Working Paper No. 52/03

Human Capital and Risk Aversion in Relational Incentive Contracts

by

Ola Kvaløy

SNF Project No. 7220
Gassmarkeder, menneskelig kapital og selskapsstrategier

The project is financed by The Research Council of Norway

FOUNDATION FOR RESEARCH IN ECONOMICS AND BUSINESS ADMINISTRATION
BERGEN, NOVEMBER 2003
ISSN 1503-2140
Human Capital and Risk Aversion in Relational Incentive Contracts

Ola Kvaløy*

Stavanger University College**

September 2003

Abstract: This paper examines a self-enforced relational incentive contract between a risk neutral principal and a risk averse agent where the agent’s human capital is essential in ex post realization of values. I analyse the effect of outside options on the optimal bonus level, showing how the presence of ex post outside options may impede desirable degrees of performance pay. The effect of risk aversion and incentive responsiveness is analysed by allowing for linear contracts. I show that the first order effect of these parameters are the same as in verifiable contracts, but second order effects show that the optimal bonus level’s sensitivity to risk aversion and incentive responsiveness increases with the discount factor. The analysis has interesting implications on firm boundaries and specificity choices.

* I thank Frøystein Gjesdal, Kjell Hausken, Petter Osmundsen and especially Trond E. Olsen for valuable comments and suggestions.
** Address: Stavanger University College, PO.Box 2557 Ullandhaug, NO-4004 Stavanger, Norway. Phone: 47 51 83 15 82. Fax: 47 51 83 15 50, E-mail: Ola.Kvaloy@oks.his.no
1. Introduction

The risk averse possessors of human capital experience a tight spot: as capital owners they are automatically exposed to the incentives of the market. But as opposed to the owners of physical capital, they cannot share the risk, and as risk averse agents they may prefer a secure employment relationship with high fixed salary and a low degree of performance pay. An optimal incentive contract will insulate the economic behaviour within the employment relationship from the temptations of the outside market. An optimal contract can ensure a wage scheme that optimally balance the need for incentives with the need for insurance, and the risk averse agent can enjoy a high degree of fixed salary, and a lower degree of performance pay.

But this is difficult. An incentive contract deterring any opportunistic behaviour must contain objective verifiable criteria that are enforceable by a court of law. In most employer-worker relationships, however, it is difficult to find objective verifiable performance measures. This is especially the case in human capital-intensive industries. It is complicated to verify the performance of a worker that creates values for the firm through the production of knowledge. Hence, verifiable contracts are seldom feasible. But relational contracts are always feasible, constrained though by the requirement of being self-enforcing. This constraint may impede the contract from implementing optimal solutions.

This paper studies a repeated employer-worker relationship where the worker uses his human capital in order to generate values for the employer. I model the employment contract within a repeated game framework where the present value of the ongoing relationship determines the players’ choice of Honouring or reneging on the contract. The model is in this respect similar to standard models of relational contracts (Klein and Leffler, 1981; Shapiro and Stiglitz, 1984; Bull 1987; Kreps 1990; Baker, Gibbons and Murphy 1994, 2002). MacLeod and Malcomson (1989) generalizes the case of symmetric information, while Levin (2003) makes a general treatment of relational contracts with both symmetric and asymmetric information, allowing for incentive problems due to moral hazard and hidden information.
To my knowledge, the present paper is the first to analyse relational contracts that includes both asymmetric information, in the form of unobservable effort, and risk aversion. It is complicated to make definite treatments of risk aversion in repeated game models of relational incentive contracts, but I allow for an approximation, studying repeated linear incentive contracts with bounded support on the noise-variable. This makes it possible to study the effect of risk aversion and incentive responsiveness within relational contracts with asymmetric information.

The model emphasizes the role of human capital. The challenge of contracting on human capital lies in the subtle balance between the residual control right of the worker and the authority of the employer. According to the standard view of ownership, it is the owner of an asset who has residual control right over the asset; that is “the right to decide all usages of the asset in any way not inconsistent with a prior contract, custom or law” (Hart, 1995). If the asset involved in the worker’s production is his own mind and knowledge; that is his own human capital, then he also is to decide all non-contractual usages. This complicates the very nature of the employment relationship, which can be seen as an implicit contractual transfer of residual control rights from the worker to the employer. Initially, ex ante any contractual relationship, the worker is a ‘free agent’ who can choose whatever behaviour he wants in order to manage his human capital. If the agent enters into an employment relationship, however, he accepts the employer to select his behavioural pattern. In other words: he accepts the employer to manage his human capital. The behavioural pattern or range of actions that the employer might require the worker to undertake is unclear and unspecified. Hence, for the employment contract to be meaningful, the employer has to be given some rights to decide non-contractual usages. But this right automatically conflicts with the residual control rights of ownership. Even if the worker accepts the employer to exercise authority, the worker still owns the asset in question, and thus has the residual control right of how to decide any non-contractual usage.

Analytically this problem can be solved by separating two types of rights that often is considered to be interlinked: the right to decide the management of the asset, and the right to decide the usage of the values created by that asset. In an employment relationship where the worker creates values for the firm with his human capital, the employer is given the right to decide how the worker shall manage his human capital. He cannot choose the level of the
worker’s effort (due to problems of observing effort), but he can choose the tasks on which the worker shall put his effort. Still the worker has the residual control right of the ideas he produces and thus have the chance to offer his value-added in an alternative market.

