Relational Incentive Contracts
and
Allocation of Ownership Rights

Four Essays

by

Ola Kvaløy

A dissertation submitted for the degree of

doctor oeconomiae

2003

Acknowledgements

The supervision committee of this dissertation has consisted of Professor Trond E. Olsen, Norwegian School of Economics and Business Administration, Professor Petter Osmundsen, Stavanger University College, and Professor Frøystein Gjesdal, Norwegian School of Economics and Business Administration.

In the role as my main supervisor, Professor Trond E. Olsen has guided me through the critical stages of the dissertation. His encouragements, comments and suggestions have been crucial to my doctoral work, and the dissertation would not have been where it stands without his generous help.

Professor Petter Osmundsen was the one who encouraged me to write this dissertation, of which I am grateful. Having an office next to him at Stavanger University College has led to many discussions of both 'related and unrelated matters'. Petter has given me a considerable degree of 'job discretion', but at the same time helped me to keep the research focused.

Professor Frøystein Gjesdal entered the committee at a later stage, but has influenced the dissertation by valuable comments and suggestions.

In addition, fellow student Frode Meland, Associate Professor Iver Bragelien, Professor Kjell Hausken, and my supervisor at 'hovedfag', Professor Gaute Torsvik, have given helpful comments.

During my three years of doctoral work, I have been happy to share office with fellow student Sigbjørn Tveterås. I am grateful for hours of sense and nonsense. I am also grateful to Professor Frank Asche and Professor Ragnar Tveterås, whom together with Sigbjørn and Petter have formed an inspiring and stimulating circle of economists here at Stavanger University College.

The three-year doctoral grant was provided by Norwegian School of Economics and Business Administration. I am grateful to the Department of Finance and Management Science for administrative help, and the Department of Petroleum Technology at Stavanger University College, for accommodating office and equipment.

I am also grateful to Barbara Beuche and Professor Walter W. Powell for their hospitality during my stay at Scancor, Stanford University, July-October 2002. The stay at Stanford was financially supported by L. Meltzers Høyskolefond , Ingjerd og Arne Skaugs forskningsfond, Grosserer E.A Smith og konsul L. Cappelen Smiths forskningsfond, Andreas Amlands handelshøyskolefond and Wilhelm Keilhaus minnefond.

I thank my mother and father for their encouragement and support through years of education. And in particular, I thank my dear wife, Marie, for her love and patience, and my daughter, Hanna, for reminding me about the enjoyment of leisure.

Stavanger, September 2003

Ola Kvaløy

Contents

Introduction

Consider a shipper who hires a trucker to ship goods from A to B. The shipper may provide the truck, or the trucker may come with his own. The quality of the trucker's effort and output is subject to influence. The trucker may choose how to maintain the truck (how to drive), how cautious he is with the goods he ships, how fast he ships the goods, and so on. In a full information economy, the question of who should own the truck is a question of technology. Differences in ability to exploit the values of the truck decide efficient allocation. For instance, if the shipper can exploit economies of scale in truck preservation and repairs (if he owns several trucks), or economies of scope in networking, logistic, customer relations etcetera, he should own the truck. Any concerns about the trucker's effort and output can be perfectly solved by contracts.

Full information is, however, not realistic. Effort is often hard to observe. Output may often be observable, but still difficult to verify by the court, since the assessment of an output's quality may be complicated and also rather subjective. So even if it is possible for the parties to contract on key output variables, they cannot solely rely on court enforcement.

So assume that it is difficult, or costly, for the parties to contract on maintenance. An efficient way for the shipper to induce the trucker to carefully maintain the truck could then be to let the trucker own it. As owner, the trucker becomes residual income claimant, which implies that he gets the money when he sells the truck. That way truck ownership gives the trucker an incentive to yield effort in truck maintenance. But by owning the truck, he can decide non-contractual usages, for instance to offer truck services to other shippers. Even if total surplus (shipper's + trucker's surplus) is maximized if the trucker concentrates on doing an optimal job for the shipper, he may instead be tempted to engage in rent seeking while he works for the shipper (e.g. spending time on searching for alternative customers), in order to maximize his own surplus at the expense of the shipper. Hence, there are costs and benefits with letting the trucker own the truck. Since it is impossible, or at least very expensive, to write legally enforceable contracts, the trucker and the shipper must to a certain extent trust each other in order to reduce the costs and exploit the benefits of the relationship. If they transact with each other only once, it may be difficult to establish trust, but if they transact with each other repeatedly, trust abuse can be costly, since it may ruin a profitable long-term economic relationship.[1]

The objective of this dissertation is to decide optimal incentives and efficient allocation of ownership rights in transactions that cannot be fully protected by legal contracts. In the shipper-trucker example the questions are: Generally, what are the feasible incentive schemes that the shipper can offer the trucker? And specifically, who should own the truck?

Even if a transaction cannot be fully protected by legal contracts, it does not mean that the transaction cannot take place. The question is how to efficiently implement and protect these kinds of transactions. I consider two solutions: efficient design of 'relational contracts' and efficient allocation of ownership rights.

Relational contracts: Transacting parties can write contracts even if the contract cannot be enforced by the legal system. The main attribute with contracts that are not legally enforceable is that it includes elements that cannot be verified by parties that are not involved in the transaction, e.g. the court. In the economics literature, a contract that cannot be legally enforced is most often referred to as a 'relational contract', but the term 'implicit contract' is also used. In this dissertation I will for the most part use 'relational contract'.

A relational contract is a response to the necessity of incomplete contracts. For a contract to be complete, each transacting party must be able to foresee, and accurately describe, all relevant contingencies that might happen during the course of the contract. Moreover, they must determine and agree upon an efficient course of action for each possible contingency. This is at best very costly, but in most cases unrealistic. Then, instead of trying to approach complete contracting, transacting parties often choose the relational contracting approach that

[1] The shipper/trucker example is inspired by Baker and Hubbard (2003).

focuses on the goals and objectives of the transaction without trying to specify all contingencies. In this dissertation, it is the problem of contracting on the quality of the agents' output that makes complete contracting impossible, and relational contracting necessary.

Since a relational contract cannot be legally enforced, it must be self-enforcing. That is, the parties that engage in the contract must have economic incentives to honour it. The more frequently the parties transact, in other words, the longer the term of the economic relationship, the higher are the incentives to honour the contract, since contract deviation can ruin future trade. A self-enforcing relational contract is thus modelled in a repeated game framework where the contracting parties write a contract on future transactions, and where the present value of honouring the contract versus the present value of reneging decides the contract's self-enforcing conditions. In such, the repeated game approach is a formalization of a rational trust-concept: an agent honours trust if it is profitable.

