

Tilburg University

Insurance and Information

Bovenberg, A.L.; Teulings, C.N.

Publication date:
2002

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Bovenberg, A. L., & Teulings, C. N. (2002). *Insurance and Information: Firms as a Commitment Device*. (CentER Discussion Paper; Vol. 2002-36). Macroeconomics.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



No. 2002-36

**INSURANCE AND INFORMATION: FIRMS AS A
COMMITMENT DEVICE**

By A.L. Bovenberg and C.N. Teulings

April 2002

ISSN 0924-7815

Discussion paper

Insurance and Information: Firms as a Commitment Device

A.L. Bovenberg* and C.N. Teulings[†]

Keywords: insurance, implicit contracts, moral hazard,
principal agent, commitment, shirking

JEL code: D23, D82

April 2002

Abstract

We explore the role of firms in insuring risk-averse workers. As a device that allows workers to commit to the delivery of their output, the firm arises endogenously as an alternative to the spot market if workers are sufficiently risk averse and the firm can base incentive payments on good information. Competition, however, may allow the spot market and explicit contracts to crowd out implicit insurance provided by the firm, even though the latter yields higher welfare. We explain why different governance structures coexist in quite homogeneous industries.

1 Introduction

Why do firms exist? Since Coase (1937) posed this fundamental question, it has still not fully been resolved. This paper views the firm as a device that enables risk averse suppliers to commit output to a specific buyer prior to its production. The implied relationship-specific nature of output allows insurance but harms incentives by creating

*CentER, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands, Email: A.L.Bovenberg@kub.nl

[†]Tinbergen Institute, Erasmus University Rotterdam, Burg. Oudlaan 50, 3062 PA Rotterdam, The Netherlands, Email: Teulings@few.eur.nl

a hold up problem. By facilitating insurance of risk averse suppliers, firms may create value depending on whether explicit and implicit contracts can alleviate the associated incentive problems. We explore when firms endogenously arise as an alternative to the spot market in which suppliers transfer property rights to a specific buyer only after the output has been produced.

We emphasize the relationshipspecific nature of output produced within a firm. In this connection, one can distinguish three alternative cases for the appropriability of output. First, output may be relationshipspecific: prior to its production, output is committed to one specific buyer. Services like hair cutting and house cleaning are typical examples in which output is necessarily relationshipspecific by the technological nature of the production process. Second, suppliers may retain property rights over the output during the production process so that it can be sold to any buyer after it has been produced. We therefore refer to this appropriation mode as the "spot market." We will show that the implied non-relationshipspecific nature of output rules out insurance of risk-averse suppliers by a specific buyer. When output happens to be high due to a favorable state of nature, the supplier sells the output on the spot market rather than delivering it to the insurer (i.e. the buyer). Accordingly, the insurer obtains the output only in bad states of nature and is thus not willing to offer an insurance contract.

We focus on a third case in which the supplier can choose whether output is relationshipspecific (so that output is committed to one specific buyer when production occurs) or not (so that output is not committed to a specific buyer so that the producer can sell the output to any buyer after it has been produced). Rather than imposed by the technological nature of the production process, the relationshipspecific nature of output thus becomes an endogenous response to incentive problems. We call the institutional arrangement that makes output relationshipspecific a "firm." The firm effectively transfers the property rights of the output to the buyer before the output is actually produced. On the one hand, the transfer of property rights implies that the buyer faces an agency problem in encouraging the supplier to provide effort. It also exposes the supplier to the temptation of the buyer to renege on the promise to provide incentive payments; the supplier may be held up. On the other hand, ex ante transferring the property rights of output to the buyer opens up scope for insurance by taking away the temptation of the supplier not to deliver the output in case of a good state of nature. Whether the supplier finds it optimal to expose herself to being held up within a firm depends on the value of insurance versus the cost of addressing agency and hold-up problems.

In traditional shirking models with risk-neutral agents, a firm is an exogenous invention that harms efficiency by creating a temptation for the firm to hold up risk-neutral workers. If a supplier would retain property rights over the output during the production process so that it could be sold to any buyer, the threat of selling the output on the spot market would induce the buyer to pay for the output. In this way, competition would eliminate the hold-up problem. Indeed, the contracting problems analyzed in shirking models (see e.g. Shapiro and Stiglitz (1984) and MacLeod and Malcolmson (1989)) originate in the assumed relationship-specific nature of output. In our model with risk-averse agents, in contrast, a firm is an endogenous innovation that allows agents to insure themselves by providing a commitment to deliver the output to a specific buyer, see Marglin (1970). In particular, a firm allows the employer to appropriate output, thereby in effect enforcing his property rights over the output. By building a factory hall and storing output in this hall, for example, the employer can prevent workers from stealing the output and selling it on the market. The institutional innovation of a firm facilitates production if suppliers feature such a strong aversion against production risk that they would refrain from producing for the risky spot market.

Another feature of the firm we emphasize is the implicit nature of the contracts that are used to address the agency and hold-up problems caused by the relationship-specific nature of output. In this connection, we depart from the standard principal-agent model (Holmstrom and Milgrom, 1987, 1991), which assumes that output is verifiable while effort cannot be observed. The latter model has a hard time explaining why low-powered incentive schemes rather than explicit piece rate contracts are typically observed in employment relationships. By allowing for only non-verifiable information on output and some, albeit imperfect, information on effort, we explain the observed prominence of fixed remuneration contracts, a phenomenon that is not easily explained within the standard principal agent framework. This suggests that lack of verifiable information rather than non-observability of effort is a major problem in employment relationships. Indeed, employers typically employ non-verifiable information to subjectively evaluate the effort level of their workers while the most important mutual obligations in employment relationships can not be enforced in court. We also show that only lack of verifiable information can explain self employment and the boundaries of the firm.

By focussing on the role of non verifiable information, we extend the literature on self-enforcing contracts in shirking models. In particular, we introduce the insurance motive in implicit contracts, thereby exploring the feasibility of self-enforcing insurance contracts. Insurance contracts are viewed as repeated prisoners' dilemmas, where the

prospect of future surpluses makes the contracts self-enforcing. We focus on the equilibrium in which the supplier provides effort first and the buyer subsequently meets the non-verifiable obligation to reward the worker for the effort.¹ Accordingly, incentives are costly because they require the supplier to concede part of the surplus to the firm so that the firm has an interest in rewarding the worker after the worker has provided the effort. Intuitively, the firm levies a charge on the supplier to offset the temptation to renege on the implicit contract. This charge distorts resource allocation by crowding out welfare-enhancing trade. Relationships that yield only marginal rents cannot be sustained because self-enforcing cooperation requires non-marginal rents from the relationship to keep honest the party who is tempted to renege on non-verifiable obligations.

By viewing the firm as an institution that allows workers to ex ante transfer the property rights over their output, we find conditions under which the firm endogenously arises as an alternative to the market. In particular, a firm survives competition with the spot market if workers are sufficiently risk averse while the firm bases its incentive payments on good information containing contracting costs. We thus reverse the Coasian perspective on the firm according to which the firm dominates the market if the market suffers from substantial transaction cost. Instead, the firm is competitive if the price charged by the firm to make the insurance contract self-enforcing is smaller than the insurance premium risk-averse workers are willing to pay. Only then do workers choose to transfer the property rights over the fruits of their labor before production actually occurs. In this connection, we demonstrate that non-verifiable information explains why the spot market can be more competitive than the firm so that self employment may survive competition from contractual employment within a firm. The reversal of the Coasian perspective is also consistent with the empirical evidence by Allen and Lueck (1999) that riskier output does result in a more widespread application of cropsharing contracts, as predicted by the standard principal agent model. Allen and Lueck suggest that this may be due to riskier output providing more scope for cheating on contractual obligations. Hence, greater riskiness might proxy for less verifiable information.

The composition of the worker's pay off is discontinuous at the point at which the worker is indifferent between the spot market and the firm. In particular, if the worker becomes marginally less risk averse or the quality of the non-verifiable information on effort worsens somewhat, the spot market crowds out the firm. An upward jump in the expected monetary pay-off of the worker offsets the additional risk assumed by the

¹This is in fact the vacancy equilibrium analyzed by MacLeod and Malcomson (1989) rather than the unemployment equilibrium explored by Shapiro and Stiglitz (1984).

worker. We can thus explain discrete regime switches in governance structures as a result of marginal changes in parameters. Accordingly, different governance structures may coexist in a single industry if such an industry operates around the point at which the firm is about as competitive as the spot market (see Krueger (1991)).

Our welfare analysis shows that competition may allow the spot market to break up the firm, even though the firm yields higher overall welfare. Intuitively, in competing for workers, firms do not take into account that they reduce the rents of other firms. As a result of this business stealing effect, competition destroys firms, thereby eliminating the rents from insurance. As the market crowds out bilateral relationships that enforce implicit insurance contracts, competition thus results in an excessively low level of insurance. Hence, contrary to the popular judgement that market competition resulting in vertical disintegration is beneficial, competition crowding out implicit contracts enforced by the firm may well harm efficiency by reducing the value that firms create by enforcing these implicit contracts. In line with common wisdom, spot market relationships do provide better incentives for effort, but they do so at the cost of a lower level of valuable insurance.

The mechanism described here resembles the concern expressed by Shleifer and Summers (1988) about competition in the market for corporate control crowding out valuable implicit contracts. Also they maintain that competition threatens the commitment to implicit contracts, thereby potentially threatening the value that firms create by enforcing implicit contracts. Whereas Shleifer and Summers focus on the crowding out of quasi rents compensating relationship-specific investments, we emphasize the reduction of rents that induce firms to comply with implicit insurance contracts with their workers. Through both these mechanisms, valuable implicit contracts are destroyed.

Although we cast our analysis in the context of employment relations between workers and firms, our main message has wider applicability. Former civil servants, who left their government job and are now hired by the same government on the basis of spot market relations, have given up insurance in favor of stronger incentives and higher expected pay. The tax payer is worse off, since the expected pay in the market exceeds the wage of the civil servant by more than the value of the additional effort extracted. Similar mechanisms are at work in the transition from defined benefit to defined contribution pension systems. Through defined benefit pensions, firms typically provide workers with substantial insurance of macro-economic risks, such as the rate of return on savings. These insurance systems are more and more crowded out by competition from defined-contribution schemes.

Similar to the value of competition between the spot market and the firm, the introduction of explicit contracting by making information verifiable is a mixed blessing. On the one hand, explicit contracts may result in more trade, thereby increasing rents of marginal firms. On the other hand, by crowding out non-marginal implicit contracts, explicit contracts destroy rents enjoyed by non-marginal firms who offer implicit insurance contracts. This latter effect can explain the complaints about the juridification of contracts, that is explicit contracts enforced by the courts crowding out self-enforcing, implicit contracts based on long-term relationships. More generally, we show that society may overinvest in additional information if the existing information is already of reasonable quality. Intuitively, the main effect of additional information in these circumstances is to redistribute rents towards those investing in the information, thereby harming other traders. If initial information is poor, in contrast, agents may underinvest in information because additional information in that case creates additional trade, thereby also benefiting potential trading partners.

The rest of this paper is structured as follows. After section 2 discusses the model, section 3 considers the case in which suppliers output is relationshipspecific so that output is committed to a specific buyer when it is being produced. Depending on the quality of the available information, it explores the contracts that can be used to address the agency and hold-up problems associated with this commitment. This section offers a integrative framework of which both the standard principal agent model of Holmstrom and Milgrom (1987) and the shirking model of Shapiro and Stiglitz (1984) are special cases. Subsequently, section 4 turns to the case in which suppliers cannot commit output to a specific buyer and shows that this rules out insurance. Section 5 considers the case in which the supplier has the option of committing output to a specific buyer. It analyses the conditions under which suppliers choose to commit output prior to its production so that the firm endogenously arises as an alternative to the spot market. Section 6 introduces the supplier's outside option of choosing leisure so that the supplier faces three alternatives: production with commitment, production without commitment, and no trade. It also investigates the welfare implications of competition between these three alternatives. Section 7 concludes.