In the present paper, the worker is in some respects modelled as an independent supplier: the worker has residual control right of ex post values since he has the opportunity to sell his value added in the alternative market. But in some respects, he is modelled as a typical employee: he is a risk averse agent facing a risk neutral principal. He is giving the principal the authority to decide on his behaviour, that is, he is only allowed to exert effort along one dimension; hence he cannot take alternative actions that exclusively improve his bargaining position.

The large literature discussing the role of human capital in the modern corporation tends to focus on the problem of expropriation.¹ When knowledge is the critical resource of the firm, it may be easy for the employees to steal ideas and start their own business. The firm then has to find ways to avoid this expropriation. Rebitzer and Taylor (1997) argue that it may be necessary to reward those employees with the highest threat of expropriation with higher rents. Rajan and Zingales (2001) show how the problem of expropriation may determine the organizational structure. They argue that human capital intensive industries will develop flat organizations with distinctive technologies and cultures in order to avoid expropriation. The human capital focus in this paper is different. Instead of focusing on the firm’s ‘battle’ against expropriation or opportunism, I focus on how the risk averse employee’s possession of human capital constrains the feasible intensity of incentives in the employment contract.

The results of the analysis can be summarized as follows: First, the model shows how outside options constrain the feasible levels of performance pay. If the value of the worker’s outside alternatives are low, it may impossible to implement high-powered incentives, since high bonuses may lead the employer to renegotiate the terms of the contract ex post value realizations. But the existence of risk aversion captures a maybe more interesting result, not discussed in the literature: If the value of the worker’s outside alternatives are high, it may be impossible to implement contracts with low-powered incentives, since the worker, if he has

done a good job, has an incentive to renege on the contract and plea for a renegotiation. Hence, even though the worker prefers a wage contract with a higher fixed salary, the existence of good outside options creates a lower bound on the bonus level that lies above the desirable level. This reduces the feasible fixed salary that the employer can afford to pay.

Second, comparative static shows that the optimal bonus of the relational contract is a negative function of risk aversion and a positive function of incentive responsiveness. Hence, the repeated game approach is robust to the standard results from linear static incentive contracts. But, in contrast to static contracts, the optimal bonus of the relational contract is affected by the value of future surplus. Second-order effects show that the optimal bonus level’s sensitivity to risk aversion and incentive responsiveness increases with the discount factor.

Third, by elucidating the dual strategic property of outside options, the model makes it possible to systematically study the costs and benefits of relationship specificity. In particular, the model shows that relationship specificity can lead to more efficient incentive schemes.

Finally, the analysis shows that assumptions concerning ex post bargaining positions is crucial to statements on optimal firm boundaries. Baker, Gibbons and Murphy (2001, 2002) provide an answer to the famous ‘Williamson puzzle’ (1985), by showing that incentives from the spot market cannot always be replicated in a relational contract inside the firm, due to problems of contract enforcement. The model in this paper shows that this argument depends on the assumption that the worker has no control rights ex post value realizations. If the worker’s human capital is essential for ex post realizations, the firm can always replicate the market, but the market cannot always replicate the firm.

In the next section I will present the model. Comparative analysis is made in Section 3, while Section 4 discusses the model’s implications on firm boundaries. Section 5 concludes.
2. The model

Consider an employer and a worker, who together form what we can call a firm. The worker makes an unobservable choice of effort $e$, which stochastically determines the worker’s output. A random variable $x$ with mean zero and variance $V$ represents noise between the level of effort $e$ and the observed output $Y(e,x) = e + x$. I assume that $x$ has bounded support: $x \in (x_L, x_H)$.

The worker’s wage is linear in $Y$ and given by

$$w = \alpha + \beta Y(e, x),$$

where $\alpha$ is a fixed salary which is paid ex ante the production of $Y$, and $\beta Y(e, x)$ is paid ex post the production of $Y$. Holmström and Milgrom (1987) showed that normally distributed noise terms are necessary for linear incentive contracts to be optimal. Here, the noise term $x \in (x_L, x_H)$ does not fulfil this requirement. But even so, the choice of linear contracts can still be justified both on theoretical and empirical grounds. First, non-linear incentive contracts have the disadvantage of being susceptible to gaming. As Gibbons (2002) argues, the main contribution of the Holmström-Milgrom model is not that it justifies linear contracts, but rather that it implicitly demonstrates the gaming-problem of non-linear contracts. For example, Mirrlees’ (1974) famous step contract, where the agent earns $w_H$ if $Y \geq \bar{Y}$, but $w_L$ if $Y < \bar{Y}$, would induce no effort once the worker’s aggregate output passes $\bar{Y}$. Linear incentive contracts have the advantage of preventing these kinds of dynamic moral hazard problems within a period. A growing body of evidence is consistent with the prediction that non-linear contracts create history-dependent incentives, see for instance Healy (1985) on bonus plans with ceilings and floors, and Asch (1990) and Oyer (1998) on bonuses tied to quotas.