Ownership rights: There are two important qualities with the institution of ownership that makes the allocation of ownership rights crucial for efficient implementation of non-verifiable contracts: first, ownership, accompanied by secure property rights, gives residual control rights that, to a certain extent, can substitute for complete contracts. Residual control right over an asset is the right to make any decisions concerning the asset's use that are not explicitly controlled by law or assigned to another contract. Since the owner of an asset has this right to decide any non-contractual usages, the institution of ownership can substitute for legal contracts when disputes take place. Second, ownership rights act as an incentive-device when complete contracts are impossible. The owner of an asset does not only possess the residual control rights, but he is also the residual income claimant of that asset. The residual return from an asset is the amount that is left over after all contractual obligations have been paid. Residual income together with residual rights gives the owner of an asset incentive to manage the asset optimally.

In this dissertation I mainly consider transactions that are subject to asymmetric information. Asymmetric information implies that that there are relevant variables that are known to some, but not all the agents involved in the transaction. This implies either that transacting agents cannot observe each other's actions – inducing problems of moral hazard, or ability – inducing problems of adverse selection. In this dissertation, the informational focus is on non-observable effort decisions and thus the problem of moral hazard. With moral hazard follows the need for incentives. The relational contract must therefore be able to implement proper incentive schemes. The strength and structure of these incentives are decisive for the relational contract's self-enforcing conditions. The challenge is to combine efficient design of self-enforcing relational incentive contracts with efficient allocation of ownership rights in a way that maximizes the surplus of the transaction.

Literature

The first formal analysis of the ownership rights solution to incomplete contracts was made by Grossman and Hart (1986), and further developed by Hart and Moore (1990). Grossman, Hart and Moore (GHM) demonstrate how residual ownership rights can substitute for legal contracts and act as an incentive-device. In a formal analysis, GHM show how transacting parties can, prior to the transaction, allocate ownership rights in a way that maximizes the surplus of the transaction, in particular in a way that reduces the level of relationship specific under- or over-investment. Since GHM interpret ownership rights to physical assets as the defining characteristic of the firm, their theory was in fact a first formal analysis of firm boundaries that did not take the production function as a starting point.

GHM's important contribution, known as the property rights approach, was in many ways a formalization of the seminal work of Coase (1937), Williamson (1975,1979,1985) and Klein, Crawford and Alchian (1978). Coase asked the simple question: "why do we have firms?" and answered that firms will only exist in environments where firms perform better than markets. He argued that some transactions are so costly that they cannot be efficiently implemented in markets. Coase especially discussed the cost of discovering relevant prices, and the costs of "negotiating and concluding a separate contract for each exchange transaction..."and argued that firms (or ownership integration, to use the language of Grossman, Hart and Moore) greatly reduces the number of necessary contracts. Williamson developed the 'transaction cost theory' by identifying more thoroughly the conditions that create transaction costs. Instead of focusing on the direct costs of contracting, as Coase did, he discussed strategic transaction costs such as the problem of imperfect contracting in terms of opportunism and relationship specific under (or over) investments. The latter was particularly discussed by Klein et. al., who demonstrated how specific investments by A in B creates specialized quasi rents that can be appropriated by B after the investment is sunk, and how this cause A to under invest in B. They argued that this kind of transaction costs could be reduced by ownership integration.

A critique against transaction cost theory was that even if it demonstrated the costs of contracting in the market (between firms) it did not elucidate the benefits of organizing the transactions within the firm. This was the key objective to GHM; to formally compare the costs and benefits of ownership. They showed that ownership integration does not necessarily reduce the problem of relationship specific investments. The benefit of integration is that it increases the acquiring firm's incentive to make relationship specific investments since it will receive a greater portion of the ex post surplus from the transaction. But the cost of integration is that the acquired firm's incentive to make specific investments decreases since it will receive a smaller fraction of the surplus.

In order to insulate the effect of ownership, the property rights approach considered a static environment absented from contracts. An important point from Williamson (inspired by Simon, 1951; Macaulay, 1963 and MacNeil, 1974) was in fact that firms could use relational contracts in order to deal with the excessive costs of complete contracting. Williamson did not, however, define relational contract in a way that could be used in more formal analysis. The first repeated game analysis of relational contracts, where the present value of honouring the contract versus the present value of reneging decides the terms of the contract, was made by Klein and Leffler (1981). They analysed reputation effects in assuring product quality. The buyer pays a price premium to the supplier to assure that the supplier exert effort in producing good quality. If the supplier reneges on the contract, all his potential customers get to know this, and the supplier therefore loses all future sales. Further progress to the relational contracting modelling was made by Shapiro and Stiglitz (1984), Bull (1987), Klein (1990) and especially MacLeod and Malcomson (1989). Common for these models is, however, that the parties have symmetric information; hence the moral hazard/hidden information problem, which is considered to be the classic impediment to effective contracting (e.g. Holmström, 1979), does not appear in these models. Relational contract models with asymmetric information was developed by Baker, Gibbons and Murphy (1994, 2002) and especially Levin

(2003) who makes a general treatment of the self-enforcing relational contract model with asymmetric information and risk neutral agents.

The first paper to comprehensively combine the formalization of relational contracts, with the property rights approach, is Baker, Gibbons and Murphy's "Relational Contracts and the Theory of the Firm" (2002). Their main result is that asset ownership affects transacting parties temptations to renege on relational contracts. They deduce several implications from this result, where the most important are: i) high-powered incentives create bigger reneging temptations under integration than under non-integration. ii) it is impossible to mimic the spot-market inside the firm, because the reneging temptation then becomes too big. iii) vertical integration is an efficient response to widely varying supply prices because varying supply prices create bigger reneging temptations when the parties are non-integrated.

The present dissertation draws both on the property rights approach and the relational contracting literature, and thus especially on Baker, Gibbons and Murphy's important contribution. Methodologically, I apply non-cooperative game theory, and especially theories of repeated games, using Nash's (1950) equilibrium concept, refined by Selten (1965). In addition I use calculus to solve constrained optimisation problems.

The essays

The dissertation consists of four essays, or 'papers', that each deals with transactions that cannot be fully protected by the legal system. But even if the papers are founded on the same assumption of non-verifiable contracts, they each emphasize different topics within what broadly can be termed organizational economics, and can thus be read independently from each other.

Asset Specificity and Vertical Integration

This paper seeks to modify the conventional hypothesis in the literature that increased asset specificity leads to more vertical integration. Klein et. al. (1978) note that asset specificity creates appropriable specialized quasi rents and claim that "integration by common or joint ownership is more likely the higher the appropriable specialized quasi rents of the assets involved." Williamson (1985) argues that the high-powered incentives of markets may impede the efficient coordination required to implement relationship specific investments, since the involved parties want to appropriate as much as possible of the coordination gains. Vertical integration is a way of reducing these kinds of co-ordination problems. Grossman, Hart and Moore (1986,1990) show that if complementary or co-specialized assets operate under separate ownership, the parties owning the assets will under invest in the relationship.