2 The model

Consider an economy populated by a fixed, large number of two types of infinitely lived players, namely risk-averse suppliers and risk-neutral buyers. Suppliers exert effort in

order to produce output that is consumed by buyers. Transactions between buyers and suppliers occur only at fixed time intervals—the so-called transaction period. All deliveries of output and all payments occur at the end of this transaction period. Without loss of generality, we normalize the time length of the transaction period to unity. Output in transaction period t , x_t , satisfies

$$x_t = e_t + z_t,$$

where e_t denotes the level of effort chosen by the supplier in transaction period t . z_t represents a normally distributed random variable, which measures shocks on output. We normalize $E[z_t] = 0$ and² $\text{Var}[z_t] = 1$ without loss of generality.³ The pay-offs of suppliers and buyers, r_t^i , $i = s, d$, are given by

$$r_t^s = y_t - \frac{1}{2\varepsilon}e_t^2, \tag{1}$$

$$r_t^d = x_t - y_t, \tag{2}$$

where y_t denotes the money buyers transfer to suppliers in exchange for output. The quadratic specification of effort cost for the suppliers implies that supply is less than infinitely elastic, with the elasticity of effort depending on the parameter ε . Demand, in contrast, is infinitely elastic due to the linear form of the pay off. A buyer may trade with several suppliers simultaneously. However, with perfectly elastic demand, a buyer's trade with one particular supplier does not affect the buyer's valuation of trade with other suppliers.

Suppliers exhibit constant absolute risk aversion. Their expected lifetime utility u^s is related to the pay offs in the various transaction periods as

$$U^s = - \sum_{t=1}^{\infty} \frac{1}{(1+\delta)^t} E[\exp[-\theta r_t^s]]$$

where θ , $0 \leq \theta \leq \varepsilon$ stands for the coefficient of absolute risk aversion and $\delta > 0$ represents the discount rate. For notational convenience, we define a transformed utility criterium $u^s = -\frac{\delta}{\theta\varepsilon} \ln[-U^s]$. With a stationary, normal distribution of r_t^s , this transformed expected life-time utility is given by

$$u^s = \frac{1}{\varepsilon} E \left[y_t - \frac{1}{2\varepsilon} e_t^2 \right] - \frac{\theta}{2\varepsilon} \text{Var} \left[y_t - \frac{1}{2\varepsilon} e_t^2 \right] \tag{3}$$

²Throughout the paper, square brackets denote the arguments of functions while round and curly brackets represent the ordering of arithmetic operations.

³The implications of the model would be unaffected if one would change the standard deviation inversely to the coefficient of absolute risk aversion θ (defined below).

where we used equation (1) to eliminate r_t^s .

Expected lifetime utility of the risk-neutral buyers u^d amounts to

$$u^d = \frac{1}{\varepsilon} \sum_{t=1}^{\infty} \frac{\delta}{(1+\delta)^t} \mathbb{E}[r_t^d] \quad (4)$$

With a stationary distribution of r_t^d , we arrive at

$$u^d = \frac{1}{\varepsilon} \mathbb{E}[e_t - y_t] \quad (5)$$

where we used equation (2) to eliminate r_t^d and $\mathbb{E}[x_t] = \mathbb{E}[e_t]$.

Each transaction period consist of three phases: a bargaining phase, a production phase, and a transaction phase. During the bargaining phase, traders must make a number of decisions. First, if a contract matched them to a specific trading partner during the previous transaction period, they must decide whether they continue that relationship. Second, if they are not matched with a trading partner, suppliers have to decide whether they produce or collect their outside option during that transaction period. If they produce, they can opt for either first producing output and looking for a buyer on the spot market only afterwards or they can first seek a trading partner, negotiate an (implicit or explicit) contract and only subsequently start production. These steps during the bargaining phase are costless and do not take any time.

At the beginning of the production phase, the state of nature is revealed to the producing suppliers. Next, these suppliers decide on their level of effort. Finally, the transaction phase occurs at the end of the production period. During this phase, suppliers deliver the output to the buyers and buyers submit their payments to suppliers. Also the events during the transaction phase are costless and do not take any time. Time elapses thus only during the production phase.

We distinguish three modes of information for effort and two information modes for output. At one extreme, effort can be verified by a third party so that effort is contractible. Verifiable information is denoted by v . At the other extreme, neither a third party nor the buyer can observe effort; *no* represents this case of non-observable effort. In the intermediate case, effort cannot be verified by a third party but can be observed by the buyer. This case of non-verifiable information is denoted by nv . Output can be either verifiable or non-verifiable by a third party. Output is always observable by both transaction partners. The supplier knows both his own level of effort and the state of nature z_t and thus also the output level. The buyer is able to evaluate the output when it is delivered at the end of the transaction period. Together, the three information modes on effort and two on output, yields six possible combinations,

denoted for example by $\{e, x\} = (v, nv)$ for the case where effort is verifiable and output is non-verifiable.

As discussed in the introduction, we distinguish three modes of appropriability of output. The next three sections analyze each appropriation mode in turn. First, we explore the case in which output is relationship-specific: at the beginning of a transaction period, a supplier has to commit the output to a specific buyer. We then turn to the case in which output is non-relationship specific: a supplier cannot commit at the beginning of a transaction period to which buyer to deliver her output. Subsequently, we investigate the case in which the supplier can choose whether or not to commit the output to a specific buyer at the start of a transaction period.

3 Equilibrium with commitment

The section explores optimal insurance contracts if output is relationship-specific. We derive the optimal contract for each of the six conceivable information modes. Table 1 summarizes the outcome for each information mode. The subsequent subsections provide a verbal discussion of the relevant contract for each mode. The proof of their optimality is delegated to the appendix.⁴ We first explore the two information modes that do not involve non-verifiable information, that is $\{e, x\} = (v, v), (no, v)$. These purely explicit contracts serve as benchmarks for implicit contracts, which employ non-verifiable information and are discussed in Section 3.2. Since contracts and equilibrium strategies are stationary across transaction periods, we simplify notation by omitting the time subscript t .

3.1 Explicit contracts

3.1.1 Mode $\{e, x\} = (v, v)$: full insurance

Competition between infinitely elastic buyers implies that the equilibrium contract maximizes suppliers' utility while buyers' utility corresponds to that in the outside option, $u^d = 0$: suppliers switch to the buyer who offers the contract that yields highest utility to the suppliers. The risk-neutral buyer offers the risk-averse supplier full insurance of

⁴For two information modes $\{e, x\} = (nv, nv), (nv, v)$, we cannot prove optimality. Given a (truncated) linear form of the contracts, we derive the optimal parameters associated with this particular functional form. The presented incentive compatible contracts serve as lower bounds for the optimal contracts.

the production risk, conditional on the supplier exerting the efficient level of effort, ε . This outcome is first best for the supplier, conditional on the production technology and the participation constraint of the buyer. In the sequel, we use the first-best level of effort to normalize the degree of absolute risk aversion: $0 \leq \eta \equiv \theta/\varepsilon \leq 1$. Hence, η can be interpreted as the degree of "relative" risk aversion, namely absolute risk aversion relative to the first-best level of output.

3.1.2 Mode $\{e, x\} = (no, v)$: explicit piece rate

Without any information on effort, incentives must be based on output so that risk is transferred to the risk-averse supplier, implying inefficient risk sharing. The optimal contract therefore trades off sharing risk and providing efficient incentives. This contract, which is linear and thus features a fixed piece rate b , offers only partial insurance to the supplier. With risk-averse suppliers, $\eta > 0$, incentives are not strong enough to induce the supplier to select the first-best level of effort (i.e. $b < 1$ so that $e^* < \varepsilon$). If suppliers are risk neutral (i.e. $\eta = 0$), the optimal contract can costlessly provide first-best incentives so that effort and utility converge to the full insurance case.

3.2 Implicit contracts

We now turn to contracts employing non-verifiable information. In contrast to explicit contracts, implicit or self-enforcing contracts can be based on non-verifiable information. These contracts rely on a repeated game with trigger strategies to enforce the cooperative equilibrium. In particular, if a player complies to the non-verifiable obligations contained in the implicit contract, the trading partner rewards this behavior by continuing the relationship. If a player shirks on these obligations, however, the trading partner punishes the shirker by terminating the relationship. MacLeod and Malcomson (1989) show that (at least) two cooperative equilibria exist. In one equilibrium, buyers first pay and are thus exposed to the supplier shirking on the non-verifiable obligation to provide effort. Suppliers do not shirk only if they are rationed, for example because they cannot instantaneously find a new trading partner after being fired. Shapiro and Stiglitz (1984) focus on this 'unemployment' equilibrium. In the other equilibrium, suppliers first deliver effort and then must wait and see whether the buyer complies to the non-verifiable obligation to pay. The buyer can credibly commit to comply only if he is rationed so that he cannot find another supplier if the current trading partner decides to quit. In the context of the labor market, this equilibrium is known as the 'vacancy'

equilibrium. We focus on this latter equilibrium.⁵ In order to save on notation and without loss of generality, we abstract from exogenous turnover. Hence, a buyer who loses a supplier by failing to comply to his non-verifiable obligations will never find another trading partner, so that his expected life time utility equals the instantaneous pay off of an unmatched buyer $u_{un}^d = 0$.⁶

For this type of implicit contract to be self enforcing, two incentive compatibility constraints must be satisfied – one for the supplier and one for the buyer. In deriving these non-shirking conditions, we decompose the monetary transfer into a contractible part and a non-enforceable part: $y[x, e] = y_c[x, e] + y_n[x, e]$, where the contractible part $y_c[x, e]$ can be contingent only on verifiable information. The non-shirking condition for the supplier can be written as

N.s.c. I: take the money and run

$$\forall z : y_c[x, e_{sh}] - \frac{1}{2\varepsilon} e_{sh}^2 \leq y_c[x, e^*] + y_n[x, e^*] - \frac{1}{2\varepsilon} e^{*2}, \quad (6)$$

where e_{sh} stands for the effort level if the supplier is shirking and e^* denotes the equilibrium value of effort. The buyer pays the non-verifiable bonus payment $y_n[\cdot]$ only if the supplier does not shirk on non-verifiable obligations. Since the supplier sets effort knowing the realization of z , the non-shirking condition must be satisfied for all values of z . Expression (6) states that the value of shirking should not exceed the value of complying to the contractual obligations, since the supplier is not rationed and can therefore not be punished for shirking by the buyer ending the relationship.⁷

We now turn to the non-shirking condition for the buyer. The buyer faces a trade-off between the short-run gain of shirking on the obligation to pay $y_n[\cdot]$ and the long-run gain of holding on to the surplus from the relationship by complying to his non-verifiable obligations. At the beginning of the transaction period, the buyer's lifetime utility when

⁵This choice can be justified by assuming that suppliers feature substantially higher discount rates than buyers do (see e.g. Baker, Gibbs and Holmstrom (1994)). This assumption implies that, compared to buyers, suppliers need a larger surplus to support the relationship.

⁶Exogenous turnover would merely complicate notation without generating new insights. Alternatively, the absence of exogenous turnover can be justified by assuming that past behavior of buyers is known to all potential suppliers, who participate in punishing shirking buyers by not engaging in future trade. With this collective punishment in place, shirking on implicit agreements harms one's future trading opportunities, see Greif (1993) and Greif, Milgrom and Weingast (1994).