Moreover, the simplicity of linear contracts makes it reasonable to believe that costs associated with the implementation of such contracts are lower than the costs associated with more complex non-linear contracts. The gaming-problem can also contribute to excessive costs due to the implementation of non-linear contracts. The popularity of linear contracts
makes it reasonable to believe that excessive costs associated with non-linear contracts exist, especially since it is hard to find empirical evidence for one of the optimum conditions of linear contracts; normally distributed noise. Hence, for the rest of this paper I will assume that excessive costs associated with the implementation of non-linear incentive contracts exceed the benefits. This assumption is particularly reasonable in risk averse environments as considered in this paper. Since risk aversion and variance increases the complexity of non-linear contracts, and also make the gaming problem more severe, the costs associated with implementing non-linear incentive contracts are most likely a positive function of these variables.

Assume that the worker’s utility from wage is given by $u(w)$, where $u$ is three times differentiable, and the expected wage is equal to its mean, that is $\bar{w} = E[w]$. The worker’s certainty equivalent is then assumed to be

$$CE_w = \alpha + \beta e - C(e) - \frac{1}{2} r\beta^2 V,$$

where $r = r(\bar{w}) = -u''(\bar{w})/u'(\bar{w})$ is the worker’s coefficient of absolute risk aversion, $V = Var(w)$, and $C(e)$ is the personal cost of making effort, where $C'(e) > 0$ and $C''(e) > 0$ (the formulation of the certainty equivalent is a Taylor approximation).

The employer’s certainty equivalent can now be written

$$CE_e = e - (\alpha + \beta e),$$

and total certainty equivalent (TCE) is then $CE_w + CE_e$, that is

$$TCE = e - C(e) - \frac{1}{2} r\beta^2 V.$$
2.1 Verifiable contract

If the parties could write a verifiable contract on output level and the ownership of the output, they could easily implement the optimal division of incentives and insurance. The worker maximizes his certainty equivalent. The first order condition yields the following incentive constraint:

\[ \dot{\beta} = \frac{\partial C}{\partial e} \]

The employer now maximizes the total certainty equivalent by choice of \( \beta \), subject to the incentive constraint. That is

\[ \max_{\beta} \ (e - C(e) - \frac{1}{2} r\beta^2 V) \]

subject to (1)

Solving this for \( \beta \) yields

\[ \hat{\beta} = \frac{1}{1 + rV C''}, \]

where \( \frac{1}{C''} \) can be interpreted as the worker’s responsiveness to incentives( \( \frac{de}{d\beta} = \frac{1}{C'(e)} \)). From (2) we obtain the classical result that the optimal level of performance pay is a negative function of risk aversion and variance and a positive function of incentive responsiveness.

2.2 Relational Contract

Assume now that the worker’s output is not verifiable, and thus not enforceable by a court of law. Further on, the parties cannot write verifiable contracts ex ante on ownership rights ex post. The parties then have to agree on a self-enforcing relational contract. The worker’s choice of effort is equivalent to an investment in human capital that is essential in the ex post realization of output \( Y \), and there exist no verifiable contract that can force the worker to
realize internal trade. Hence, the worker can threaten ex post to trade the output with external trading partners. Assume that the alternative market values the effort of the worker to be $\theta Y(e, x)$ where $\theta \in (0,1)$.

The game between the worker and the employer now proceeds as follows: first the employer offers a compensation package $(\alpha, \beta)$, where $\alpha$ is a fixed salary to be paid ex ante the production of $Y$, and $\beta Y$ is the bonus meant to be paid ex post the realization of $Y$. Second, the worker makes a choice of effort $e$. Third, the employer and the worker observe $Y$. They now decide if they still want to accept the bonus element $(\beta)$ of the compensation package, or if they want to renegotiate the compensation scheme.

Assume that 50:50 Nash bargaining decides the price of the good if one of the parties chooses to renegotiate the contract.² The price is then $\frac{Y+\theta Y}{2}$, leaving a bonus element equivalent to $\frac{1+\theta}{2} = \gamma$. In a single-period relationship, the worker will choose to renegotiate if $\beta < \gamma$, and the employer will choose to renegotiate if $\beta > \gamma$, so the players will ex ante agree to a 50:50 Nash compensation $\gamma Y$. In other words: a relational contract where $\beta \neq \gamma$ is not enforceable. To be able to implement a relational contract, the players must have an infinite horizon (or an uncertainty with respect to when the relationship ends). To formalize this, I consider an infinitely repeated relationship between the worker and the employer, where they both play trigger strategies. The employer begins by offering a compensation package $(\alpha, \beta)$. The employer will continue to do so unless the worker or the employer chooses to renegotiate ex post, in which case they refuse to agree on anything else than the 50:50 Nash compensation $\gamma Y$, hereafter called a spot contract, forever after.³ (Note that even if we now enter into the

² The 50:50 Nash bargaining solution is quite common in the literature (see e.g. Grossman and Hart 1986; Baker, Gibbons and Murphy 2002). Most bargaining solutions are ex post Pareto-optimal as long as bargaining is costless and information is symmetric (see e.g. Rubenstein 1982). Anyway, the qualitative results in this paper will not change if we allow for another division of the surplus.