I note that if the level of asset specificity increases, the relative value of external trade is reduced. In a modified version of Baker, Gibbons and Murphy's (BGM) model (2002), I show that in a relational contract between two independent parties of a relationship specific input, increased specificity reduces the temptation to renege on the relational contract, since the benefit of external trade is reduced. This creates scope for higher-powered incentives. If the buyer owns the seller (vertical integration), however, the buyer has the residual control right to the good produced, so the seller cannot hinder the buyer to force internal trade. Hence, in this case asset specificity does not affect the self-enforcing conditions of the contract. Non-integration can therefore be an efficient response to asset specificity.

The result finds empirical support, for instance in the offshore industry where the oil companies and their main suppliers, who design and build installations that the oil companies use to extract oil, always operate with separate ownership even though the suppliers manage capital stock and produce inputs that are highly specific to the buying oil companies.

The modification to BGM is mainly in terms of player strategies. In their set up, the parties play grim trigger strategies in which deviation from the relational contract results in spot governance forever after. In spot governance, the parties cannot contract ex ante on ex post realizations, but they can negotiate ex post over the price of the good. I analyse so-called carrot and stick strategies where contract deviation results in a one-period trade in the alternative market before return to the relational contract. These kinds of strategies, also called 'mutual punishment', are more complex to analyse, but in settings with high levels of asset specificity, they are still more realistic than the standard grim strategies.

Human Capital and Risk Aversion in Relational Incentive Contracts

This paper examines a relational incentive contract between a risk neutral principal (or employer) and a risk averse agent (or worker/employee) where the agent's human capital is essential in ex post realization of values. I analyse the effect of outside options on the optimal degree of performance pay, showing how the presence of ex post outside options may impede desirable degrees of performance pay. In particular, I show that if the value of the agent's outside alternatives are high, it may be impossible to implement contracts with low-powered incentives, since the agent, if he has done a good job, has an incentive to renege on the contract and plea for a renegotiation. Hence, even though the agent 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. I show how the parties can eliminate this problem by choosing a level of asset specificity that enables the parties to implement the desirable degree of performance pay.

The paper is, to my knowledge, the first to analyse relational contracts that includes both asymmetric information, in the form of unobservable effort, and risk aversion. It is difficult 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, variance and incentive responsiveness within relational contracts with asymmetric information. 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 paper 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. In order to capture this balance, I analytically separate 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: the principal has the authority to decide on the agent's behaviour, that is, the agent is only allowed to exert effort along one dimension; hence he cannot take alternative actions that exclusively improve his bargaining position. But the agent has residual control rights of ex post values since he has the opportunity to sell his value added in the alternative market. The analysis shows that this right is crucial to statements on optimal firm boundaries. BGM (2001, 2002)) provide an answer the famous 'Williamson puzzle' 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 agent has no control rights ex post value realizations. If the agent is risk averse, and his human capital is essential for ex post realization of values, the firm can always replicate the market, but the market cannot always replicate the firm.

Team Incentives in Relational Contracts

This paper compares three types of incentive schemes for teams: an incentive scheme based on joint performance evaluation (JPE) compensates members of team based on the team's overall performs. Hence, a worker is rewarded if his peers perform well. An incentive scheme based on relative performance evaluation (RPE) rewards team members that perform relatively better than their peers. Hence, a worker is rewarded if his peers perform badly. An incentive scheme based on independent performance evaluation (IPE) rewards the team members independently from their peers' performances.

The classic advantage with RPE is that the employer can make a workers' compensation independent from good or bad outside factors (common noise components), while the advantage with JPE is that it encourages cooperation. Che and Yoo (2001) formalize the latter by considering an implicit contract between two agents on yielding high effort. I combine the framework of Che and Yoo on team incentives with the framework of BGM (2002) on relational contracts to find under which conditions relational incentive contracts based on JPE, RPE and IPE are self-enforcing. Like Che and Yoo, I study a repeated game between one principal and two agents, but contrary to them I assume that the quality of the agents' output is non-verifiable, so that legal enforcement is impossible. Hence, I model a self-enforcing relational contract between the principal and his agents.

I deduce two results that can explain why individual compensation based on IPE and/or RPE is especially common among 'white collar' workers in the higher levels of organizations: first, I find that we can expect a relatively higher frequency of incentive schemes based on RPE and IPE when the productivity of effort is high. The higher the productivity, the easier it is for the principal to offer credible incentive schemes, hence the critical discount factors necessary to implement relational contracts, decrease. But, on low discount factors, JPE is more expensive since it is harder for the agents to enforce an implicit contract between them on low discount factors. Hence, higher productivity makes the enforceability of IPE and RPE rise relatively to JPE. Second, I find that IPE is more probable if the agents own the critical assets. By each owning a critical asset, the agents can ex post bargain over the surplus independently from each other. This opportunity is especially threatening to the relational contract when incentives are based on non-independent performance evaluation since the agents risk getting low wages from the relational contract, and high wages from independent bargaining. This problem is eliminated with IPE.

Norms Matter

The objective of the last paper is to demonstrate that norms, in terms of trust-level, trust structure and type of trust, is decisive for firm boundaries. Hart (2001) argues that it is hard to find clear-cut relationships between trust and optimal allocation of asset ownership. In the typical repeated game approach, where the discount factor is a proxy for trust, a move from a 'low-trust' environment to a 'high-trust' environment only reveals that cooperation in general is more likely to occur both between integrated parties and non-integrated parties.

I develop two simple models to show that norms actually can determine firm boundaries. First, I consider a repeated trust game (with symmetric information) between a buyer and a seller of inputs. I do not make the standard assumption that contract verifiability is exogenously given; instead I introduce endogenous probability of contract verification. The more the parties invest in contract specifications, the higher is the probability of a third party (the court) verifying and thus legally enforcing the contract. If the discount factor is sufficiently low, transactions will only be implemented if the possibility for contract verification is sufficiently high. By introducing this endogenous probability of legal contract enforcement, we get hold of a trade-off between 'direct transactions costs', discussed by Coase, and 'strategic transaction costs', introduced by Williamson et. al: the lower the degree of relationship-specific under investment, the higher is the necessary investment in contract specification.

This substitutability between direct and strategic transaction costs makes it possible to study optimal asset allocation focusing on institutional differences in possibilities of legal contract enforcement. Following, Rock and Wachter (2001), I argue that the incentives to invest in a detailed contract are weaker when parties are under the same ownership than when parties are under separate ownership. In integrated firms, the marginal effect of contract specifications on the probability of legal enforcement is reduced by the probability that the owner instead of the court will solve contractual disputes. In repeated relationships with contracts relying both on legal enforcement and self-enforcement, this implies that non-integration dominates in low trust environments: if the discount factor is sufficiently low, the contract cost necessary to implement a contract is higher when the parties are integrated, than when the parties are non-integrated.