⁷This non-shirking condition would become non-linear if we would allow for risk aversion and rationing at the same side of the market. This is the main reason why we focus on the equilibrium in which the risk-neutral buyer is rationed.

shirking reads, see equations (4) and (5)

$$u_{sh}^d [e^*, z] = \frac{\delta}{1 + \delta} \frac{1}{\varepsilon} r_{sh}^d [e^*, z] + \frac{1}{1 + \delta} u_{un}^d = \frac{\delta}{1 + \delta} \frac{1}{\varepsilon} r_{sh}^d [e^*, z], \quad (7)$$

where $u_{sh}^d [e^*, z]$ denotes the life-time utility of the buyer when shirking, conditional on the realization of the state of nature during the current transaction period, z . The first term at the right-hand side of (7) represents the net discounted value of the pay off during the transaction period. The second term at the right-hand side of (7) reflects what happens after the transaction period; the supplier punishes the buyer for shirking by ending the relationship so that the shirking buyer receives the value of an unmatched player. Since $u_{un}^d = 0$, this latter term drops out.

The corresponding expected lifetime utility of non-shirking (i.e. meeting one's non-verifiable obligations) conditional on the state of nature during the current transaction period, $u_{n-sh}^d [e^*, z]$, reads

$$u_{n-sh}^d [e^*, z] = \frac{\delta}{1 + \delta} \frac{1}{\varepsilon} r_{n-sh}^d [e^*, z] + \frac{1}{1 + \delta} u_{n-sh}^d, \quad (8)$$

where u_{n-sh}^d denotes expected life-time utility of non-shirking.

The non-shirking condition $\forall z : u_{sh}^d [e^*, z] \leq u_{n-sh}^d [e^*, z]$ ensures that the non-verifiable contractual obligations are self-enforcing. Taking expectations in equation (8), we find $u_{n-sh}^d = \frac{1}{\varepsilon} \mathbb{E}[r_{n-sh}^d [e^*, z]]$. Substituting this result into (8) to eliminate u_{n-sh}^d and subsequently substituting the resulting expression and (7) into the non-shirking condition, we arrive at

$$\forall z : r_{sh}^d - r_{n-sh}^d \leq \frac{1 - \chi}{\chi} \mathbb{E}[r_{n-sh}^d], \quad (9)$$

where $0 \leq \chi \equiv \frac{\delta}{\delta + 1} \leq 1$. Substituting $r_{sh}^d = x - y_c$ and $r_{n-sh}^d = x - y_c - y_n$ (see (2)) and using $\mathbb{E}[x] = \mathbb{E}[e^*]$, we arrive at the non-shirking condition for the buyer:

N.s.c. II: take the output and refuse to pay

$$\forall z : y_n^* \leq \frac{1 - \chi}{\chi} \mathbb{E}[e^* - y_c^* - y_n^*] = \frac{1 - \chi}{\chi} \varepsilon u^d, \quad (10)$$

where y_n^* is shorthand notation for $y_n [x, e^*]$ and where y_c^* is defined analogously. The benefit from shirking (saving the equilibrium non-verifiable payment y_n^* on the left hand side) must not exceed its cost (loosing the future rents of the relationship on the right-hand side). The non-shirking condition should hold for all possible values of this non-verifiable payment, including the largest one $\max_z [y_n^*]$. Condition (10) implies that the

non-verifiable payment y_n^* cannot be positive if $\chi = 1$: a buyer who reneges on the non-verifiable obligation cannot be punished because the future punishment is discounted away, $\delta \rightarrow \infty$.

The supplier does not obtain any rents from the relationship so that payments by the supplier to the buyer cannot be self-enforcing. Hence, the non-verifiable payment cannot be negative for any realization of z :

$$y_n^* \geq 0. \tag{11}$$

This non-negativity constraint together with (10) implies that the participation constraint for the buyer $u^d \geq 0$ is met. The participation constraint of the buyer thus does not need to be checked separately. Indeed, if the contract includes a non-verifiable bonus, $y_n^* > 0$, the buyer enjoys a positive surplus. Even though buyers in the bargaining stage are perfectly competitive, a competitive contract contingent on non-verifiable information transfers some surplus to the buyer in order to induce the buyer to actually transfer the non-verifiable payments to the supplier.

We are now in a position to derive the optimal contracts for the information modes involving non-verifiable information.

3.2.1 Mode $\{e, x\} = (nv, nv)$: fixed remuneration contract

The implicit contract stipulates a level of effort e^* and a fixed implicit bonus payment y_n^* . Buyers pay this non-verifiable bonus since they do not want to lose the future rents of the relationship, $u_{n-sh}^d = \frac{1}{2}\chi(1 - \chi)$. Effort is set below its first best level (i.e. $e^* = (1 - \chi)\varepsilon < \varepsilon$) because incentives require the supplier to concede part of the surplus to the buyer in order to ensure that the buyer has an interest in complying to the incentive payments. Just as in an explicit piece rate contract, therefore, incentives are costly to suppliers. In an explicit piece rate, incentives requires inefficient risk sharing while in an implicit fixed remuneration contract, credible incentives require a transfer of part of the surplus to the buyer.

The fixed remuneration contract resembles a standard employment contract: wages are fixed and unrelated to output, but employers require workers to maintain a certain effort level. Indeed, one of the major puzzles in labor economics is why so many low powered incentive compensation schemes are observed. The standard principal agent model with contractible effort can explain low powered incentives only if suppliers are extremely risk averse ($\eta \rightarrow 1$). The non-verifiability of output and effort provides a simple explanation for the prominence of low powered incentive schemes without having

to rely on extremely risk-averse workers. Since fixed remuneration is a more common transaction mode for employment relationships than explicit piece rate contracts, the lack of verifiability of output seems to be the major problem in employment relationships. The standard principal agent model appears to be too pessimistic about the observability of effort but overly optimistic about the verifiability of output.

3.2.2 Mode $\{e, x\} = (no, nv)$: implicit piece rate *ipr*

This mode yields the most complicated contract. In order to avoid both negative and large positive non-verifiable payments, we consider an implicit contract with a truncated linear relationship between output x and the non-verifiable payment y_n , with both a maximum and minimum for y_n . The contractible payment y_c is constant since neither output nor effort are verifiable. The provision of incentives for effort requires the non-contractible bonus y_n to depend on x . Hence, y_n is stochastic. However, y_n cannot depend linearly on x for two reasons. First, $y_n < 0$ cannot be supported in an equilibrium in which suppliers are not rationed (see (11)). Second, y_n cannot be too large because a large non-verifiable bonus is credible only if suppliers transfer a large surplus to buyers (see (10)), which is contradictory to the interest of the supplier.

Consider a supplier who decides on effort e , taking as given a truncated linear contract and knowing the state of nature z .⁸ Figure 1 shows the situation graphically. z is on the horizontal axis while the upper panel shows effort e and the lower panel the non-verifiable payment y_n . If $z \in [z^0, z^+]$, the decision of the supplier is unaffected by the non-linearities in the contract. Just as in a regular linear explicit piece rate contract, therefore, the optimal effort level amounts to $e^* = b\varepsilon$, where b denotes the piece rate. If $z < z^0$, y_n is less than the cost of effort $\frac{1}{2\varepsilon}e^2$ when obeying to the linear rule $e = b\varepsilon$. In these unfavorable states of nature, therefore, the supplier faces no incentives to provide any effort, and hence $e^* = 0$. If $z \in (z^+, z^{\hat{}})$, the supplier gets the maximum value of y_n , even if he supplies less effort than implied by the linear rule $e = b\varepsilon$. Since additional effort does not yield a higher payment, the supplier sets $0 < e^* < b\varepsilon$. In favorable states of nature $z \geq z^{\hat{}}$, the non-verifiable bonus hits its ceiling even if no effort is provided so

⁸Some commentators wondered whether the model would not be simplified by assuming that suppliers set e not knowing z . This suggestion invokes two remarks. First, though e were non-stochastic in that case, y_n would still be non-linear, for the same reasons as discussed in the text, leading to a model of comparable complexity. Second, the assumption that suppliers do not know z when setting e would imply that a non-linear contract involving extreme punishments for low levels of output could approximate the first best in $\{e, x\} = (no, v)$ (see Mirrlees (1976) and Holmstrom and Milgrom (1987)).

that $e^* = 0$.

Since the expressions for the contract's parameters are rather complicated in the general case, we focus on the special case $\varepsilon \rightarrow \infty$, holding η constant (hence $\theta \equiv \eta\varepsilon \rightarrow \infty$). By $\varepsilon \rightarrow \infty$, optimal expected output increases as the marginal cost of effort falls while the variability of the state of nature remains the same. Relative to the expected level of output, therefore, the variability of the non-verifiable payment drops to zero. The utility cost of risk compared to the expected value of output, in contrast, is unaffected by the drop in relative riskiness because the coefficient of "relative" risk aversion η remains unaffected by the simultaneous increase θ and ε .

In this limiting case $\varepsilon \rightarrow \infty$, the optimal contract becomes linear. The intuition for this result is as follows. As all contracts based on non-verifiable information, the optimal contract must strike a balance between providing sufficient incentives and containing the surplus that must be conceded to the buyer. On the one hand, the non-verifiable payment must be large enough to compensate for the cost of effort in order to satisfy the suppliers' non-shirking condition (6). On the other hand, in order to meet the buyer's non-shirking condition (10), the payment cannot be too large for that would require a large share of the surplus to be conceded to the buyer. When the non-verifiable payment is stochastic, both conditions are more difficult to satisfy simultaneously, since they have to apply for both extremes, the "worst" case $z = z^0$, where y_n is just sufficiently large to compensate the supplier for the cost of effort $\frac{1}{2\varepsilon}e^2$, and for the "best" case $z = z^+$, where the expected future surplus is just enough to let the buyer comply with the non-verifiable obligation to pay y_n . This limits the linear range of the contract $[z^0, z^+]$ in which full incentives can be provided. The larger this range becomes, the more rents have to be conceded to the buyer. By increasing ε , however, the relative variability of output decreases and hence the relative variability of y_n . Hence, the optimal range $[z^0, z^+]$ increases so that more states of nature are covered by the linear part of the contract. In the limit $\varepsilon \rightarrow \infty$, non-linearities lose their importance altogether so that all states of nature fall in the linear part of the contract. At the same time, the variability of y_n (as a ratio of expected output) converges to zero, so that only the expected value of y_n matters for the non-shirking condition.

An implicit piece rate contract combines two imperfections, namely a lack of insurance (since the contract is contingent on output rather than effort) and a lack of verifiability (so that part of the surplus has to be conceded to the buyer). From the point of view of suppliers, incentives are thus costly for two reasons, namely, first, the risk imposed on risk-averse suppliers, and, second, the rents that must be transferred to

buyers in order to enforce non-verifiable payments. Both these reasons dull incentives. Indeed, the equilibrium contract with non-verifiable output and non-observable effort implies lower incentives and a lower equilibrium effort level e^* than in case output is verifiable or effort is observable.

3.2.3 Mode $\{e, x\} = (nv, v)$: mixed piece rate

A mixed piece rate contract combines an explicit piece rate b on output and a fixed non-verifiable payment for exerting a particular level of effort e^* . The explicit part of the contract reduces the non-verifiable bonus in the implicit contract y_n^* that is required to induce the supplier to supply a given amount of effort e^* . When shirking, optimal effort e_{sh} amounts to $b\varepsilon$, just as in a pure explicit piece rate contract. The non-verifiable transfer should compensate the supplier for the cost of effort $\frac{1}{2\varepsilon}e^{*2}$ net of the cost of effort when shirking $\frac{1}{2\varepsilon}e_{sh}^2 = \frac{\varepsilon}{2}b^2$ and of the reward provided by the explicit contract for the additional effort $b(e^* - e_{sh}) = be^* - \varepsilon b^2$. The non-verifiable bonus thus fills the hole left by the explicit contract; the less incentives are provided through the explicit piece rate contract, the more incentives should be provided through the implicit contract.

The explicit part of the contract, the piece rate b , is important compared to the fixed implicit bonus payment y_n , if the cost of the explicit part (assignment of risk to the supplier) is small compared to that of the implicit part (the surplus transferred to the buyer). In the context of the mixed piece rate, the explicit and implicit part of the contract are thus substitutes; if the implicit part is costly, agents rely more heavily on the explicit part to provide incentives.