³ This trigger strategy has the advantage of being simple to analyse, but it also has the disadvantage of not regarding the issues of optimal punishment and renegotiation. Abreu (1988) shows that the highest equilibrium pay offs require the strongest credible punishment. In the model in section 2 the punishment of deviation is not the strongest, but the results of the model would hold even with optimal punishment, since the simple idea is that cooperation depends on the present value of the relationship. See Baker, Gibbons and Murphy (1994) for a similar argument.

The problem of renegotiation is that renegotiation from punishment is Pareto-efficient. One can meet this problem by arguing that a new relational contract, after deviation and renegotiation, could not be established on the same self-enforcing terms, since the threat of infinite punishment would not seem credible. See Fudenberg and Tirole (1991) for a discussion on renegotiation proofness.
study of repeated relationships, the moral hazard problem cannot be solved as in Radner, 1981, Rogersen, 1985, and Fudenberg, Holmström and Milgrom, 1990, since, in contrast to these models, the parties cannot write verifiable contracts.)

Given the employer’s strategy, if the worker accepts the bonus element of the contract, the present value of his expected profit is given by

\[ \beta (e^R + x) + \frac{\delta}{1-\delta} CE_w^R, \]

where superscript, \( R \), denotes relational contract, \( \delta \) denotes the discount factor and \( e^R \) maximizes the certain equivalent such that \( CE_w^R = Max(\alpha + \beta e - C(e) - \frac{1}{2} r\beta^2 V) \). If the worker reneges on the contract, and calls for a renegotiation, the present value of his expected profit is given by

\[ \gamma (e^R + x) + \frac{\delta}{1-\delta} CE_w^S, \]

where superscript, \( S \), denotes spot contract, and \( e^S \) maximizes the worker’s surplus from spot transactions, such that \( CE_w^S = \max_e (\gamma e - C(e) - \frac{1}{2} r\gamma^2 V) \).

The worker will stick to the original compensation package if

\[ \beta (e^R + x) + \frac{\delta}{1-\delta} CE_w^R \geq \gamma (e^R + x) + \frac{\delta}{1-\delta} CE_w^S \quad \forall x \].

Given the worker’s strategy, if the employer sticks to the original compensation package, the present value of his expected profit is given by

\[ (1-\beta)(e^R + x) + \frac{\delta}{1-\delta} CE_e^R, \]

where \( CE_e^R = e^R - (\alpha + \beta e^R) \). If the employer reneges on the contract, and calls for a renegotiation, the present value of his expected profit is given by
(7) \((1 - \gamma)(e^R + x) + \frac{\delta}{1 - \delta} CE^S_e\)

where \(CE^S_e = e^S - \gamma e^S\).

The employer will stick to original compensation package if

(8) \((1 - \beta)(e^R + x) + \frac{\delta}{1 - \delta} CE^R_e \geq (1 - \gamma)(e^R + x) + \frac{\delta}{1 - \delta} CE^S_e \quad \forall x\).

Combining (5) and (8) yields a necessary and sufficient condition for the relational contract to be self-enforcing:

\(|\gamma - \beta| \Delta x \leq \frac{\delta}{1 - \delta} (e^R - C(e^R) - \frac{1}{2} \gamma^2 V) - (e^S - C(e^S) - \frac{1}{2} \gamma^2 V),

where \(\Delta x = x_H - x_L\).

That is

(9) \(|\gamma - \beta| \Delta x \leq \frac{\delta}{1 - \delta} (TCE^R - TCE^S),

The parties can choose the fixed salary, \(\alpha\), to make the condition sufficient.

3. Comparative Analysis

From (9) we observe that there are upper and lower bounds on the feasible level of performance pay. Define \(\beta^R \in (\beta_L, \beta_H)\) as the feasible levels of performance pay in a relational linear incentive contract.

**Proposition 1:** The feasible levels of performance pay \(\beta \in (\beta_L, \beta_H)\) in a relational linear incentive contract are given by (9).
The proposition clarifies the limits of relational contracting. In a verifiable contract, any level of $\beta \in (0,1)$ is feasible, and the optimal choice is independent on outside options and discount factors. In a relational contract relying on self-enforceability, however, ex post outside options and the value from future trade, constrains the feasible $\beta$. In spirit, the proposition is similar to Levin (2003). He shows that if the agent is risk neutral, the optimal relational incentive contract is non-linear, where a bonus is paid if output exceeds a critical level. Due to risk neutrality, the strongest possible incentives are desirable, but self-enforcement imposes a lower and an upper bound on the critical output level. I show that if the agent is risk averse, and the parties stick to linear contracts, the feasible levels of performance pay have a lower and an upper bound $\beta \in (\beta_L, \beta_H)$. From the concavity of TCE, we have

**Lemma:** The optimal bonus level of a relational linear incentive contract is given by $\hat{\beta}$ iff $\hat{\beta} \in (\beta_L, \beta_H)$, $\beta_L$ iff $\hat{\beta} < \beta_L \leq \gamma$ and $\beta_H$ iff $\hat{\beta} > \beta_H \geq \gamma$, where $\beta \big| \beta_L \geq \hat{\beta}$ is given by

$$v(\gamma - \beta_L) = TCE^R - TCE^S, \quad v = \Delta x^{\frac{1-\delta}{\delta}}.$$

and $\beta \big| \beta_H \leq \hat{\beta}$ is given by

$$v(\beta_H - \gamma) = TCE^R - TCE^S.$$

Hence,

**Corollary:** There exist levels of $\gamma, \Delta x, \delta, r, V$ and $C''$ where the optimal level of performance pay in a verifiable linear incentive contract cannot be implemented in a relational linear incentive contract, that is $\hat{\beta} \not\in (\beta_L, \beta_H)$. 