The second model is a static version of the same game, but now with non-verifiable actions, non-rational, reciprocal agents and asymmetric information. I show that if the trust game is played only once, the party with the best reputation of being trustworthy should own the asset. As residual claimant, the asset owner has the power to either honour or abuse trust, when contracts cannot be verified. If the asset owner is considered trustworthy, it relaxes the trusting party's participation constraint, and increases the surplus from the transaction. Hence, if reciprocity is the norm, then the parties' reputation of obeying this norm is decisive for optimal asset allocation, since the asset owner's reputation decides the surplus from the transaction.

REFERENCES

Baker, George, Robert Gibbons and Kevin J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts." *Quarterly Journal of Economics* 109: 1125-56.

Baker, George, Robert Gibbons and Kevin J. Murphy. 2002. "Bringing the Market Inside the Firm." *American Economic Review Papers and Proceedings* 91: 212-18.

Baker, George, Robert Gibbons and Kevin J. Murphy. 2002. "Relational Contracts and the Theory of the Firm." *Quarterly Journal of Economics* 117: 39-94.

Baker, George and Thomas Hubbard. 2003: "Make versus Buy in Trucking: Asset Ownership,

Job Design and Information." *American Economic Review* 93: 551-72.

Bull, Clive. 1987. "The existence of Self-Enforcing Implicit Contracts." *Quarterly Journal of Economics* 102: 147-59.

Che, Yeon-Koo and Seung-Weon Yoo 2001. Optimal Incentives for Teams. *American Economic Review* 91: 525-40.

Coase, Ronald H. 1937. "The Nature of the Firm." *Economica* 4: 386-405.

Grossman, Sanford and Oliver Hart. 1986 "The Costs and Benefits of Ownership: A theory of Lateral and Vertical Integration." *Journal of Political Economy* 94: 691-719.

Hart, Oliver and John Moore. 1990. "Property Rights and the Nature of the Firm." *Journal of Political Economy* 98: 1119 – 58.

Hart, Oliver. 2001. "Norms and the Theory of the Firm." *University of Pennsylvania Law Review* 149: 1701-15.

Holmström, Bengt. 1979. "Moral Hazard and Observability." *Bell Journal of Economics,* 10: 74-91.

Klein, Benjamin, Crawford Robert G. and Alchian, Armen A. 1978. "Vertical Integration, Appropriable Rents, and the Competitive contracting Process." *Journal of Law and Economics* 21: 297-326.

Klein, Benjamin and Keith Leffler. 1981. " The Role of Market Forces in Assuring Contractual Performance." *Journal of Political Economy* 89:615-41.

Kreps, David M., 1990. "Corporate Culture and Economic Theory." In J. Alt and K. Shepsle, eds. *Perspectives on Positive Political Economy.* Cambridge University Press.

Levin, Jonathan. 2003. "Relational Incentive Contracts." *American Economic Review* 93: 835-57.

MacLeod, Bentley and James Malcomson. 1989. "Implicit contracts, Incentive Compatibility, and Involuntary Unemployment." *Econometrica* 57: 447-80

Macauley, Stewart. 1963. "Non-Contractual Relations in Business: A Preliminary Study," *American Sociological Review* 27:55-67.

Macneil, Ian. 1978. "Contracts: Adjustments of Long-Term Econmic Relations under

Classical, Neoclassical, and Relational Contract Law." *Nortwestern University Law Review* 252: 854-906

Nash, John. 1950. "Equilibrium Points in n-Person Games." *Proceedings of the National Academy of Sciences* 36:48-49.

Rock, E. and M. Wachter. 2001. "Islands of Conscious Power: Law, Norms and the Self Governing Corporation." *University of Pennsylvania Law Review* 149: 1619-77.

Selten, R. 1965. "Spieltheoretisce Behandlung eines Oligopol-modells mit Nachfragetrageheit" *Zeitschrift für Gesamte Staatswissenschaft,* 121:301-24

Shapiro, Carl and Joseph E. Stiglitz. 1984. "Equilibrium Unemployment as a Worker Discipline Device." *American Economic Review 74:* 433-44.

Simon, Herbert. 1951. "A formal theory of the employment relationship." *Econometrica* 19: 293:305.

Williamson, Oliver E. 1975. *Markets and Hierarchies: Analyses and Antitrust Implications.* New York: The Free Press.

Williamson, O. E. 1979."Transaction-Cost Economics: The Governance of Contractual Relations." *Journal of Law and Economics* 22: 233-261.

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism.* New York: The Free Press.

Essay

I

Asset Specificity and Vertical Integration

Ola Kvaløy*

Norwegian School of Economics and Business Administration

Abstract: *Asset specificity is usually considered to be an argument for vertical integration. The main idea is that specificity induces opportunistic behaviour, and that vertical integration reduces this problem of opportunism. In this article I show that asset specificity actually can be an argument for non-integration. In a repeated game model of relational contracts, based on Baker, Gibbons and Murphy, 2002, I show that asset specificity affects the temptation to renege on relational contracts between non-integrated parties, but not between integrated parties. If the parties are non-integrated, higher levels of specificity can provide relational contracts with higher-powered incentives.*

* I thank Iver Bragelien, Frøystein Gjesdal, Kjell Hausken, Frode Meland, Petter Osmundsen, Gaute Torsvik and especially Trond E. Olsen for valuable comments and suggestions.

1. Introduction

The transaction cost theory entered the stage in the mid 1970s, partly as an attempt to explain the fundamental Coaseian question: Why do we have firms? The question acted as a headline for the general problem of economic organization: How do we explain the various observed ways of organizing economic activity? The factors leading to vertical integration have been a central issue in this literature. And a factor that has received a lot of attention is the degree of asset specificity. The traditional hypothesis is that asset specificity leads to vertical integration. This hypothesis is formulated through different lines of thought. Klein, Crawford and Alchian (1978) emphasize the problem of "hold-up". A party that has invested in specific assets may be forced to accept a worsening of the terms of the relationship after the investment is sunk. Hence, asset specificity creates appropriable specialized quasi rents. Klein et al. claim that "integration by common or joint ownership is more likely the higher the appropriable specialized quasi rents of the assets involved." Williamson (1985, 1991) emphasizes the problem of maladaptation. As investments in specific assets increase, disturbances requiring coordinated responses become more numerous and consequential. The high-powered incentives of markets may impede efficient coordination, since both parties want to appropriate as much as possible of the coordination gains. Vertical integration is a way of reducing this kind of maladaptation. The "property rights approach", developed by Grossman, Hart and Moore (GHM) (1986, 1990), does not formulate an explicit hypothesis concerning asset specificity, but states that if assets are strictly complementary, then some form of integration is optimal. GHM show that if complementary or co-specialized assets operate under separate ownership, the parties owning the assets will underinvest in the relationship.

