considered investment range because \( b^* < \bar{b}(e) \). Since \( B \)'s payoff over the range \( b \geq \bar{b}(e) \) is continuous and decreasing in \( b \), the local optimum is found at the interval boundary, \( b = \bar{b}(e) \). Conversely, if \( B \) undertakes an investment \( b < \bar{b}(e) \) when \( e < \bar{e} \), buyer and seller renegotiate to a no-trade outcome. In this case, \( B \)'s payoff with renegotiation becomes

\[ V(b,e) - t + (1 - \gamma)[c(e) - V(b,e)] - \phi(b). \quad (20) \]

Denote the maximizer of this function as \( B \)'s defective effort \( b_{sh} = \arg\max \gamma V(e,b) - \phi(b) \).

This implies the interior maximum \( b_{sh} \leq b^* < \bar{b}(e) \) (and \( b_{sh} < b^* \) unless \( \gamma = 1 \)). Note that

\(^{28}\)We have \( \bar{b}(e) = b^* \) by definition of \( \bar{e} \). Also, note that \( \bar{b}(e^*) < b^* \) because \( S(e^*,b^*) > 0 \).
the buyer’s local payoff is falling for any \( b > b_{sh} \), and his payoff function is continuous at the boundary \( \bar{b}(e) \) because the renegotiation gain \( [c(e) - V(b, e)] \) converges to zero for \( b \to \bar{b}(e) \). Recalling that the boundary investment \( \bar{b}(e) \) constitutes his local optimum over the complementary range \( b \geq \bar{b}(e) \), B’s globally optimal investment for \( e < \bar{e} \) is \( b_{sh} \leq b^*(\leq \bar{b}(e)) \) for any \( (e, \gamma) \). □

Lemma 1 conveys that the buyer will never try to ‘save the day’ by overinvesting into the relationship. While the seller’s underinvestment imposes a negative externality on B as it reduces his trade benefit (and therefore, his default payoff under a fixed price contract), the buyer realizes that he cannot overcome this externality by exerting a larger investment.²⁹

To continue the analysis, remember that when the seller indeed undertakes \( e \geq \bar{e} \), the buyer chooses \( b^* \) and a trade outcome \( (e^*, b^*) \) is achieved. Consider a potential equilibrium in which instead, \( S \) invests some \( e < \bar{e} \). While B will then have to make a minimum investment \( \bar{b}(e) > b^* \) to facilitate equilibrium trade, the Lemma suggests that he will rather undertake the defective effort \( b_{sh} \) which induces the parties to rescind their trade contract. This logic leads to the following conclusion: if the seller finds shirking attractive for a buyer investment of \( b = b^* \), the buyer will not counteract, and trade will not occur in equilibrium. When the seller shirks, she will choose \( e_{sh} = \arg\max \{\gamma V(e, \cdot) - (1 - \gamma)C(e)\} - \psi(e)(< e^*) \) as in the case of one-sided investments.³⁰ Specifically, suppose that

\[
t - c(e^*) - \psi(e^*) \geq t - c(e_{sh}) + \gamma[c(e_{sh}) - V(e_{sh}, b)] - \psi(e_{sh})
\]  

(21)

for \( b = b^* \). If this applies, trade supported by investment levels \( (e^*, b^*) \) is an equilibrium outcome. In contrast, if (21) is violated at \( b = b^* \), S chooses the shirking effort \( e_{sh} \), and

²⁹This would be different if agent B invested prior to the seller. In this latter case, overinvestment may arise as a commitment device to mitigate the seller’s shirking incentives.

³⁰Since \( e_{sh} \) does not depend on \( b \) (recall \( V_{sh} = 0 \)), its size is identical to the baseline model in Section 2 where only \( S \) invests. Also, \( e_{sh} < \bar{e} \) in analogy to the arguments in fn. 17 above.
agent $B$’s response is the effort level $b_{sh}$ by Lemma 1.\footnote{In particular, notice that when (21) is violated for $b = b^*$, it is violated even more so for $b = b_{sh}$. Since $e_{sh} < \hat{e}(< \hat{e}(b_{sh}))$ and $b_{sh} < \hat{b}(\hat{e})(< \hat{b}(e_{sh}))$ is $B$’s best response to $e_{sh} < \hat{e}$ by Lemma 1, $(b_{sh}, e_{sh})$ are mutual best responses.} The unique equilibrium then features investments $(e_{sh}, b_{sh})$ and trade does not arise.

We now show that shirking can arise in additional situations as well. Suppose condition (21) applies for $b = b^*$. The seller then does not shirk when $B$ expends $b^*$, and a trade equilibrium with investments $(b^*, e^*)$ exists. Even in such a situation, though, there may exist a second equilibrium in which again, both parties shirk and trade will be rescinded. To see this, suppose that while being satisfied for $b = b^*$, (21) is violated for $b = b_{sh}(< b^*)$.$\footnote{Since the right hand side of (21) decreases in $b$ (while the left hand side does not change), there must exist a non-empty range of parameterizations for which (21) holds for $b = b^*$, but not for $b = b_{sh} < b^*$.} $Interestingly, in this latter situation a shirking equilibrium $(e_{sh}, b_{sh})$ coexists with the trade equilibrium $(b^*, e^*)$: if the seller expects the buyer to shirk, she will shirk by the incentive constraint (21). Moreover, if $S$ defects in this way, $B$’s best response is to choose $b_{sh}$ by Lemma 1. These arguments demonstrate that two-sided investments give rise to multiple equilibria, with opposite economic outcomes.