12
It is naturally most interesting to study the properties of relational incentive contracts when \( \hat{\beta} \in (\beta_L, \beta_H) \).\(^4\) When \( \hat{\beta} > \beta_H > \gamma \) the employer has short-term gains from contract deviation. In order to commit to the contract, the employer cannot provide the worker with sufficiently high-powered incentives. This point is made in Baker, Gibbons and Murphy (2002): High-powered incentives cannot be implemented if the reneging temptations are too large.

But the problem can also be that the employer cannot provide the worker with incentives that are too low-powered. When \( \hat{\beta} < \beta_L < \gamma \), it is the worker who has the short-term gains from contract deviation. In order to deter deviation, the employer must offer \( \beta = \beta_L \) ex ante to meet the worker’s ex post outside opportunities. In order to earn profit he then has to reduce the fixed salary \( \alpha \). Hence, if the worker is risk averse, and \( \hat{\beta} < \beta_L \), then good ex post outside options is a ‘burden’ for the worker: even though the worker prefers a wage contract with a higher fixed salary, the ex post realization of value added automatically creates a lower bound on the bonus level, which again reduce the feasible fixed salary that the employer can afford to pay. In such, the model explains the existence of excessive bonuses in human capital-intensive industries where the workers are highly exposed to the incentives of the market (see e.g. Blair and Roe, 1999). Moreover the model cast light on the modern stress phenomenon in human capital intensive industries where employees experience so-called ‘burnout’ after working ‘24 hours a day’ (see e.g. ZDnet.com or MetaGroup.com for reports on this phenomenon).

By modelling the relational contract as a linear incentive contract, we are able to make comparative static on the effect of risk aversion, variance and incentive responsiveness on the optimal bonus level when \( \hat{\beta} \notin (\beta_L, \beta_H) \). Let \( k \) be a parameter in the cost function, and a measure of incentive responsiveness in the following sense: For \( e(\hat{\beta}, k) \) given by 

\[
\beta = \frac{ae}{e} (e, k), \quad \text{we have} \quad \frac{\partial e}{\partial \hat{\beta}} > 0.
\]

That is, the incentive responsiveness \( \frac{\partial e}{\partial \hat{\beta}} \) increases with increasing \( k \) (see appendix for more details). We obtain

\(^4\) It can be objected that the linear contract approximation is unrealistic when the parties cannot even implement the optimal slope of the linear contract. But the corollary above applies especially for higher levels of \( r \) and/or \( V \), and as previously argued, it is reasonable to believe that costs associated with implementing non-linear incentive contracts is a positive function of risk aversion and variance.
**Proposition 2:** The optimal bonus level of a relational linear incentive contract is a negative function of risk aversion and variance, and a positive function of incentive responsiveness. That is \( \frac{\partial}{\partial v} = \frac{\partial}{\partial v} < 0, \frac{\partial}{\partial k} > 0 \) and \( \frac{\partial}{\partial v} = \frac{\partial}{\partial v} < 0, \frac{\partial}{\partial k} > 0, i = H, L. \)

**Proof:** See appendix

This is not a surprising result as it replicates the standard result from verifiable linear incentive contracts. Nevertheless, it demonstrates the robustness of the infinite repeated game approach.

In relational contracts, as opposed to verifiable contracts, the optimal bonus’ sensitivity to changes in risk aversion, variance and incentive responsiveness is affected by the discount factor. On low discount factors, the relational contract is weaker, and the range of feasible bonus levels is smaller (\( \beta_H - \beta_L \) is smaller). This implies that the optimal bonus level is less sensitive to parameter-changes when \( \beta \in (\beta_L, \beta_H) \), and contrary:

**Proposition 3:** When \( \beta \notin (\beta_L, \beta_H) \), the higher the discount factor \( \delta \), the stronger is the effect of risk aversion, variance and incentive responsiveness on the optimal bonus level of the relational contract. That is \( \frac{\partial^2 \beta}{\partial \delta^2} > 0 \) and \( \frac{\partial^2 \beta}{\partial \delta \partial v} > 0, i = H, L. \)

**Proof:** See appendix

Proposition 3 implies that the ‘burden of outside options’ is hardest in low trust environments (see Hart, 2001, on interpreting the discount factor as a proxy for trust). If \( \beta < \beta_L \leq \gamma \) and the parties heavily discount the relationship’s future surplus, high levels of risk aversion or low levels of incentive responsiveness cannot ‘free’ the worker from high levels of performance pay.
3.1 Relationship specificity

When the optimal bonus of the verifiable contract, $\hat{\beta}$, cannot be implemented, the parties have incentives to adjust the specificity of the relationship in order to implement more efficient incentive schemes. That is, the parties have incentives to take investments that adjust $\gamma$. They can reduce $\gamma$ by relationship specific investments, for instance in firm specific training programs. And they can increase $\gamma$ by standardizing output or generalizing the skill of the worker. Of course, the parties must balance the gains from adjusting $\gamma$ with its costs.