The three approaches introduced here share the common belief that there is a correlation between the degree of asset, or investment, specificity and the appearance of vertical integration. In the April 2000 edition of Journal of Law and Economics, Klein states that "*...the rigidity costs associated with long term contracts increase as relationship-specific investments increase (...). Therefore, the greater the relationship-specific investments present in an exchange, the more likely vertical integration (that avoids the rigidity costs associated with long term contracts) will be chosen as the self-enforcing arrangement. All that is* *required for this positive relationship between specific investments and the likelihood of vertical integration is that the relative inefficiency costs from weakening of incentives is not systematically positively related to the level of specific investments, and there is no reason to believe they are."*

In the present article I will show, however, that there may be a reason to believe that the "relative inefficiency costs from weakening of incentives" are systematically positively related to the level of asset and investment specificity. The analysis draws on a repeated game model of relational contracts developed by Baker, Gibbons and Murphy (BGM), 2002. A relational contract[1] contains rules or standards that cannot be legally enforced. Hence, the contract must be self-enforceable in the sense that the present value of honouring the contract must be greater than the present value of reneging. BGM show how asset allocation matters in the presence of long-term relational contracts. An important result is that incentives in relational contracts between firms can be higher-powered than incentives in relational contracts within firms. In a modified version of BGM's model, I show that this difference in incentive intensity is positively related to the degree of asset and investment specificity.

The repeated game model is one in which an upstream party in each period uses an asset to produce a good that could either be used in a specific downstream party's production process, or put to an alternative use. Asset ownership conveys ownership of the good produced, so if the upstream party owns the asset (non-integration),[2] the downstream party cannot use the good without buying it from the upstream party, whereas if the downstream party owns the asset (integration), then he already owns the good. Since the good's value to the downstream party exceeds its value in the alternative market, the parties agree on a relational contract where the downstream party pays bonuses to make the upstream party improve the specific quality of the good. In order to analyse asset specificity within this framework, it is necessary to make modifications to BGM's model. In their set up, the parties play grim trigger strategies in which deviation from the relational contract results in spot governance forever after. In spot governance, the parties cannot contract ex ante on ex post realizations, but they can negotiate ex post over the price of the good. BGM assume Nash bargaining, so the price depends on

[1] Relational contracts are also called 'implicit' contracts (e.g. MacLeod and Malcomson, 1989).
[2] Following Grossman and Hart's (1986) terminology, seller ownership is called "non-integration"; buyer ownership is called "integration".

bargaining positions, but not necessarily on the level of asset specificity (since high and low levels of specificity can yield the same spot price). Asset specificity clearly matters, however, if the parties face the possibility of actual trade in the alternative market. I analyse so-called carrot and stick strategies where contract deviation results in a one-period trade in the alternative market before return to the relational contract. These kinds of strategies, also called mutual punishment (Myerson, 1997), are more complex to analyse, but are still more realistic than the standard grim strategies.

When the alternative market is a real alternative and the parties can choose between a relational employment contract (integration) and a relational outsourcing contract (non-integration), high levels of asset specificity induce relational outsourcing.[3] The reason is that increased specificity reduces the temptation to renege on a relational outsourcing contract, since the benefit of external trade is reduced. In a relational employment contract, however, the downstream owner has the residual control right to the good produced, so the upstream party cannot hinder the downstream party to force internal trade. Hence, asset specificity does not affect the self-enforcing conditions of the employment contract. This difference between employment contracts and outsourcing contracts makes the relative efficiency of non-integration increase with the level of asset specificity. The reduced temptation to renege on the relational outsourcing contract, due to increased specificity, makes it possible to design higher-powered incentive schemes without running the risk of opportunistic behaviour.

This link between asset specificity, contract efficiency and asset allocation seems not to be addressed in the theoretical part of the literature. Repeated game models of economic organization acknowledge that relational contracts may be a substitute for vertical integration in dealing with the problem of opportunism. They also recognize the role of reneging temptation in the design of efficient incentive contracts. But the absence of a formal comparison of relational contracts between firms and relational contracts within firms, prior to Baker, Gibbons and Murphy's important contribution, has made the separating effect of specificity hard to identify. Klein and Leffler analyse reputation effects in assuring product quality in their seminal 1981 paper. The buyer pays a price premium to the supplier to ensure that the supplier exerts effort to produce good quality. If the supplier reneges on the contract,

[3] The terms 'relational employment' and 'relational outsourcing' stem from BGM

all his potential customers get to know, and the supplier therefore loses all future sales. Hence, the alternative market disciplines against opportunistic behaviour. But Klein and Leffler do not compare relational contracts between independent parties with contracts between vertically integrated parties. Halonen (2002) recognizes the importance of reducing outside options in order to reduce the gain from contract deviation, in her dynamic version of the Hart/Moore (1990) game. But she does not relate the outside options to the difference between specific and alternative use. Since the supplier in her model only makes specific investments in *human* capital, she relates the outside option to the investing party's dependency on the asset he invests in, i.e., to what extent it is important that the investing party manages the asset. Hence, Halonen does not make any statements concerning asset specificity and vertical integration, but she recognizes that separation of strictly complementary assets can be beneficial in providing maximum punishment for deviation.

The idea that putting parties in more adverse situations may promote efficiency is also discussed in Klein (1980) and Williamson (1983). Klein refers to the case where franchisers require franchisees to rent from them, rather than own the land on which their outlet is located. This prevents opportunism since the franchiser can require the franchisee to move if the franchisee cheats. Williamson uses the concept of hostages to emphasize the importance of credible commitment. By posting hostages, that is posting a value before the transaction in order to commit to the other party, one can reduce the possibility of opportunistic behaviour and negotiate a contract with better terms. Chiu (1998) relates the importance of credible commitment directly to the concept of investment specificity. He claims that "the theoretical prediction that integration is more likely in the presence of relationship-specific investments is not as robust as previously thought". He shows that specific investments cause a threat to the relationship when outside options are attractive, not when outside options are unattractive, as the traditional hypothesis implies. But Chiu does not compare the effect specificity may have on contracts between integrated parties with the effect on contracts between non-integrated parties.

The rest of the paper is organized as follows: Section 2 briefly discusses the empirical research on the determinants of vertical integration. Section 3 presents the model. A comparative analysis is made in Section 4, while section 5 concludes.