The mixed piece rate contract offers a better description of the contract faced by salesmen than the standard principal agent model. A firm employing a salesman who sells too few pieces has no incentive to comply to its non-verifiable obligations. Hence, the salesman is punished not only by reducing the contractible piece rate reward but also by losing non-verifiable payments, i.e. benefits that he can reasonably expect when providing sufficient effort but which are not included in his legal labor contract.

3.2.4 Mode $\{e, x\} = (v, nv)$: full insurance

The non-verifiability of x does not change anything compared to the mode $\{e, x\} = (v, v)$ because information on output does not play a role in the full insurance contract. Hence, also the information mode $\{e, x\} = (v, nv)$ yields the first-best full insurance contract.

3.3 The role of information content

We are now in a position to discuss Table 1 in greater detail. The first and second line contain the utility of, respectively, a supplier and a buyer engaging in trade. The third line provides total welfare (defined as the sum of the utilities of the supplier and the buyer). The fourth line indicates the incentives for effort provided by each contract by presenting the effort the contracts elicit. The fifth line contains the piece rate used for that purpose. Finally, the sixth line lists the expected non-verifiable payment. As a consequence of the buyers' non-shirking condition (10), this payment is directly proportional to the surplus reaped by the buyers (i.e. the second line).

Proposition 1 *The parameter $\chi \in [0, 1]$ can be interpreted as a general measure of the quality of information on output or effort, where $\chi = 0$ coincides with verifiable information and $\chi = 1$ with non-observable information (this latter case applies only to effort since output is always observable by both transaction partners).*

The proof of this proposition is derived from Table 1 by considering all information modes involving non-verifiable information and setting χ equal to 0 and 1 respectively. We first consider non-verifiable information on effort, starting with the information mode $\{e, x\} = (nv, nv)$ yielding a fixed remuneration contract. Setting $\chi = 0$, we observe that this contract corresponds in all dimensions to the full insurance contract for the information mode $\{e, x\} = (v, nv)$. The other information mode involving non-verifiable information on effort is $\{e, x\} = (nv, v)$, yielding a mixed piece rate contract. At $\chi = 0$, this contract yields the same outcomes as in $\{e, x\} = (v, v)$, while at $\chi = 1$ the outcomes coincide with those if $\{e, x\} = (no, v)$.

Now turn to the information mode with non verifiable information on output $\{e, x\} = (no, nv)$, yielding an implicit piece rate contract. At $\chi = 0$, this contract produces the same outcome as the explicit piece rate contract corresponding to the information mode $\{e, x\} = (no, v)$. \square

The information modes of verifiability and non-observability can thus be viewed as extreme cases of a general measure of information problems χ . Verifiable information is the limiting case of ever declining verifiability problems ($\chi = 0$); the trigger strategy based on this information becomes as effective in enforcing cooperation as a legal contract verifiable in court. Non-observable information is the opposite extreme of ever increasing verifiability problems ($\chi = 1$). The trigger strategy that is conditioned on the available information does not add any value. Intuitively, enforcement problems increase with the discount rate δ (implying that the long-run gains from complying to

non-verifiable obligations are discounted more heavily), or equivalently, with the length of the transaction period (with a constant discount rate per period of time, a longer transaction period corresponds to a larger discount rate per transaction period δ). A longer transaction period implies that mutual obligations are cleared at a lower frequency so that the buyer has to wait longer to observe the non-verifiable information. The parameter χ can thus be viewed as summarizing the quality of non-verifiable information. We shall therefore refer to the parameter χ as an inverse measure the quality information: if $\chi = 0$, the quality of information is perfect; if $\chi = 1$, the quality of information is so bad that it does not add any value.

The parameter χ can alternatively be interpreted as a parameter indicating how serious the hold-up problem is in the relationship between buyer and seller. With relationship-specific output, effort is a specific investment that gives rise to a hold-up problem: the supplier must invest effort before the output can be sold to the buyer to whom the output is committed. The larger $\chi \equiv \frac{\delta}{1+\delta}$ and therefore the lower the frequency at which mutual obligations are cleared, the more important is the investment character of effort and thus the more serious is the hold-up problem associated with this specific investment is.

Proposition 2 *If the quality of information on effort is not inferior to that on output, then information on output is irrelevant.*

The only information mode in which the information on effort is superior to that on output is $\{e, x\} = (v, nv)$. This information mode yields the same outcome (i.e. a full insurance contract) as the mode in which the information on output is upgraded to that on effort, namely $\{e, x\} = (v, v)$. \square

The intuition for this proposition is that, in order for risk-neutral buyers to shield risk-averse suppliers from production risk, the contractual payments should depend on effort rather than output. Hence, information on effort is more valuable than information on output so that contracts should employ information on output only if adequate information on effort is lacking.

Proposition 3 *Supplier's utility and effort are strictly increasing in the quality of information on effort. They are strictly increasing in the quality of information on output, whenever this information is superior to the information on effort.*

The proof follows directly from Table 1 (by differentiating the entries for u^s in the last three columns (corresponding to the implicit contracts) with respect to χ) and Proposition 2. \square

Information on effort is always valuable to suppliers because it allows buyers to insure suppliers at lower costs to the suppliers. Information on output is valuable only if it provides superior information than the information on effort that is available. The better the quality of information becomes, the cheaper it is to provide incentives and the higher the equilibrium level of effort is.

Proposition 4 *Buyer's utility is positive if and only if the contract employs non-verifiable information. In that case, it is hump shaped in the quality of information χ .*

The proof follows directly from Table 1 by differentiating the entries for u^d in the last three columns (corresponding to the implicit contracts) with respect to χ . \square

Whenever a contract involves implicit incentive payments, $y_n^* > 0$, buyers should extract rents from the continuation of the relationship in order to enforce their non-verifiable obligations, see condition (10). The second part of the proposition implies that buyers have an interest in an intermediate degree of non-verifiability. Only in that case, buyers are able to obtain a significant share of a substantial surplus. Too much verifiability eliminates the reason for paying buyers rents and therefore reduces the share of the surplus that accrues to buyers. Too little verifiability crowds out trade by reducing the incentives for effort and thus decreases the overall surplus from trade. Beyond a certain degree of non-verifiability, the negative impact on the buyer's pay off of the smaller overall surplus dominates the positive impact of a larger surplus share.

The hump-shaped curve for the buyer's utility resembles the Laffer curve for government tax revenues as a function of the tax rate. Indeed, if the information becomes more non verifiable (as indicated by a larger parameter χ), the buyer has to impose a larger implicit 'tax' on suppliers in order to offset his short-run gain from renegeing on the obligation to pay the incentive payments. This implicit tax can be viewed as the price suppliers have to pay buyers to ensure that the buyers are committed to the contract. Starting from a zero tax rate, the revenues from this implicit tax (i.e. the surplus that accrues to buyers) first increase with the level of the imposed tax but eventually decrease as the adverse incentive effects of higher implicit tax rates erode the base of the tax.

Our analysis assumes that the quality of information is an exogenous variable. In practice, however, agents may invest in the quality of information in order to alleviate incentive problems so that the information quality becomes an endogenous variable. The following proposition addresses the competitive value of information on effort and output.

Proposition 5 *The competitive value of information on effort is strictly positive whenever the supplier is risk averse. The competitive value of information of output is strictly positive whenever this information is superior to the information on effort.*

Competition forces rationed buyers to employ the quality of information that maximizes suppliers' utility. Hence, the competitive value of additional information is determined by its effect on u^s . The proof follows directly by differentiating the entries for u^s in the last three columns (corresponding to the implicit contracts) with respect to χ . \square

With risk-neutral suppliers, the quality of information on effort is irrelevant whenever verifiable information on output is available, as in the information mode $\{e, x\} = (nv, v)$. In that case, information on output crowds out the information on effort. If effort is verifiable, better information on non-verifiable output does not generate any competitive value, reflecting Proposition 2.

While proposition 5 investigated the competitive value of additional information, we can also analyze the impact of additional information on overall welfare defined as the sum of u^s and u^d .⁹ The competitive value of additional information is given by its impact on only suppliers' welfare. Accordingly, the difference between the welfare impact of additional information and its competitive value is given by its effect on the utility of buyers. The last part of Proposition 4 indicates that, if information is poor (as indicated by a high level of χ), investments in additional information generate positive external effects by increasing the rents accruing to buyers. If information is almost verifiable (as reflected by a low level of χ), in contrast, additional information reduces the pay-off of buyers so that additional information generates negative external effects. These negative external effects are due to a business stealing effect: in competing for suppliers to trade with, buyers fail to take into account that they reduce the rents of other buyers. This results in the following proposition:

Proposition 6 *Whereas society underinvests in additional information if information is poor, it overinvests in additional information if information is already of good quality.*

The intuition behind this proposition is that additional information yields positive external effects if it generates substantial additional trade. If it mainly redistributes rents away from buyers to sellers, in contrast, it yields adverse external effects. This latter effect can be so strong that making output verifiable may even harm overall welfare

⁹This definition is natural since the marginal utility of income equals unity for both players. However, since the external effects involve buyers' welfare only, our qualitative results go through as long as a positive weight is attached to the welfare of the buyers.

(as measured by the sum of the utilities of the supplier and buyer). In particular, Table 1 shows that moving from the information mode $\{e, x\} = (nv, nv)$ to the information mode $\{e, x\} = (nv, v)$ by making output verifiable harms overall welfare if $\eta(1 + \eta) > (\chi + \eta)^2$. Accordingly, if the information on effort is reasonable good and risk aversion is large, the introduction of verifiable information on output reduces welfare. This is a typical second best result: improving information may harm welfare. Reflecting proposition 5, the competitive value of verifiable information on output is positive. The intuition for the different signs of the competitive value and the welfare impact of additional information is a business stealing effect. In particular, competing buyers invest in verifiable information on output to attract suppliers but in the process reduce the rents accruing to other buyers.

4 Equilibrium without commitment

If the supplier is not committed to deliver the output to a specific buyer, he can sell the output on the spot market. The buyer's utility function implies that equilibrium price level on the spot market is unity: $y = x$. With a normally distributed pay off, the supplier sets effort so as to maximize (3) subject to the constraint $y = x$. The resulting equilibrium level of effort $e^* = \varepsilon$ coincides with the first-best level in a full insurance contract under commitment. Trade in risk, however, is inefficient. Since the risk-averse supplier bears all the production risk, his utility in the spot market, $u^s = \frac{1}{2}(1 - \eta)$, is below that in the first best (i.e. $\frac{1}{2}$).

The following proposition states that the spot market is the only feasible equilibrium in the case without commitment:

Proposition 7 *Insurance is ruled out if suppliers cannot commit to a specific buyer.*

The proof of this proposition applies an additional non-shirking condition. If suppliers cannot commit to deliver all output to a specific buyer, this condition prevents the supplier from selling (part of) the produced output on the spot market:

N.s.c. III: take the output and sell it on the market

$$\forall x^m > 0 : y_n^*[x^* - x^m, e] + x^m \leq y_n^*[x^*, e]$$

where x^m denotes the output that is sold by the supplier on the spot market. By dividing through by x^m and taking the limit of $x^m \rightarrow 0$, we can write the non-shirking condition above as $\frac{dy_n^*[x, e]}{dx} \geq 1$. The piece rate should always be greater or equal to the market

price so that the non-contractible payment exhibit as least as much variation as output. This rules out insurance. \square

Insurance demands some form of commitment from the supplier to transfer all output to the buyer. Since the supplier is not rationed, he does not lose future rents by renegeing on this commitment. Hence, he has no incentives to comply. Anticipating this behavior, the buyer does not offer any insurance in the first place. In the absence of commitment, therefore, the spot market is the equilibrium outcome for all information modes so that the quality of information is irrelevant for the outcome.