Figure 1 shows (9) for $\hat{\beta} < \beta_L < \gamma$. The curved line shows $TCE^R$. The horizontal line shows $TCE^S$. These lines intercept where $TCE^R = TCE^S$ and $\beta = \gamma$. The chord shows the left hand side of (9) multiplied with $\frac{1-\delta}{\delta}$, where $\frac{1-\delta}{\delta} \Delta x = v$ decide its gradient. The feasible $\beta$ is in the region where the curved line lies above the chord, that is between $\beta_L$ and $\gamma = \beta_H$ on the horizontal axis, where $\beta_L$ is decided by the parameters.
From figure 1 we see if $\hat{\beta} < \beta_L < \gamma$, a marginal increase in $\gamma$ would reduce $\beta_L$ (since the gradient of the chord is unaffected) and thus increase social surplus. An increase in $\gamma$ increases the worker’s short-term gain from deviating (given positive realizations of output $Y$), but it also makes the future spot contract less attractive. On high discount factors, this strengthens the relational contract and makes it possible to negotiate a better-termed incentive scheme. If the discount factor is sufficiently low (the chord sufficiently steep), however, then the only feasible incentive scheme has bonus equal to $\gamma$, and the parties can only increase social surplus by lowering $\gamma$. Hence, in high trust environments, the parties would increase outside options, i.e. reduce the level of relationship specificity in order to implement more efficient incentive schemes, while in low trust environments, the parties must reduce outside options, i.e. increase the specificity level in order to increase social surplus. This relationship prevails when it is the worker who has short-term incentives to deviate from the contract; that is when $\hat{\beta} < \beta_L < \gamma$. When $\gamma < \hat{\beta}$, the story is reverse:

![Figure 2](image-url)
Figure 2 shows (9) when $\gamma < \beta_H < \hat{\beta}$. Now we see that a marginal decrease in $\gamma$ would increase $\beta_H$, and thus increase social surplus. Here, an increase in $\gamma$ would make the spot contract more attractive, and hence decrease the flexibility of the relational contract. If the discount factor is sufficiently low, however, the parties can only increase social surplus by increasing $\gamma$. Hence, in high trust environments, the parties would reduce the worker’s outside options, i.e. increase the level of relationship specificity in order to increase social surplus, while in low trust environments, the parties must increase outside options, i.e. reduce the specificity level in order to increase social surplus.

Figure 1 and 2 show the costs and benefits of relationship specificity. It can lead to opportunism, which is emphasized by transaction cost economists (see e.g. Klein, Crawford, Alchian, 1978), but relationship specificity can also be a commitment device, and lead to more efficient incentive schemes. Moreover, the analysis complements parts of Milgrom and Holmström (1991). They argue that the principal must restrict outside activities in order to implement efficient incentive schemes, especially when performance in the tasks that benefits the firm are hard to measure and reward. I show that not only the principal, but also the risk averse agent with essential human capital may have incentives to reduce outside options if it enables the principal to commit to a higher fixed salary.

Formally, figure 1 and 2 show:

**Proposition 4:** If $\hat{\beta} < \beta_L < \gamma$, then there exist a discount factor $\delta > \bar{\delta}$ so that $\frac{\partial C}{\partial q} > 0$ and $\delta < \bar{\delta}$ so that $\frac{\partial C}{\partial q} < 0$. If $\gamma < \beta_H < \hat{\beta}$, then there exist a discount factor $\delta > \bar{\delta}$ so that $\frac{\partial C}{\partial q} < 0$, and $\delta < \bar{\delta}$ so that $\frac{\partial C}{\partial q} > 0$.

The proposition exposes an interesting relationship between trust-level, reneging temptations and social surplus. In high trust environments, social surplus is increased by increasing short-term gain from reneging on the contract. Intuitively, this insight applies more generally. Increased outside temptations may strengthen an already ‘solid’ relationship.
4. The boundaries of the firm

As I indicate, the simple economic environment outlined in this paper may cast light on the puzzle of firm boundaries. When I introduce Section 2 saying that the employer and the worker “form what we can call a firm”, I anticipate what is usually called an employment relationship. This may seem inaccurate since the worker has the residual control right. Following Baker, Gibbons and Murphy (2002), the relationship considered in this paper should be described as relational outsourcing: the parties engage in a relational contract, not a spot contract, and the worker (the upstream party) has the residual control right of the asset. When I still choose to characterize the relationship as an employment relationship, and thus a “firm”, it comes from the assumption that the worker cannot take actions that exclusively change the value of the outside option. Hence, in this setting, as long as the parties engage in a relational contract, I will interpret the relationship as a firm. Once the parties decide to deviate from the contract and instead engage in a spot contract, however, the relationship can be considered as a market transaction. The question is then: when will the parties form a firm?