2. The empiricism of vertical integration

There is an impressive body of empirical research that supports predictions of transaction cost economics (see Joskow 1988, and Shelanski and Klein 1995 for an overview). I believe, however, that the empirical work does not verify the hypothesis that asset specificity leads to vertical integration. It is a fact that a number of quantitative case studies and cross-sectional econometric analyses show a positive correlation between asset specificity and vertical integration. But these studies do not prove that asset specificity *leads* to vertical integration. The econometric models assume that organizational form is a function of asset specificity, uncertainty, complexity and frequency. Organizational form is the dependent variable while asset specificity is one of the independent variables. The causality between the variables is in general not discussed.

Even though many transaction cost economists claim that the vertical integration hypothesis has a substantial empirical foundation, a number of prominent economists question the empirical validity of the hypothesis. Ronald Coase has all since his famous contribution "The Nature of the Firm" (1937) doubted the importance of asset specificity in bringing about vertical integration. He is in fact sceptical to the concept of opportunism in analyses of economic organization. He argues (1988) that the importance of reputation makes it unlikely that a party would act opportunistically even if assets are specific. His experience is that businessmen find contractual arrangements to be a satisfactory answer to the possible problems of asset specificity. Holmström and Roberts (1998) point out "many of the hybrid organizations that are emerging are characterized by high degrees of uncertainty, frequency, and asset specificity, yet they do not lead to integration. In fact, high degrees of frequency and mutual dependency seem to support, rather than hinder, ongoing cooperation across firm boundaries."

The economic organization of the international oil industry may serve as good example of separated specific assets. The oil companies and their main suppliers, who design and build installations that the oil companies use to extract oil, always operate with separate ownership. But the suppliers manage capital stock and produce inputs that are highly specific to the buying oil companies. The inputs may be valuable to a competing oil company, but the technology is often tailor-made for a specific field or a specific company. The parties usually agree on a so-called EPCI-contract, in which the main suppliers are responsible for engineering, procurement, construction and installation. The parties normally agree on an even split of cost overruns and savings relative to a target sum. Hence, the contracts contain high-powered incentive schemes (for more details see Osmundsen, 1999). It is reasonable to assume that these incentive schemes would not have been feasible in an integrated solution. The specificity of the assets and the dependency between the parties makes it possible for the oil companies to design strong incentives without the risk of hold-up behaviour.

The classical empirical case of vertical integration has been the General Motors' (GM) acquisition of Fisher Body in 1926. The standard view has been that GM merged vertically with Fisher Body because of concerns over specific investments and hold-up behaviour. Several economists now question this explanation. Coase (2000) points out that GM already owned 60 percent of the shares of Fisher Body before they acquired the remaining 40 percent. He claims that there is no evidence that hold-up occurred before the merger took place. Freeland (2000) states that "far from reducing opportunistic behaviour, the vertical integration in fact increased GMs vulnerability to rent seeking behaviour based in human asset specificity". Casadesus and Spulber (2000) argue that the merger reflected economic considerations specific to that time, not some immutable market failure. The contractual arrangements and working relationship prior to the merger, they claim, exhibited trust rather than opportunism.

3. The model

Baker, Gibbons and Murphy analyse an economic environment consisting of an upstream party (U), a downstream party (D) and an asset, where both parties and the asset live forever or cease to exist simultaneously at a random date. The parties are risk neutral and share the discount factor, δ , per period. The upstream party uses the asset to produce a good that could either be used in the downstream party's production process, or put to an alternative use. In each period the upstream party chooses a vector of n actions (or investments) $\mathbf{a} = (a_1, a_2, ..., a_n)$ at a cost $c(\mathbf{a})$ which affects the value of the product both for the downstream party (Q) and for the alternative market (P). The downstream value is either high or low, where $q(\mathbf{a})$ is the probability that a high value Q_H will be realized and $1-q(\mathbf{a})$ is the probability that a low value Q_L will be realized. The alternative-use value can also be either high or low, where $p(\mathbf{a})$ is the probability that a high value, P_H , will be realized and $1-p(\mathbf{a})$ is the probability that a low value P_L will be realized. Given the upstream party's actions, the downstream and the alternative-use values are conditionally independent. It is assumed that $c(0)=q(0)=p(0)=0$, so when the upstream party decides not to take actions, he bears no costs but also has no chance of realizing the high values. It is further assumed that $P_L < P_H < Q_L < Q_H$ so that the value to the downstream party always exceeds its value in the alternative use. In other words, the asset is relationship specific. The first-best actions, $\mathbf{a}^*$, maximizes the expected value of the good in its efficient use minus the cost of action, hence the total surplus from the transaction is given by

$$S^* = \underset{a}{Max}\left[Q_L + q(\mathbf{a}^*)\Delta Q - c(\mathbf{a}^*)\right], \text{ where } \Delta Q = Q_H - Q_L.$$

The actions are unobservable to anyone but the upstream party, so contracts contingent on actions cannot be enforced. It is assumed that Q and P are observable, but not verifiable, so it is possible to design self-enforceable contracts, but not to contract on Q or P in a way that a third party can enforce.

The parties can organize their transactions through different choices of contract governance and ownership structure. With respect to ownership structure, it is assumed that asset ownership conveys ownership of the good produced, so if the upstream party owns the asset (non-integration), the downstream party cannot use the good without buying it from the upstream party, whereas if the downstream party owns the asset (integration), then he already owns the good. With respect to contract governance, the parties can agree on either a spot contract or a relational contract. In a spot contract, a spot price is negotiated for each period and is determined by ownership structure and bargaining positions. If the upstream party owns the asset, 50:50 Nash bargaining over the surplus from trade decides the spot price. If the downstream party owns the asset, he can just take the realized output without paying, so the upstream party will refuse to take costly actions. In a relational contract, the parties agree on a compensation contract $(s, b_L, b_H, \beta_L, \beta_H)$ where salary s is paid by downstream to upstream at the beginning of each period, and b_i is supposed to be paid when Q_i is realized, *(i = H, L)* and β_j when P_j is realized, *(j = H, L)*. For example: If the upstream party produces a good which yields a high value in the specific relation, Q_H, and a low value in the alternative market, P_L, the downstream party should, according to the contract, pay the bonuses $b_H + \beta_L$ to the upstream party.[4] Such a contract induces the upstream party to yield effort even if he doesn't own the asset. Since the contract cannot be enforced by a third party, the parties will honour the contract only if the present value of honouring is greater than the present value of reneging.