5 Equilibrium if suppliers can choose to commit

The previous two sections treated commitment and non-commitment of output as an exogenous condition beyond the control of the supplier and the buyer. In practice, the agents can invest in commitment. To illustrate, building a factory hall allows the employer to prevent workers from selling the produced output to a third party on the outside market. By enforcing the property rights of the buyer, a factory hall can thus be viewed as an institutional device that allows suppliers to commit their output to a specific buyer. In some industries, a factory hall may not be needed to enforce this commitment. In the case of consultancy and other professional services, for example, a legal contract stipulating that employees are not allowed to provide services on their own account may suffice. The visibility of these particular services makes such a contract enforceable even though production does not occur in a factory hall. The commitment to refrain from selling output on the outside market does not make the quality of the output verifiable. It rules out only that suppliers obtain an alternative income from their effort by selling their output elsewhere but it does not necessarily imply that the courts can verify the quantity of the produced output.

This discussion results in the following definition of a firm:

Definition 8 *A firm is a physical or contractual arrangement between a supplier and buyer that enforces the property rights of the buyer over all output of the effort by the supplier. A firm thus allows a supplier to commit to the delivery of all output of his effort to one specific buyer.*

Appropriability of output by the buyer provides a more convincing rationale for the existence of firms than reputations based on implicit contracts alone. Indeed, also many transactions on the spot market are based on non-verifiable obligations supported by

reputation. Hence, reputations do not require firms. Having defined firms as devices to commit the delivery of output, we shall in the rest of this paper refer to the commitment case as the firm. Likewise, having established that the spot market is the only feasible outcome in the non-commitment case, we associate this case with the spot market. The previous two sections thus discussed the outcomes for the firm and the spot market, respectively. The comparison of these two cases allows us to explore whether the firm survives competition from the spot market if suppliers can choose whether to commit output or not. Competition forces rationed buyers to accept the appropriability mode preferred by suppliers. Accordingly, in order to explore the competitiveness of the firm, we compare suppliers' utility in the firm (as contained in Table 1 for various information modes) with that on the spot market (as discussed in Section 4).

Proposition 9 *In the presence of some verifiable information and risk-averse suppliers, the firm is always more competitive than the spot market.*

The proof follows by comparing u^s for the modes featuring some form of verifiable information in Table 1 (i.e. $\{e, x\} = (v, v), (v, nv), (no, v), (nv, v)$) with suppliers' utility $\frac{1}{2}(1 - \eta)$ in the spot market. \square

The intuition behind this proposition is that the firm allows a welfare improving transfer of risk from the supplier to the buyer if information is verifiable. Some small transfer of risk is always efficient, even though this trade in risk harms the incentives for effort if only output is verifiable: a small amount of supplier insurance yields first-order welfare gains while the welfare losses on account of lower incentives are only second order.

This proposition has important implications for the theory of the firm because it implies that a firm using an explicit piece contract always dominates the spot market. This limits the relevance of the standard principal agent model for exploring the boundaries of the firm vis-a-vis the spot market. In fact, spot markets can survive only under one of three conditions:

- the enforcement of the delivery of output within a firm requires costly investments;
- writing an explicit contract is costly;
- both output and effort are not verifiable.

The costly investments mentioned in the first condition may involve building a factory hall. We return to these investments at the end of this section. The second

condition is closely related to the third condition: an explicit contract is costly because it requires resources to be invested in the verifiability of the information. Non verifiable information (or costly verifiable information) thus is the key to understanding why spot markets may be more competitive than firms are.

We now turn to the case in which verifiable information is absent.

Proposition 10 *In the absence of any verifiable information and in the presence of risk-averse suppliers, the firm is more competitive than the spot market if either suppliers are sufficiently risk averse or the non-verifiable information is of sufficient quality.*

The proof follows from Table 1: suppliers' utilities u^s in the information modes $\{e, x\} = (nv, nv), (no, nv)$ exceed suppliers' utility $\frac{1}{2}(1 - \eta)$ in the spot market if $\chi \rightarrow 0$ or $\eta \rightarrow 1$. \square

In contrast to the case with explicit insurance contracts, the costs to the supplier of implicit insurance contracts are first order even if only small amounts of insurance are provided. Indeed, these costs are related to the incentive payments (which must be made credible by transferring rents to the buyer) rather than to the insurance provided. Hence, the spot market is more competitive than the firm if risk aversion is small, so that the costs to the supplier of the implicit contract outweigh the gains from insurance. Intuitively, the price buyers charge to make incentive payments credible exceeds the maximum insurance premium suppliers are willing to concede in order to be shielded from spot market risk.¹⁰

Shapiro and Stiglitz (1984) and MacLeod and Malcolmson (1989) analyze the commitment case with risk neutral suppliers and non verifiable effort (and output). In that setting, the relationshipspecific nature of output is exogenous and is problematic because it gives rise to agency and hold up problems that can be addressed only through costly implicit contracts associated with the underutilization of productive resources (either unemployment or vacancies). Suppliers would be unambiguously better off if they could get rid of their obligation to supply the output to the firm so that they could sell output to the highest bidder. In this way, competition would eliminate the incentive problems, thereby eliminating the underutilization of resources. Allowing for risk averse suppliers, however, proposition 10 reveals that commitment within a firm is not necessarily less

¹⁰Even though it employs non verifiable information on effort, the mixed piece rate contract (corresponding to the information mode $\{e, x\} = (nv, v)$) always survives competition with the spot market due to Proposition 9. With verifiable information on output, small amounts of insurance yield only second-order losses in suppliers' welfare even if the information on effort is poor.

desirable for suppliers than the spot market, even though information is non verifiable so that explicit contracts can not address incentive problems associated with commitment. Although the firm implies the absence of competition during the transaction period, it may nevertheless be a blessing rather than a curse because it allows suppliers to insure themselves. Competition between various modes of organizing production may thus induce investments in the limitation of competition by making output relationships specific within the firm.

Coase (1937) argued that the firm would emerge in equilibrium only if the spot market suffers from substantial transaction cost. Our model, in contrast, implies that the firm is more competitive than the spot market only if the firm does not suffer from excessive contracting problems due to the low quality information on either effort or output. Thus, we reverse the Coasian perspective on the theory of the firm. The firm survives the competition from the spot market only if workers (i.e. the suppliers) are willing to pay a large enough insurance premium to the employer (i.e. the buyer) to make the insurance contract credible. The value of the firm is equal to the discounted value of these insurance premiums. It is this value that is traded on the stock market. Our model thus explains why the stock market value of firms may substantially exceed the tangible assets owned by these firms.

Figure 2 shows graphically how risk aversion and the quality of information impact the competitiveness of the firm in the absence of verifiable information. The diagonal, $\chi = \eta$, represents the parameter combinations where suppliers are indifferent between the firm and the spot market for the information mode $\{e, x\} = (nv, nv)$. To the north west of this diagonal, the risk aversion parameter η exceeds the parameter indicating the lack of verifiability of effort, χ . In this area, the risk premium the suppliers are willing to concede to be shielded from market risk exceeds the price they have to pay the firm in order to make incentive payments credible. Hence, the firm is more competitive than the spot market in this area.

The dotted curve, $\chi = \frac{\eta^2}{1-\eta(1-\eta)}$, represents the parameter combinations where suppliers are indifferent between the firm and the spot market for the information mode $\{e, x\} = (no, nv)$. Again, to the north west of the curve, the firm is more competitive than the spot market. This area is smaller than the corresponding area for the information mode $\{e, x\} = (nv, nv)$. Intuitively, compared to the case in which it observes also effort, the firm can offer the worker less insurance for the same costly incentive payments. With insurance thus becoming more expensive, the supplier finds the firm less attractive compared to the spot market.

The following proposition characterizes the loci for which the firm and the spot market are equally competitive:

Proposition 11 *In the absence of verifiable information, the utility of the buyer, the equilibrium level of effort and the allocation of risk are discontinuous at the point where the firm and the spot market are equally competitive. In particular, in the spot market the utility of the buyer is lower, while effort and the risk borne by the supplier are higher.*

The proof follows by comparing the equilibrium outcomes for the spot market with those for the information modes $\{e, x\} = (nv, nv), (no, nv)$ in Table 1 at the parameter combinations of η and χ where suppliers' utility $\frac{1}{2}(1 - \eta)$ in the spot market equals suppliers' utility under the two information modes in Table 1. \square

At the point where the firm is as competitive as the spot market, the level of the buyer's utility is discontinuous. This is closely related to Proposition 4: whereas the buyer obtains his outside option on the spot market, a contract relying on non verifiable information is credible only if the buyer enjoys a surplus. Hence, moving from the firm to the spot market, the buyer experiences a drop in utility. Suppliers' utility, in contrast, remains constant, but its composition changes discontinuously. In particular, higher expected monetary incomes compensate for a loss in insurance as suppliers no longer have to transfer part of the surplus to the buyer to make the firm's incentive payments credible. Without these costs (to the supplier) of providing incentive payments, effort jumps upward to its first best level; compared to the provision of additional effort within a firm, the provision of additional effort is less costly to the supplier because it does not require the supplier to transfer more surplus to the buyer. The costs of implicit incentive payments explains why effort in the firm cannot exceed that in the market.

These discrete regime switches offers an explanation for the observation by Krueger(1991) that two rather different transaction modes coexist in the fast food industry, namely low powered incentive schemes with a fixed hourly wage (i.e. the fixed remuneration contract) and high powered incentive scheme where the owner of a franchise absorbs all the establishment specific risk (i.e. the spot market). Apparently, this industry operates around the point at which an implicit insurance contract is about as competitive as the spot market. Our model confirms the popular judgments about fixed remuneration contracts versus franchises: fixed remuneration contracts provide less incentives but more insurance than franchises do. Hence, workers provide more effort in franchises.

The introduction of a firm to enforce the property rights of the buyer on the output may require costly investments in the organization of production, for example the

construction of a factory hall. Comparing suppliers' utilities for the firm and the spot market, one can compute the competitive value of such investments.

Proposition 12 *The competitive value of investments in a firm is non decreasing in risk aversion and the quality of information.*

Compute the difference between u^s in Table 1 (for various information modes) and suppliers' utility without commitment (i.e. $\frac{1}{2}(1-\eta)$) and define $\hat{u}^s \equiv \max\{u^s - \frac{1}{2}(1-\eta), 0\}$ as the utility gain from the option to create a firm. Taking the first derivative of \hat{u}^s with respect to η and χ , one finds $\frac{\delta \hat{u}^s}{\delta \eta} \geq 0$ and $\frac{\delta \hat{u}^s}{\delta \chi} \geq 0$. \square

The intuition behind this proposition is that the benefits of a firm rise with risk aversion while the costs of a firm decline with quality of information. In particular, investments in commitment within a firm allow insurance and are thus especially valuable if suppliers are risk averse. Moreover, the incentive problems associated with this commitment become less serious if better information allows for better contracts. Investments in the quality of information and commitment are thus complements: better information makes the firm more attractive (see Proposition 3) while commitment within a firm makes good information important so that contracts can alleviate incentive problems within substantial transaction costs.

6 Outside option: no trade

The previous section showed that insurance harms effort. This result may no longer hold if suppliers can opt for a third alternative in addition to the firm and spot market, namely a riskless outside option of enjoying leisure so that neither production nor trade occur (i.e. $x_t = y_t = 0$). The utility value of leisure amounts to $\frac{1}{2}\beta$, $0 \leq \beta \leq 1$.¹¹ The model is thus fully parameterized by the quality of the outside option β , the degree of risk aversion η , the information modes on effort and output, and the length of the transaction period as indicated by χ . Figure 3 graphically compares the competitiveness of the various transaction modes. The various indifference lines depict, for each information mode (and assuming a particular value for χ), the combinations of β and η for which the supplier is indifferent between, on the one hand, the firm, and, on the other hand, either

¹¹There is a difference between engaging in production and supplying no effort (i.e. $e_t = 0$) and taking advantage of the outside option. If working at $e_t = 0$, the risk-averse supplier is still subject to production risk z_t . If enjoying leisure, in contrast, the supplier collects a risk free payoff $\frac{1}{2}\beta$. The choice to enjoy leisure is assumed to be verifiable.

the spot market or no trade. In addition, the broken 45° degree line MN ($\eta = 1 - \beta$) divides the plane into an area below the diagonal where the spot market dominates no trade and an area above the diagonal where no trade dominates the spot market.