Assume that there is a direct cost of relational contracting. This can be, for instance, the cost of finding the right balance between bonus level and fixed salary. The model predicts that the parties will form a firm when the gains of writing contracts exceed the cost. What, then, are the gains of relational contracting? Well, it enables the parties to implement more efficient incentive schemes than the spot market agreement. But as we have seen in section 2, these gains vary. If the optimal bonus level is equal to the incentives of the spot market; that is \( \hat{\beta} = \gamma = \beta^{\delta} \), then there is no need for a relational contract to implement it. Hence, the parties will not form a firm. If \( \hat{\beta} \neq \beta^{\delta} \), then there exists a gain from engaging in a relational contract. If \( \hat{\beta} \) is close to \( \beta^{\delta} \), then the gains from relational contracting may be rather small. Also, if \( \hat{\beta} < \beta_{L} \) and the efficiency loss from not being able to implement \( \hat{\beta} \), is great, the gains from relational contracting will be small. We can formulate the following proposition:

---

5 This interpretation corresponds to Herbert Simon’s (1951) conception of the employment relationship: the parties engage in an employment relationship if the worker accepts the employer to exercise authority over the worker; that is the employer is given the authority to select the worker’s behavior pattern.

6 It is common to contrast the costly verifiable contracts with the costless relational contracts, but there may be direct costs associated with any kinds of contracting, regardless of how the contract is enforced.
Proposition 5: Let $\Phi$ denote the direct cost of relational contracting. The gains from relational contracting are given by $TCE^R - TCE^S = \Omega$. The parties will form a firm if $\Omega > \Phi$.

Proposition 5 is indeed in the spirit of Coase (1937) and Williamson (1985) as it sees firm boundaries as a question of transactional and contractual costs and benefits. But as oppose to Williamson, I do not claim that the possibility of opportunism (deviation) is greater in the market than in the firm. In fact, the possible gains from opportunism may be greater in the relational contracting of the firm. Avoiding the possibility of hold up is inevitable, since the worker’s human capital is essential in the ex post realization of firm value. Also, I do not claim that the incentives are necessarily more powerful in the market. The point is that the incentives of the spot market are more or less costless to implement, but less flexible in range. This is an amendment to Baker, Gibbons and Murphy’s (2001, 2002) solution to the Williamson puzzle, asking why one cannot replicate the market inside a firm. BGM show that incentives from the spot market cannot always be replicated in a relational contract inside the firm, due to problems of contract enforcement. The model in this paper shows that this argument depends on the assumption that the worker has no control rights ex post value realizations. If the worker’s human capital is essential for ex post realizations, the firm can always replicate the market, but the market cannot always replicate the firm. Hence, in human capital-intensive industries, the question is not if it is possible to provide “spot market incentives” inside the firm, but how costly it is to develop optimal incentive contracts instead of relying on the costless high-powered incentives of the spot market.

5. Concluding remarks

The problem of human capital is usually considered to be a problem of expropriation. This is primarily a problem in a risk neutral environment. If the worker is risk neutral, or if he has the chance to share risk with other employees, he may be tempted to take a good idea with him and start his own business. But the worker is often risk averse, and he cannot easily share the risk of possessing his own human capital. Still, if the worker cannot write verifiable contracts with his employer, the threat of expropriation or incessant renegotiation is underlying the employment relationship. The goal with this paper has been to show how a problem of writing verifiable incentive contracts with risk averse possessors of human capital constrains the
feasible intensity of incentives, and moreover how this have implications for specificity choices and firm boundaries.

The choice of analysing the employment relationship in the framework of an infinitely repeated game deserves a comment: This approach rests on the assumption of the self-interested rational “economic man”. Empirical research suggests, however, that individuals often behave in a more reciprocal manner (see Fehr and Gachter, 2000, for an overview). Reciprocity may imply co-operation in the one shot trust game, or a smaller threshold for cooperation in the dynamic game, but it may also imply a more severe punishment than what can be expected from the rational agent. Introducing reciprocity in the environment outlined in the previous section could moderate the predictions. If the employer offered a compensation package with a high fixed salary and a smaller bonus, the reciprocal employee could choose to accept the compensation even if the ex post outside opportunities were huge. This behaviour could stem from the employee’s loyalty to the employer. Such kind of loyalty may explain the stable long-term employment relationships one has observed in many industries. Recent studies suggest, however, that loyalty is eroding, especially in the human capital-intensive industries (see O’Connor, 1993 and Capelli, 2000). The self-interested rational agent may therefore work as a useful abstraction in dealing with human capital in the modern corporation.

**APPENDIX**

1. **Deducing (9)**

Since \( x \) is continuous, (5) and (8) includes infinite number of restrictions. But using bounded support on \( x \), we can find the binding constraints, analysing (5) and (8) for extreme realizations of \( x \).

When \( \beta \leq \gamma \), (5) is weakest for \( x = x_H \) and (8) is weakest for \( x = x_L \). The binding constraints are thus
A necessary condition for the relational contract to hold is that the sum of (A.1) and (A.2) holds. This yields

\[(A.3) \quad (\gamma - \beta)(x_H - x_L) \leq \frac{\delta}{1-\delta} (TCE^R - TCE^S) .\]

When \( \beta \geq \gamma \), (5) is weakest for \( x = x_L \) and (8) is weakest for \( x = x_H \). The binding constraints are thus

\[(A.4) \quad \beta(e^R + x_L) + \frac{\delta}{1-\delta} CE^R + \gamma (e^R + x_L) + \frac{\delta}{1-\delta} CE^S \]
\[(A.5) \quad (1 - \beta)(e^R + x_H) + \frac{\delta}{1-\delta} CE^R \geq (1 - \gamma)(e^R + x_H) + \frac{\delta}{1-\delta} CE^S ,\]

and the sum of (A.4) and (A.5) yields

\[(A.6) \quad (\beta - \gamma)(x_H - x_L) \leq \frac{\delta}{1-\delta} (TCE^R - TCE^S) .\]