BGM's taxonomy of organizational design is summarized as follows (see BGM, *QJE* pp.46):

	Non-integration	**Integration**
Spot contract	*Spot outsourcing (SO)*	*Spot employment (SE)*
Relational contract	*Relational outsourcing (RO)*	*Relational employment (RE)*

So far, I have been following BGM's set-up. In this paper I will compare relational outsourcing with relational employment using other player strategies than the grim trigger strategies analysed by BGM. In BGM, if a party reneges on a contract, the other party refuses to enter into a new relational contract with that party. Instead, they agree to trade in spot governance forever after. In this paper, however, if one of the parties reneges, they first agree on a spot price (as in BGM). In the next period, the party who did not renege punishes the other party by refusing to enter into any agreement (including a spot agreement) and instead chooses to trade in the alternative market. After this "punishment phase" the parties return to a relational contract (see strategy specifications below). These kinds of trigger strategies are in the literature referred to as mutual punishment strategies, carrot and stick strategies, or two-phase punishment strategies (see Gibbons, 1992).

BGM's strategy specifications have the advantage of both being simple to analyse and making it possible to compare all four organizational forms within the same framework. In the modification studied here, it is simply assumed that specificity deters spot contracting from being a long-term option. Still, there are several reasons for making this modification.

[4] BGM start up with a more general contract $(s, b_{HH}, b_{HL}, b_{LH}, b_{LL})$, *(i,j=H,L)*, but restrict it to $(b_{HH} = b_H + \beta_H, b_{HL} = b_H + \beta_L, b_{LH} = b_L + \beta_H, b_{LL} = b_L + \beta_L)$ in order to simplify the comparative analysis.

First, it can be argued that the carrot and stick strategy is more realistic than the grim strategy, especially in buyer/supplier relationships with high levels of asset or investment specificity. It is difficult to understand why the parties would stick to spot governance forever after a contract breach when specificity makes relational contracting significantly more efficient than spot contracting. Carrot and stick strategies are more in line with actual economic behaviour off the cooperation path. In the offshore industry, for example, contract breach often results in operators i) renegotiating the terms of the current project, (ii) searching for new long term trading partners, while trading directly in inferior spot markets iii) entering into a new long term contracts with either the old trading partner or a new one (see e.g the Norsok reports, 1995). Second, analysing carrot and stick strategy equilibria is more appropriate if asset specificity is regarded as a significant explanatory variable. In order to analyse the effect of asset specificity in long term contracts, the alternative market must be modelled as a real threat point, not merely a reference point for spot negotiations. In BGM the level of asset specificity does not affect the robustness of relational contracts. In the present paper, however, asset specificity does affect the parties' temptation to renege on relational contracts. Third, both grim strategies and carrot and stick strategies yield the same surplus, for given actions, in equilibrium. But for sufficiently high levels of specificity, efficient relational contracts can be implemented for lower discount factors when the parties play the carrot and stick strategy than when they play the grim strategy. This provides an argument for studying carrot and stick strategies in the presence of specificity.

In this paper, the strategy for U (D) is specified as follows:

1. In period t, honour the terms of the relational contract $(s, b_L, b_H, \beta_L, \beta_H)$ if D (U) honoured in period t-1.
2. In period t, honour the terms of the relational contract $(s, b_L, b_H, \beta_L, \beta_H)$ if there was no trade with D (U) in period t-1.
3. In period t, refuse to trade with D (U) if the trade between the parties in period t-1 was accomplished by spot contracting.

To "honour the terms of the relational contract" means for the upstream party to accept the bonuses offered and for the downstream party to pay the promised bonuses. We enter this game ex post quality realizations in period t. When the parties are to decide whether to honour or renege on the contract, they know the quality realizations of period t, but can only have expectations regarding the remaining periods. The parties honour the contract if the present value of honouring exceeds the present value of reneging. A relational contract is self-enforcing if both parties choose to honour the contract $(s, b_L, b_H, \beta_L, \beta_H)$ for all possible realizations of Q_i and P_j. The critical part of the analysis is to deduce the conditions for when the relational employment contract and the relational outsourcing contract are self-enforcing. Technically, these are conditions for when the strategies specified above constitute subgame perfect Nash equilibria of relational contracts. See appendix on subgame perfection.

Before we proceed, consider four additional assumptions: First, it is assumed that both parties incur a switching cost v by trading in the alternative market when the product has already been produced for the purpose of trading in the specific relation.[5] They avoid this cost if they know ex ante that no trade will occur between the parties. Second, in contrast to BGM, it is assumed that ownership is fixed on the "punishment path". This seems realistic as long as the strategies, in case of deviation, specify only one period of spot governance. Only small negotiation costs would make a one-period ownership transfer inefficient (Halonen (2002) fixes ownership forever after deviation even in grim trigger strategies). Third, BGM assume that C(0) yields Q_L always. But it is more realistic and thus assumed in this paper, that if the upstream party takes no costly actions, he can choose between realizing Q_L and realizing zero values. This gives the upstream party a punishment possibility even if the downstream party owns the asset.[6] Fourth, it is assumed that the downstream party's valuation of the alternative market goods is equal to the price he has to pay. Hence, if the downstream party buys the good in the alternative market, he earns no surplus from this trade.[7] None of these assumptions changes the quality of the results in this paper, but they are made both for analytical convenience and in order to make the upstream-downstream relationship as realistic as possible.

[5] We can view these costs as time costs or extra transport costs associated with the unexpected move from relational trade to alternative market trade.

[6] This assumption will not change the downstream payoff function in relational contracts since Q_i will always be realized in relational contracts equlibria. Also note that this does not mean that upstream can hold-up the good in relational employment ex post realization. The choice of realizing Q_L or zero is taken ex ante.

[7] BGM say nothing about this since the parties never trade in the alternative market in their model.

3.1 Relational employment

If the upstream party is confident that the downstream party will honour the contract $(s, b_L, b_H, \beta_L, \beta_H)$ then the upstream party will choose actions $\mathbf{a}^{RE}$ that solve: $\underset{a}{Max}\,(s + b_L + \Delta b q(\mathbf{a}) + \beta_L + \Delta\beta\, p(\mathbf{a}) - c(\mathbf{a})) \equiv U^{RE}$ where $\Delta b = b_H - b_L$, $\Delta\beta = \beta_H - \beta_L$, and superscript (RE) denote relational employment. The expected downstream payoff is then $E(Q_i - b_i - \beta_j - s | \mathbf{a} = \mathbf{a}^{RE}) = Q_L + \Delta Q q(\mathbf{a}^{RE}) - s - b_L - \Delta b q(\mathbf{a}^{RE}) - \beta_L - \Delta\beta p(\mathbf{a}^{RE}) \equiv D^{RE}$ so total surplus under relational employment is $S^{RE} \equiv U^{RE} + D^{RE} = Q_L + q(\mathbf{a}^{RE})\Delta Q - c(\mathbf{a}^{RE})$.