For the four information modes involving verifiable information, the firm always dominates the spot market, see Proposition 9. Hence, only the indifference loci between the firm and no trade are relevant. The firm dominates no trade in the area to the south-west of the loci. The areas are larger than the area where the spot market dominates no trade, consistent with the observation that the firm dominates the spot market for these information modes. The vertical line MP ($\beta = 1$) represents the indifference locus for the information modes $\{e, x\} = (v, v), (v, nv)$. Since full insurance is first best, this information mode yields the largest area in which production dominates no trade. MS ($\eta = \frac{1-\beta}{\beta}$) depicts the locus for the information mode $\{e, x\} = (no, v)$, which yields an explicit piece rate contract. The area to the south west of this curve is the smallest of all verifiable information modes since it yields the lowest utility for the worker. The indifference curve MQ ($\eta = \frac{(1-\beta)\chi}{\beta-(1-\chi)}$), which corresponds to $\{e, x\} = (nv, v)$, lays in between MP and MS .

In the two information modes employing only non-verifiable information, the firm is subject to competition not only from the east (no trade), but also from the south (the spot market). The square in the north west delineated by the lines VR ($\beta = 1 - \chi$) and VX ($\eta = \chi$) represents the area where the firm is the most competitive for the information mode $\{e, x\} = (nv, nv)$, yielding a fixed remuneration contract. VR and VX are the loci of indifference of the firm with no trade and the spot market, respectively. The parameter space is thus divided into three parts corresponding to production within the firm (the square $NRVX$), production for the spot market (the area $VMOX$), and the absence of any trade (the area $RPMV$). In the area RNV , production would not occur without firms facilitating insurance. The insurance provided by the firm encourages suppliers to deliver effort in a risky environment. In this way, the invention of the firm has stimulated production. Hence, contrary to the conclusion of the previous section which compared the firm to the spot market only, the firm might actually raise effort.¹² If quality of information improves (and thus χ declines), the area RNV where the firm enhances trade becomes larger.

If only output is observable (i.e. $\{e, x\} = (no, nv)$), the area where the firm is most

¹²Assuming that production risk increases with effort, Sinn (1995) argues that insurance may enhance incentives. We arrive at similar conclusions as Sinn (1995) in a model in which production risk does not rise with the level of effort.

competitive, $NTWY$, is smaller compared the case in which also effort is observable. This is because non-verifiable information on output allows the firm to offer less insurance than non-verifiable information on effort. Nevertheless, also under this information mode, the firm opens up the possibility of production in the area NTW , where otherwise suppliers would prefer not to produce at all.

Whereas proposition 11 investigates the loci of indifference of the firm with the spot market, the following proposition analyzes the loci of indifference of the firm with no trade.

Proposition 13 *If the contract operated by the firm employs non-verifiable information, the utility of the buyer is discontinuous at the point where the supplier is indifferent between the firm and no trade, being larger for the firm.*

The proof is similar to that of Proposition 11. \square

Propositions 11 and 13 can be used to explore the welfare implications of competition between the spot market, no trade and the firm.

Proposition 14 *If the contract operated by the firm employs non-verifiable information, competition between the firm, the spot market, and no trade may harm welfare by crowding out the firm.*

This proposition is a direct implication of the discontinuity of u^d and the continuity of u^s at the loci of indifference between the firm and the spot market (see Proposition 11) and between the firm and no trade (see Proposition 13). As the firm is replaced by the spot market or no trade, the drop in buyer's utility is first order while the change in supplier's utility is only second order. \square

Giving the supplier the option to sell output on the spot market may involve costs. Proposition 14 suggests that society may overinvest in this option. The proposition is illustrated by considering the information mode $\{e, x\} = (nv, nv)$. As discussed in Section 5, suppliers are indifferent between the firm and the spot market for $\chi = \eta$. If χ marginally exceeds η , the supplier prefers the spot market above the firm, while the utility of the buyer jumps discontinuously from a positive value under the firm to zero in the spot market. Competition between the various transaction modes implies that the spot market crowds out the firm even though the latter yields higher overall welfare. This is because suppliers do not take into account the negative externality on buyers' rents of their decision to quit the firm and switch to the spot market. Whereas the utility gain for suppliers is only second order, the loss in rents accruing to the firm is first order.

7 Conclusions

We established a theory of the firm based on the imperfect ability of risk-averse suppliers to commit to an insurance contract in which they promise to deliver all output to the buyer. The firm is viewed as a device enabling the supplier to commit to this delivery, either by physical means, such as a factory hall, or by contractual means, as in case of consultancy firms. This commitment device introduces specificity in the relation between the supplier and the buyer. Rather than some exogenous technological requirement, specificity is thus deliberately introduced to make insurance feasible. In addition to the classical trade off between insurance and incentives popularized by the standard principal agent model, our model features a second trade off, namely that between insurance and surplus sharing. Implicit insurance contracts can be operated only if risk-averse suppliers concede part of the surplus to buyers. In the worst case, in which effort is non-observable and – unlike in the principal agent model – output is not verifiable, incentives are dulled for two reasons, namely to reduce both the risk assumed by risk-averse suppliers and the surplus that must be conceded to the buyer. Interestingly, we established that non-verifiability of information can be viewed as an intermediate case between non-observability and verifiability, see Proposition 1.

Our analysis uncovered a fundamental difference between the "principal-agent" trade off between insurance and incentives and the "non-verifiability" trade off between insurance and surplus sharing. The cost to the supplier of a small amount of insurance is only second order in the principal agent model, whereas it is first order in the non-verifiability model. As a direct consequence, explicit insurance contracts based on verifiable information always dominate the spot market, see Proposition 9. Indeed, the spot market can dominate the firm only if information is non-verifiable, the commitment device is costly, or explicit contracts are associated with transaction costs. The transition from the spot market to a firm based implicit contract, in contrast, is discontinuous, see Proposition 11. Whereas the aggregate pay off of the supplier moves continuously, its composition changes discontinuously. Incentives and hence effort fall, while insurance jumps up, confirming the popular judgment regarding the implications of firm based employment relations versus self employment. The overall pay off of the buyer jumps up when moving from the spot market to the implicit contract. This regime switch makes the lack of verifiability model more suitable for analyzing the nature of the firm vis-a-vis the market than the principal agent model.

In the principal agent model, the contract formulated by transaction partners max-

imizes wealth conditional on the incentive constraint of the supplier. In the non-verifiability model, in contrast, this is not generally the case. The reason is that competition between buyers implies that the equilibrium contract maximizes the suppliers' pay off. Competition thus disregards external effects on the buyers' pay off, namely the surplus collected by buyers when employing an implicit contract. Accordingly, competition between the spot market and a firm using implicit contracts may crowd out the firm on account of a business stealing effect, thereby hampering wealth maximization, see Proposition 14.

We have taken the quality of information as exogenous. In reality, agents may invest in the quality of information so that the available information becomes endogenous. Our model can be fruitfully employed to analyze both the competitive and social value of such investments. When doing so, we encounter the same mechanism as in investigating the welfare implications of competition between the spot market and the firm. In particular, the competitive value of information accounts only for the effect of improved information on supplier's utility, while the social value of information includes also the impact on the buyer's pay off. The latter pay off can be viewed as resulting from an implicit "tax" imposed on the supplier by the buyer to make the latter's commitment to the insurance contract credible. Just as other, explicit tax revenues, the implicit tax revenues collected by the buyer exhibit the shape of a Laffer curve. In particular, for a small implicit tax associated with small verifiability problems, a marginal increase in the tax rate corresponding to a small increase in verifiability problems raises the revenues accruing to the buyer. Beyond a certain point, however, the verifiability problems become so grave that the adverse incentive effects of a higher tax rate on the tax base dominate the positive revenue effects of a further increase in the tax rate, see Proposition 4. Since suppliers ignore these effects on the rents (i.e. tax revenues) collected by buyers and since the actual contract maximizes suppliers utility due to competition between buyers, society underinvest in information if the initial information is poor. If information is good to start with, in contrast, too many resources are spend on enhancing the quality of information further, see Proposition 6. Intuitively, suppliers improve information to avoid having to pay implicit taxes. Hence, additional information redistributes rents instead of creating additional value for society at large. The overall welfare effect of more juridification (i.e. more explicit contracting based on better, namely verifiable, information replacing implicit contracts based on long-term relationships) is therefore ambiguous: it may either create value by facilitating trade or destroy value by using scarce resources to redistribute rents. The latter effect most likely dominates if most potential trade has

already been exploited by implicit contracts based on good information.

8 References

Allen, D., and D. Lueck, 1999, 'The Role of Risk in Contract Choice,' *Journal of Law, Economics, and Organization*, Vol. 15, pp. 7-4-736.

Baker, G., M. Gibbs, and B. Holmstrom, 1994, 'Subjective Performance Measures in Optimal Incentive Contracts,' *Quarterly Journal of Economics*, Vol. 109, pp. 1125-1156.

Coase, R., 1937, 'The Nature of the Firm,' *Economics N.S.*, Vol. 4, pp. 386-392.

Greif, A., 1993, 'Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition,' *American Economic Review*, Vol. 83, pp. 525-546.

Greif, A., P.R. Milgrom and B.R. Weingast, 1994, 'Coordination, Commitment, and Enforcement: The Case of the Merchant Guild,' *Journal of Political Economy*, Vol. 94, pp. 745-776.

Holmstrom, B., and P. Milgrom, 1987, 'Aggregation and Linearity in the Provision of Intertemporal Incentives,' *Econometrica*, Vol. 55, pp. 303-328

Holmstrom, B., and P. Milgrom, 1991, 'Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership, and Job Design,' *The Journal of Law, Economics and Organization*, pp. 25-52.

Krueger, A.B., 1991, 'Ownership, Agency and Wages: An Examination of Franchising in the Fast Food Industry,' *Quarterly Journal of Economics*, Vol. 106, No. 1, pp. 75-101.

MacLeod, W.B., and J.M. Malcomson, 1989, 'Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment,' *Econometrica*, Vol. 57, pp. 447-480.

Marglin, S.A., 1974, 'What do Bosses do? The Origins and Functions of Hierarchy in Capitalist Production,' *Review of Radical Political Economics*, Vol. VI, no. 2, pp. 60-112.

Mirrlees, J.A., 1976, 'The Optimal Structure of Incentives and Authority within an Organization,' *Bell Journal of Economics*, Vol. 7, pp. 105-131.

Shapiro, C., and J.E. Stiglitz, 1984, 'Equilibrium Unemployment as a Worker Discipline Device,' *American Economic Review*, Vol. 74, pp. 433-444.

Shleifer, A., and L.H. Summers, 1988, 'Breach of Trust in Hostile Takeovers,' in A.J. Auerbach (ed.), *Corporate Takeovers: Causes and Consequences* (University of Chicago Press, Chicago)

Sinn, H.W., 1995, 'Social Insurance, Incentives and Risk Taking,' National Bureau of Economic Research Working Paper No. 5335.