The following parametric example illustrates these results. We return to the previous linear quadratic formulation but now incorporate buyer investments. Specifically, assume $V(e, b) = v + e + ab$ with $a \geq 0$ as a measure for the return on buyer investments. Let $c(e) = f - (3/2)e$ as before, and indicate investment costs as $\psi(e) = e^2/2$ and $\phi(b) = b^*/2$, respectively. Notice that $b^*(= b^{FB}) = a$ and the relationship has value for $(e^* = 3/2, b^* = a)$ whenever $S(e^*, b^* > 0$, that is, if$\footnote{As expected, a larger return to the buyer’s investments $a$ increases the range for which the relationship has value.}$

\begin{equation}
\begin{aligned}
v - f > -(21/8 + \frac{a^2}{2}).
\end{aligned}
\end{equation}

\footnotetext[31]{In particular, notice that when (21) is violated for $b = b^*$, it is violated even more so for $b = b_{sh}$. Since $e_{sh} < \hat{e}(< \hat{e}(b_{sh}))$ and $b_{sh} < \hat{b}(\hat{e})(< \hat{b}(e_{sh}))$ is $B$’s best response to $e_{sh} < \hat{e}$ by Lemma 1, $(b_{sh}, e_{sh})$ are mutual best responses.}
Along the lines of our previous analysis, the seller has no incentive to shirk whenever
\[
t - f + (9/8) \geq t - f + (3/2)e_{sh} + \gamma [f - v - (5/2)e_{sh} - ab] - (1/2)(e_{sh})^2
\] (23)
holds for \( b = b^* = a \), and \( e_{sh} = 3/2 - (5/2)\gamma \) as before. Let us focus on \( \gamma = 1/2 \) where the seller’s shirking effort is \( e_{sh} = 1/4 \). Then, (23) holds for \( b = a \) and a trade equilibrium \( (b^* = a, e^* = 3/2) \) exists whenever
\[
v - f \geq -(35/16 + a^2).
\] (24)
Comparing (22) and (24) reveals that for \( a \geq \sqrt{7/8} \) so that buyer investments are sufficiently important, a trade equilibrium exists whenever the relationship has value. Conversely, for \( a < \sqrt{7/8} \), (23) fails for \( b = b^* \) and a trade equilibrium ceases to exist over a non-empty range of parameters \( v - f \in [-35/16 + a^2, 21/8 + a^2/2] \). Over this range, the seller will shirk and choose \( e_{sh} = 1/4 \) in unique equilibrium. In response, the buyer will by Lemma 1 undertake the defective effort \( b_{sh} = \text{argmax}(1/2)ab - (1/2)b^2 = a/2 \) himself.\(^{34}\)

Finally, we demonstrate that even if \( a > \sqrt{7/8} \), the trade equilibrium as described above may co-exist with an additional shirking equilibrium in which both parties shirk and trade does not arise. Specifically, consider \( v - f \in [-35/16 + a^2, 35/16 + a^2/2] \). Within this range, the seller’s no-shirking condition (23) holds for \( b = b^* \), but it is violated for \( b = b_{sh} = a/2 \). Accordingly, the seller’s best response to \( b = a/2 \) is \( e_{sh} = 1/4 \) by (23) while by the result of Lemma 1, the buyer’s best response is to choose the defective effort \( b_{sh} \) as well.

To summarize, the feasible range is characterized by \( S(e^*, b^*) \geq 0 \), i.e., \( v - f \geq -[21/8 + a^2/2] \). When \( 21/8 + a^2/2 > 35/16 + a^2 \), i.e., \( a^2 < 7/8 \), a unique shirking equilibrium exists for \( v - f \in -[35/16 + a^2, 35/16 + a^2/2] \). Over the adjacent range \( v - f \in -[35/16 + a^2, 35/16 + a^2/2] \), trade equilibrium and a no-trade shirking equilibrium co-exist. Finally, for \( v - f \geq -[35/16 + a^2/2] \), the trade equilibrium becomes unique.

\(^{34}\)To check consistency, note that shirking becomes more attractive the smaller the buyer’s investment. Hence, if (23) does not hold for \( b = a \), it will be violated as well when \( b = b_{sh} = a/2 \).
Conversely, when \( a^2 > 7/8 \) so that \( 21/8 + a^2/2 < 35/16 + a^2 \), there are only two intervals of interest. While no unique shirking equilibrium appears, dual equilibria continue to coexist for \( v - f \in [-21/8 + a^2/2, 35/16 + a^2/2] \). Again, a unique trade equilibrium prevails when \( v - f \geq -[35/16 + a^2/2] \).

7 Conclusion

This paper investigates a natural variant of the workhorse model of bilateral trade and incomplete contracting. We show that when trade is viable only if the seller’s effort exceeds some threshold level, she may have an incentive to shirk, and to intentionally destroy the value of the relationship. This extreme hold up result provides a novel argument why relationship between economic agents sometimes fail. In particular, the seller gains from the renegotiation which is needed to rescind the existing trade agreement when the completion of trade has become inefficient. We show that shirking occurs more often when the parties carry out the trade transaction only after the buyer’s valuation has been revealed. Moreover, when both agents can invest into the relationship, an agent will never ‘overinvest’ to reduce the other party’s incentives to shirk, and multiple equilibria may arise. While results were achieved in the context of bilateral trade, it may be interesting to see whether they extend to other professional and personal bilateral relationships where specific investments often play a major role.
References