Since (A.3) is relevant for \( \beta \leq \gamma \) and (A.6) is relevant for \( \beta \geq \gamma \) we can write these to restrictions in one expression using absolutes:

\[(9) \quad |\gamma - \beta| \Delta x \leq \frac{\delta}{1-\delta} (TCE^R - TCE^S) \]

where \( \Delta x = x_H - x_L \). As noted, the parties can choose the fixed salary, \( \alpha \), to make the condition sufficient.
2. The measure of incentive responsiveness

For \( e(\beta, k) \) given by \( \beta = \frac{\partial C}{\partial e}(e, k) \), we have \( \frac{\partial^2 C}{\partial \beta \partial e} > 0 \). That is, the incentive responsiveness \( \frac{\partial^2 C}{\partial \beta \partial e} \) increases with increasing \( k \). This holds if the cost function satisfies \( \frac{\partial^2 C}{\partial \beta \partial e} - \frac{\partial^2 C}{\partial \beta \partial k} > 0 \). (Example, the condition holds for a cost function of the form \( C(e, k) = A(k)e^n, n \geq 2 \), where \( A'(k) < 0 \). With this condition, the gain from a marginal increase in \( \beta \) increases with the level of incentive responsiveness. That is \( \frac{\partial C}{\partial \beta} = (1 - \beta) \frac{\partial C}{\partial \beta} > 0 \) for \( \beta < 1 \).

3. Proof proposition 2 and 3

When \( \beta \in (\beta_L, \beta_H) \) the optimal \( \beta \) is given by (2) showing that the optimal level of performance pay is a negative function of risk aversion and variance and a positive function of incentive responsiveness. From Lemma we have that \( \beta_L \) is optimal iff \( \beta < \beta_L \leq \gamma \) where \( \beta_L \) is given by

\[
(10) \quad v(\gamma - \beta_L) = TCE^R - TCE^S, \quad v = \Delta x^{\frac{1-\delta}{\delta}}
\]

When \( \beta_L < \beta_L < \gamma \) we must have (for simplicity I exclude functional arguments):

\[
(A.7) \quad \frac{-\partial C_R}{\partial \beta} \bigg|_{\beta=\beta_L} < v
\]

This is visualized in figure (1). The chord is steeper than the \( TCE^R \) curve at point \( \beta_L \).

Differentiating (10) with respect to \( r \) yields
From (A.7), the bracket on the left hand side is negative, and the difference on the right hand side is positive since $\beta_L < \gamma$ and $\frac{\partial TCE}{\partial b_T} < 0$. This yields $\frac{\partial b_L}{\partial r} < 0$, which also implies $\frac{\partial b_L}{\partial V} < 0$. Differentiating (10) with respect to $k$ yields

$$\frac{\partial TCE}{\partial k} \frac{\partial b_L}{\partial k} = \frac{\partial TCE}{\partial k} - \frac{\partial TCE}{\partial k}$$

From (A.7), the bracket on the left hand side is negative, and the difference on the right hand side is also negative since $\beta_L < \gamma$ and $\frac{\partial TCE}{\partial b_T} > 0$. This yields $\frac{\partial b_L}{\partial k} > 0$.

From lemma we have that $\beta_H$ is optimal iff $\dot{\beta} > \beta_H \geq \gamma$, where $\beta_H$ is given by

$$(11) \quad v(\beta_H - \gamma) = TCE^R - TCE^S$$

When $\dot{\beta} > \beta_H \geq \gamma$ we must have

$$\frac{\partial TCE}{\partial b} \bigg|_{\beta = \beta_H} < v$$

This is visualized in figure (2). The chord is steeper than the $TCE^R$ curve at point $\beta_H$.

Differentiating (11) with respect to $r$ yields

$$\frac{\partial TCE}{\partial b} \frac{\partial b_L}{\partial r} = \frac{\partial TCE}{\partial r} - \frac{\partial TCE}{\partial r}$$

From (A.10), the bracket on the left hand side is positive, and the difference on the right hand side is negative since $\beta_H > \gamma$ and $\frac{\partial TCE}{\partial b_T} < 0$. This yields $\frac{\partial b_L}{\partial r} < 0$, which also implies $\frac{\partial b_L}{\partial V} < 0$. 23
Differentiating (11) with respect to \( k \) yields

(A.12) \((v - \frac{\partial TCE}{\partial \phi}) \frac{\partial \beta}{\partial k} = \frac{\partial TCE}{\partial k} - \frac{\partial TCE}{\partial \phi}\)

From (A.10), the bracket on the left hand side is positive, and the difference on the right hand side is also positive since \( \beta_H > \gamma \) and \( \frac{\partial TCE}{\partial \phi} > 0 \). This yields \( \frac{\partial \beta_H}{\partial k} > 0 \).

Proposition 3 can be verified by differentiating (A.8), (A.9), (A.11) and (A.12) with respect to \( v \) noting that \( \frac{\partial v}{\partial \phi} < 0 \).

REFERENCES