Appendix The derivation of contracts

8.0.1 Mode $\{e, x\} = (v, v)$

Competition between infinitely elastic buyers ensures that the buyers' participation constraint (use (5)) $u^d = \frac{1}{\varepsilon}E[e - y] \geq 0$ is binding so that $E[y] = E[e] = e$. Optimizing expression (3) with respect to e^* subject to the participation constraint $E[y] = e^*$, we arrive at $e^* = \varepsilon$. \square

8.0.2 Mode $\{e, x\} = (no, v)$

This information mode is assumed in the standard principal agent model. Holmstrom and Milgrom (1987) show that the optimal contract in this case is linear in output: $y = a + bx$.¹³ Accordingly, the pay off of suppliers r^s is distributed normally. Substituting $y = a + bx$ into (3) and maximizing the resulting expression with respect to e , we arrive at $e = b\varepsilon$. The participation constraint of the buyer is binding due to competition between buyers, $u^d = \frac{1}{\varepsilon}E[e - y] = 0$. This implies that $a = (1 - b)e = (1 - b)b\varepsilon$ and $y = e + bz = b(\varepsilon + z)$. Substituting this expression for y and $e = b\varepsilon$ into equation (3) to eliminate y and e and maximizing the resulting expression with respect to b , we find $b^* = \frac{1}{1+\eta}$ and $e^* = \frac{\varepsilon}{1+\eta}$. \square

8.0.3 Mode $\{e, x\} = (nv, nv)$

Since neither effort nor output is verifiable, the contractible payment y_c has to be constant. A supplier who shirks on his non-verifiable obligations thus lacks incentives to provide any effort. With $e_{sh} = 0$, the non-shirking condition for the supplier (6) simplifies to

$$\forall z : y_n^* \geq \frac{1}{2\varepsilon}e^{*2}. \quad (12)$$

¹³Whereas Holmstrom and Milgrom (1987) assume that suppliers learn about the state of nature during the course of the transaction period, we assume that suppliers know the state of nature already at the beginning of the transaction period. However, suppliers cannot exploit this additional information because the state of nature does not affect the optimal effort level under a linear contract. Hence, suppliers select exactly the same effort level as under the Holmstrom and Milgrom assumption.

Substituting (12) into the non-shirking condition for the buyer (10) to eliminate y_n^* , we arrive at

$$\forall z : \frac{1}{2\varepsilon}e^{*2} \leq \frac{1-\chi}{\chi} \mathbb{E} \left[e^* - y_c^* - \frac{1}{2\varepsilon}e^{*2} \right]. \quad (13)$$

Competition between buyers makes this constraint binding for all z , so that e^* and therefore y_n^* do not depend on z . Since e^* , y_c , and y_n^* are non-stochastic, we can apply (3):

$$\varepsilon u^s = y_c^* + y_n^* - \frac{1}{2\varepsilon}e^{*2}. \quad (14)$$

Substitution the binding versions of (12) and (13) into (14) to eliminate y_n^* and y_c^* respectively yields

$$u^s = \frac{e^*}{\varepsilon} - \frac{1}{2} \frac{1}{1-\chi} \left(\frac{e^*}{\varepsilon} \right)^2.$$

Competition between buyers ensures that the equilibrium contract maximizes u^s with respect to e^* . This results in $e^* = (1-\chi)\varepsilon$. \square

8.0.4 Mode $\{e, x\} = (no, nv)$

We consider a truncated implicit linear contract $y_n = a + b\bar{x}$, where $\bar{x} \equiv \max[x^+, \min[x, x^-]]$. Hence, the minimum and maximum bonus are given by, respectively, $\min_z [y_n] = a + bx^-$ and $\max_z [y_n] = a + bx^+$. Define: $z^{+,-} \equiv x^{+,-} - \varepsilon b$. The non-negativity constraint on y_n implies:

$$y_n \geq a + bx^- \geq 0$$

Competition between buyers makes this constraint binding so that $a = -bz^- - \varepsilon b^2$. Consider a supplier who decides on effort e facing this contract. If the decision is unaffected by the non-linearities in the contract, the optimal level is $e^* = \varepsilon b$. Whenever the non-verifiable payment is less than the cost of effort, $y_n[z] < \frac{1}{2\varepsilon}e^{*2} = \frac{1}{2}\varepsilon b^2$, the supplier is better off by not providing any effort. Let z^0 denote the value of z for which this inequality is just satisfied, $z^0 \equiv z^- + \frac{1}{2}\varepsilon b$. For any $z < z^0$, the supplier provides no effort (i.e. $e = 0$) because the effort cost dominates the non-verifiable bonus. Likewise, y_n can never be larger than $a + bx^+$. Hence, whenever $z^+ < z < z^+ + \varepsilon b$, the supplier sets effort below εb so that y_n is exactly equal to its maximum: $e = \varepsilon b - (z - z^+)$. For $z^+ + \varepsilon b \leq z$, the supplier optimally provides no effort, $e = 0$, since negative effort would be costly. Hence:

z	e	x	y_n
$(-\infty, z^0]$	0	z	0
$(z^0, z^+]$	εb	$\varepsilon b + z$	$\frac{1}{2}\varepsilon b^2 + b(z - z^0)$
$(z^+, z^\wedge]$	$\varepsilon b - (z - z^+)$	$\varepsilon b + z^+$	$\frac{1}{2}\varepsilon b^2 + b(z^+ - z^0)$
(z^\wedge, ∞)	0	z	$\frac{1}{2}\varepsilon b^2 + b(z^+ - z^0)$

where: $z^\wedge \equiv z^+ + \varepsilon b$ and:

$$\mathbb{E}[x] = \varepsilon b (\Phi^\wedge - \Phi^0) + z^+ (\Phi^\wedge - \Phi^+) - (\phi^+ - \phi^\wedge) \quad (15)$$

$$\mathbb{E}[y_n] = b(-z^- (1 - \Phi^-) + z^+ (1 - \Phi^+) + (\phi^0 - \phi^+)) \quad (16)$$

$$y_n[0] = -bz^- \quad (17)$$

where $\phi[\cdot]$ and $\Phi[\cdot]$ are the normal density and distribution function and ϕ^i and Φ^i denote $\phi[z^i]$ and $\Phi[z^i]$ for $i = -, 0, +, \wedge$. For the derivation of the expression for $y_n[0]$, we use $z^0 < 0 < z^+$.¹⁴

The non-shirking condition for the buyer (10) must be met for the largest non-verifiable bonus:

$$\max_z [y_n] = b(z^+ - z^0) + \frac{1}{2}\varepsilon b^2 \leq \frac{1 - \chi}{\chi} (\mathbb{E}[x] - \mathbb{E}[y_n] - y_c) \quad (18)$$

The utility of suppliers can be written as:

$$\varepsilon u^s = (y_c + y_n[0]) - \frac{1}{2}\varepsilon b^2 - \frac{1}{2}\theta b^2 \Psi[z^0, z^+, \varepsilon, \theta, b] \quad (19)$$

where:

$$\Psi[\cdot] \equiv \frac{2}{\theta^2 b^2} \ln \left[\int_{-\infty}^{z^0} \vec{\phi} e^{-\theta b z^0} dz + \int_{z^0}^{z^+} \vec{\phi} e^{-\theta b z} dz + \int_{z^+}^{z^\wedge} \vec{\phi} e^{-\theta(bz^+ + Q)} dz + \int_{z^\wedge}^{\infty} \vec{\phi} e^{-\theta b z^\wedge} dz \right]$$

$$Q \equiv -\frac{1}{2\varepsilon} \left((\varepsilon b - z + z^+)^2 - \varepsilon^2 b^2 \right) \geq 0$$

and $\vec{\phi} \equiv \phi[z]$. $\frac{1}{2}\theta^2 b^2 \exp[\Psi]$ is the expected value of the exponent of a truncated transformation of $-\theta b z$. Since $\mathbb{E}[e^{-\theta b z}] = \exp[\frac{1}{2}\theta^2 b^2]$, we have $\Psi[-\infty, \infty, \cdot] = 1$. The term Q reflects the utility gain obtained by spending less effort when z is above the upper threshold. Using expression (10) to eliminate y_c from (19), we arrive at:

$$\varepsilon u^s \leq (\mathbb{E}[x] + y_n[0] - \mathbb{E}[y_n]) - \frac{1}{2} \frac{1}{1 - \chi} \varepsilon b^2 - \frac{\chi}{1 - \chi} b (z^+ - z^0) - \frac{1}{2} \theta b^2 \Psi[z^0, z^+, \cdot]$$

¹⁴The solutions of the first-order conditions for z^0 and z^+ derived below for $\lim_{\varepsilon \rightarrow \infty}$ are of the form $(z^{0,+})^2 = \text{constant}$, see equation (21). These equations have a positive and a negative root. Hence, the sign of z^0 is determined by the second-order condition. Intuitively, the optimal contract subjects to the piece rate the part of the support of z that features the highest density, i.e. the area around the median $z = 0$.

Substitution of (15), (16), and (17) to eliminate, respectively, $E[x]$, $E[y_n]$, and $y_n[0]$ yields:

$$\begin{aligned} u^s \leq & b(\hat{\Phi} - \Phi^0) + \frac{z^+}{\varepsilon}(\hat{\Phi} - (1-b)\Phi^+ - b) - \frac{b}{\varepsilon}z^-\Phi^0 \\ & + \frac{1}{\varepsilon}(\hat{\phi} - (1-b)\phi^+ - b\phi^0) \\ & - \frac{1}{2} \frac{1}{1-\chi} b^2 - \frac{\chi}{1-\chi} \frac{b(z^+ - z^0)}{\varepsilon} - \frac{1}{2} \eta b^2 \Psi[z^0, z^+, \cdot] \end{aligned} \quad (20)$$

The optimal truncated linear contract maximizes u^s . Hence, the following first order conditions must apply (where we take into account $z^{\hat{}} = z^+ + \varepsilon b$, $z^- = z^0 - \frac{1}{2}\varepsilon b$ and $d\phi^i/dz^i = -z^i\phi^i$, $i = 0, +, \hat{}$):

$$\begin{aligned} \frac{du^s}{dz^+} &= b\hat{\phi} + \frac{1}{\varepsilon}(\hat{\Phi} - (1-b)\Phi^+ - b) + \frac{\hat{\phi}}{\varepsilon}(z^+ - z^{\hat{}}) - \frac{\chi_x}{1-\chi_x} \frac{b}{\varepsilon} \\ &+ \frac{b}{\varepsilon\bar{\Psi}} \left(\int_{z^+}^{z^{\hat{}}} \left(1 - \frac{1}{\varepsilon b}(\varepsilon b - z + z^+)\right) \vec{\phi} e^{-\theta(bz^+ + Q)} dz + \int_{z^{\hat{}}}^{\infty} \vec{\phi} e^{-\theta bz^{\hat{}}} dz \right) \\ \frac{du^s}{dz^0} &= -b\phi^0 - \frac{b}{\varepsilon}\Phi^0 + \frac{b\phi^0}{\varepsilon}(z^0 - z^-) \\ &+ \frac{\chi}{1-\chi} \frac{b}{\varepsilon} + \frac{b}{\varepsilon\bar{\Psi}} \int_{-\infty}^{z^0} \vec{\phi} e^{-\theta bz^0} dz \end{aligned}$$

where $\bar{\Psi} \equiv \exp[\frac{1}{2}\theta^2 b^2 \Psi]$. Let the functions $z^0[\varepsilon]$ and $z^+[\varepsilon]$ denote the values of z^0 and z^+ that solve these first-order conditions, as a function of ε . These functions satisfy:

$$\lim_{\varepsilon \rightarrow \infty} \varepsilon \phi^0 = \lim_{\varepsilon \rightarrow \infty} \varepsilon \hat{\phi} = \frac{\chi}{1-\chi} \quad (21)$$

which follows from multiplying both first-order conditions by ε and dropping the vanishing terms (using $\lim_{z \rightarrow \infty} z\phi(z) = 0$). Employing the definition of $\phi(z)$, we find from (21) that $z^{0,+}[\varepsilon]$ decrease and increase in the limit proportionally to $\sqrt{\ln \varepsilon}$, which implies that $\lim_{\varepsilon \rightarrow \infty} \frac{z^{0,+}}{\varepsilon} = 0$. Using this latter result in (20), we arrive at:

$$\lim_{\varepsilon \rightarrow \infty} u^s \leq b - \frac{1}{2} \frac{1}{1-\chi_x} b^2 - \frac{1}{2} \eta b^2$$

Due to competition, this inequality is binding. The value of b that maximizes $\lim_{\varepsilon \rightarrow \infty} u^s$ thus satisfies $b = \frac{1-\chi}{1+(1-\chi)\eta}$. \square

8.0.5 Mode $\{e, x\} = (nv, v)$

The explicit contract is linear in output $y_c = a + bx$.¹⁵ This explicit contract applies also if the supplier shirks on his non-verifiable obligations contained in the implicit contract.

¹⁵Like in the mode $\{e, x\} = (no, nv)$, we cannot prove that the linear form of the contract is optimal.

Hence, if shirking on his non-verifiable obligations, the supplier selects effort so as to maximize (substitute $y = a + bx$ into (3) to eliminate y) $\varepsilon u^s = a + be - \frac{\theta}{2}b^2 - \frac{1}{2\varepsilon}e^2$. This yields $e_{sh} = b\varepsilon$. Substituting this expression and $y_c = a + bx$ into the supplier's non-shirking condition (6) to eliminate e_{sh} and y_c , we find

$$\forall z : y_n^* \geq \frac{1}{2\varepsilon}e^{*2} - be^* + \frac{\varepsilon}{2}b^2. \quad (22)$$

Substitution of (22) into (10) to eliminate y_n^* yields

$$\forall z : \frac{1}{2\varepsilon}e^{*2} - be^* + \frac{\varepsilon}{2}b^2 \leq \frac{1-\chi}{\chi} \mathbb{E} \left[e^* - y_c^* - \frac{1}{2\varepsilon}e^{*2} + be^* - \frac{\varepsilon}{2}b^2 \right]. \quad (23)$$

Competition between buyers makes this constraint binding for all z , so that e^* and therefore y_n^* do not depend on z . Since e^* and y_n^* are thus non-stochastic and the linearity of the explicit contract implies that y_c^* is distributed normally with $\text{Var}[y_c^*] = b^2$, we can apply (3)

$$\varepsilon u^s = \mathbb{E}[y_c^*] + y_n^* - \frac{1}{2\varepsilon}e^{*2} - \frac{\theta}{2}b^2 \quad (24)$$

Substitution of the binding versions of (22) and (23) to eliminate y_n^* and $\mathbb{E}[y_c^*]$ from (24), we arrive at

$$u^s = \frac{e^*}{\varepsilon} - \frac{1}{2} \left(\frac{e^*}{\varepsilon} \right)^2 - \frac{1}{2}b^2\eta - \frac{\chi}{1-\chi} \left(\frac{1}{2} \left(\frac{e^*}{\varepsilon} \right)^2 - b\frac{e^*}{\varepsilon} + \frac{1}{2}b^2 \right)$$

Competition between buyers ensures that the equilibrium contract maximizes u^s with respect to the implicitly agreed effort level e^* and explicit piece rate b . This yields $b = \frac{\chi}{\chi+\eta}$ and $e^* = \left(1 - \frac{\eta}{\chi+\eta}\right)\varepsilon \square$

Table 1: Summary of various transaction modes

$\{e, x\}$	(v, v)	(v, nv)	(no, v)	(nv, nv)	(no, nv)	(nv, v)
contract	fi		epr	fr	ipr^{16}	mpr
u^s	$\frac{1}{2}$	$\frac{1}{2} \frac{1}{1+\eta}$	$\frac{1}{2}(1-\chi)$	$\frac{1}{2} \frac{1-\chi}{1+(1-\chi)\eta}$	$\frac{1}{2} \left(1 - \chi \frac{\eta}{\chi+\eta}\right)$	
u^d	0	0	$\frac{1}{2}\chi(1-\chi)$	$\frac{1}{2} \frac{\chi(1-\chi)}{1+(1-\chi)\eta}$	$\frac{\chi(1-\chi)}{2} \left(\frac{\eta}{\chi+\eta}\right)^2$	
$u^s + u^d$	$\frac{1}{2}$	$\frac{1}{2} \frac{1}{1+\eta}$	$\frac{1}{2}(1-\chi^2)$	$\frac{1}{2} \frac{1-\chi^2}{1+(1-\chi)\eta}$	$\frac{1}{2} \left(1 - \frac{\eta(\eta+1)\chi^2}{(\chi+\eta)^2}\right)$	
e^*/ε	1	$\frac{1}{1+\eta}$	$1-\chi$	$\frac{1-\chi}{1+(1-\chi)\eta}$	$1 - \chi \frac{\eta}{\chi+\eta}$	
b	0	$\frac{1}{1+\eta}$	0	$\frac{1-\chi}{1+(1-\chi)\eta}$	$\frac{\chi}{\chi+\eta}$	
$E[y_n^*]/\varepsilon$	0	0	$\frac{1}{2}(1-\chi)^2$	$\frac{1}{2} \frac{(1-\chi)^2}{1+(1-\chi)\eta}$	$\frac{(1-\chi)^2}{2} \left(\frac{\eta}{\chi+\eta}\right)^2$	

legend:

fi : full insurance, epr : explicit piece rate, fr : fixed remuneration, ipr : implicit piece rate, mpr : mixed piece rate, b : the piece rate

¹⁶This column assumes $\varepsilon \rightarrow \infty$ and $\eta = \theta/\varepsilon$ is constant.

Figure 1: The Relation between z and e, y_n

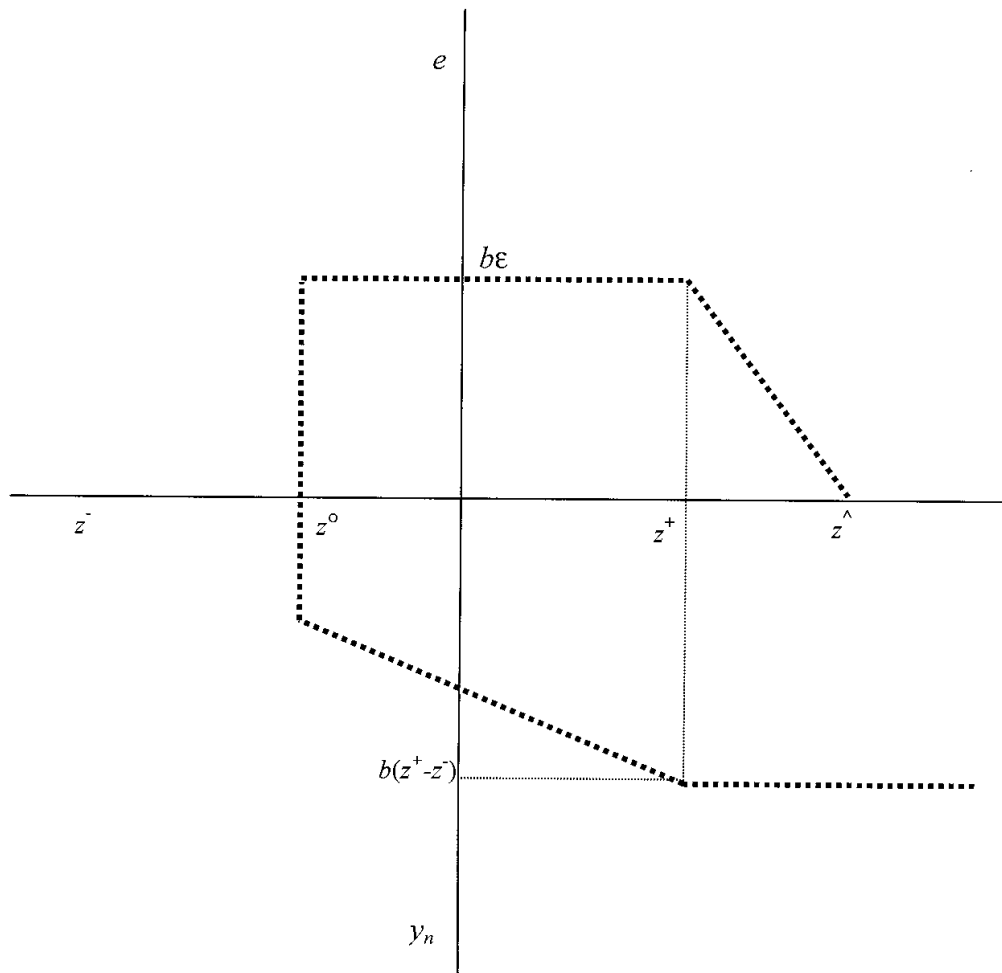


Figure 2: The Quality of Information versus Risk Aversion

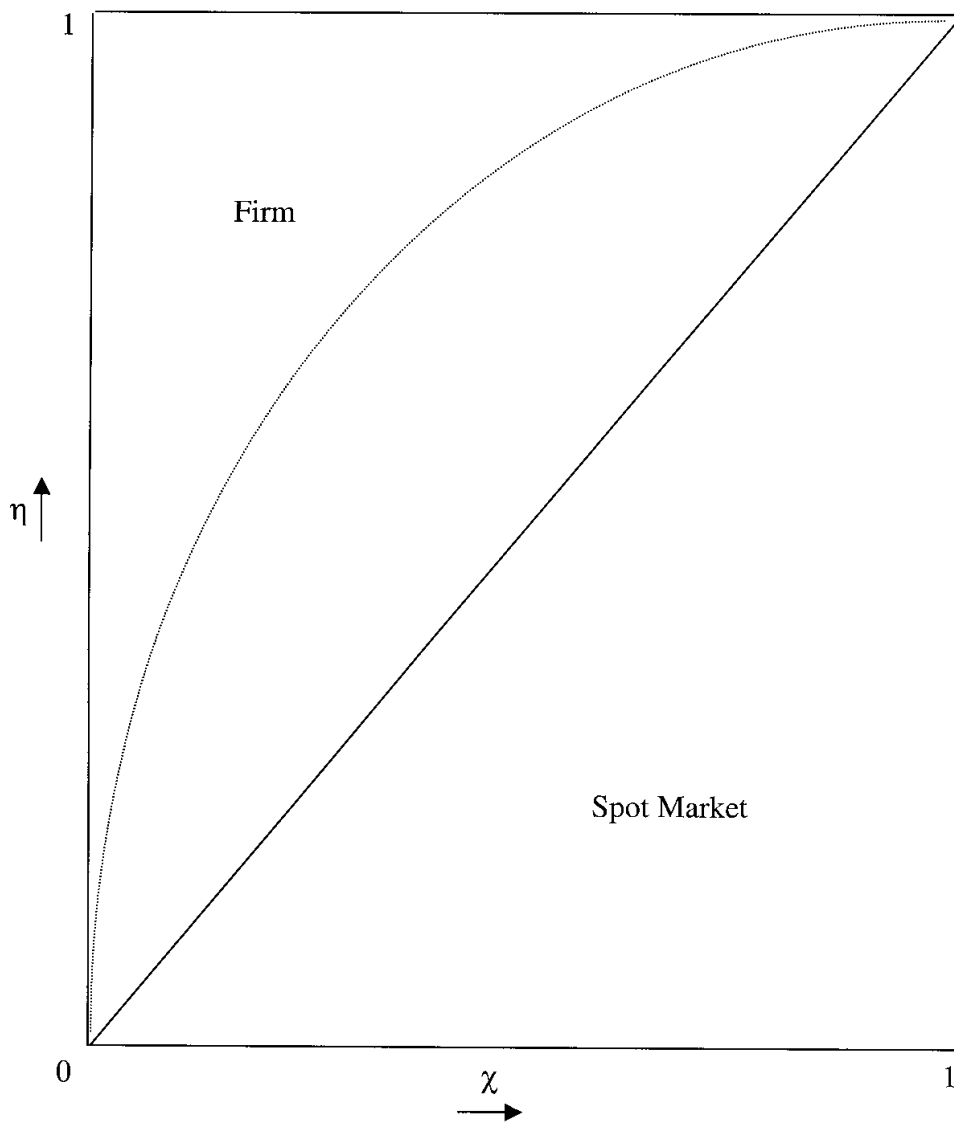


Figure 3: The Outside Option versus Risk Aversion