Given that the downstream party always honours the contract, the upstream party will earn $s + b_i + \beta_j - c\ (\mathbf{a}^{RE})$ in period *t*, and expect a total of $\frac{\delta}{1-\delta}(s + b_L + \Delta b q(\mathbf{a}^{RE}) + \beta_L + \Delta\beta\, p(\mathbf{a}^{RE}) - c(\mathbf{a}^{RE})) = \frac{\delta}{1-\delta} U^{RE}$ from future trade if he honours the contract. To make the different payoffs easy to compare, I distinguish between period *t*, period *t+1* and all the remaining periods. The present value of honouring the relational employment contract is thus written

$$s + b_i + \beta_j - c\ (\mathbf{a}^{RE}) + \delta U^{RE} + \tfrac{\delta^2}{1-\delta} U^{RE}\ .$$

If the upstream party reneges on the contract in period *t* by refusing to accept the promised payment $b_i + \beta_j$ (or refusing to make a promised payment if $b_i + \beta_j < 0$), the trade is accomplished by spot contracting, where the downstream party, as the asset owner, just takes the good and leave the upstream party with nothing. According to the specified strategies, there is no trade between the parties in period *t+1*, so the upstream party earns nothing and bears no investment costs. In period *t+2* the relational contract is re-established. The payoff after reneging is then

$$s - c\ (\mathbf{a}^{RE}) + \tfrac{\delta^2}{1-\delta} U^{RE}.$$

The upstream party will thus honour rather than renege on the relational employment contract when, for all values of *i* and *j*,

(1) $b_i + \beta_j + \delta U^{RE} \geq 0$.

Given that the upstream party always honours the contract, the downstream party's payoff from honouring the relational employment contract is

$$Q_i - b_i - \beta_j - s + \delta D^{RE} + \tfrac{\delta^2}{1-\delta} D^{RE} .$$

If the downstream party reneges on the contract in period t, he will just take the realized value, Q_i, and pay nothing. In period $t+1$ the upstream party will refuse to produce the good, so the downstream party has to buy the good in the alternative market. He will not gain a surplus on this trade, since his valuation of this non-specific good is equal to the price he has to pay. In period $t+2$ the relational employment contract is re-established. The present value of reneging on the contract is thus simply

$$Q_i - s + \tfrac{\delta^2}{1-\delta} D^{RE} .$$

The downstream party will honour rather than renege on the relational employment contract when, for all values of *i* and *j*,

(2) $b_i + \beta_j \leq \delta D^{RE}$.

(1) and (2) represent 8 constraints that have to hold in order for the relational employment contract to be self-enforcing. Combining these restrictions yields (see appendix):

(3) $|\Delta b| + |\Delta \beta| \leq \delta S^{RE}$.

This is both a necessary and a sufficient constraint for the relational contract $(s, b_L, b_H, \beta_L, \beta_H)$ to hold, since the parties can always choose a fixed salary s that satisfies both (1) and (2). The efficient relational employment contract maximizes total surplus, S^{RE}, subject to (3).

3.2 Relational outsourcing

In relational outsourcing, if the upstream party is confident that the downstream party will honour the contract $(s, b_L, b_H, \beta_L, \beta_H)$ the upstream party chooses actions $\mathbf{a}^{RO}$ that solve $\underset{a}{Max}\ (s + b_L + \Delta b q(\mathbf{a}) + \beta_L + \Delta\beta p(\mathbf{a}) - c(\mathbf{a})) \equiv U^{RO}$ where superscript RO denotes relational outsourcing. The downstream party's payoff is then $E(Q_i - s - b_i - \beta_j | \mathbf{a} = \mathbf{a}^{RO}) = Q_L + \Delta Q q(\mathbf{a}^{RO}) - s - b_L - \Delta b q(\mathbf{a}^{RO}) - \beta_L - \Delta\beta p(\mathbf{a}^{RO}) \equiv D^{RO}$, so total surplus under relational outsourcing is $S^{RO} \equiv U^{RO} + D^{RO} = Q_L + q(\mathbf{a}^{RO})\Delta Q - c(\mathbf{a}^{RO})$.

If the upstream party honours the relational outsourcing contract he will receive

$$s + b_i + \beta_j - c\ (\mathbf{a}^{RO}) + \delta U^{RO} + \tfrac{\delta^2}{1-\delta} U^{RO}\ .$$

If the upstream party reneges on the contract in period *t*, trade is accomplished by spot contracting. Since the upstream party now owns the asset, the downstream party cannot just take the good. I assume, like BGM, that the parties set prices by means of 50:50 Nash negotiations, which yields $\frac{1}{2}(Q_i + P_j - v)$.[8] In period *t+1*, downstream refuses to trade with upstream. Anticipating this, upstream chooses actions $\mathbf{a}^{AO}$, which solve $\underset{a}{Max}(P_L + \Delta P p(\mathbf{a}) - c(\mathbf{a}\ )) \equiv U^{AO}$. In period *t+2* the parties re-establish their relational contract.[9]

The upstream party's payoff after reneging is then

$$s + \tfrac{1}{2}(Q_i + P_j - v) - c\ (\mathbf{a}^{RO}) + \delta U^{AO} + \tfrac{\delta^2}{1-\delta}\ U^{RO}.$$

---

[8] The downstream party will pay the upstream party the alternative value $P_j - v$ plus half the surplus from trade with the downstream party: $\frac{1}{2}(Q_i - (P_j - v))$, i.e. $\frac{1}{2}(Q_i + P_j - v)$.

[9] The strategy in which the no-trade-punishment is deferred until period t+1 coincides with subgame perfect equilibrium for *v* exceeding a critical level (see Appendix). According to the specified strategies, the parties know that trade in the alternative market follows after spot governance. Hence, they avoid the switching cost *v* if they defer the trade in the alternative market from *t* until *t+1*.

The upstream party will thus honour the contract when, for all values of *i* and *j*,

(4) $$b_i + \beta_j + \delta U^{RO} \geq \tfrac{1}{2}(Q_i + P_j - v) + \delta U^{AO}.$$

If the downstream party honours the contract he will earn

$$Q_i - b_i - \beta_j - s + \delta D^{RO} + \tfrac{\delta^2}{1-\delta} D^{RO}.$$

If the downstream party reneges in period t, the parties agree on the 50:50 Nash price so that the downstream party earns $Q_i - s - \frac{1}{2}(Q_i + P_j - v)$. In period $t+1$ upstream refuses to trade with downstream, who has to buy the good in the alternative market and thus gains no surplus. The downstream party's payoff after reneging is then

$$Q_i - s - \tfrac{1}{2}(Q_i + P_j - v) + \tfrac{\delta^2}{1-\delta} D^{RO}.$$

The downstream party will thus honour the contract, for all values of *i* and *j*, when

(5) $$\delta D^{RO} \geq b_i + \beta_j - \tfrac{1}{2}(Q_i + P_j - v)$$

Combining (4) and (5) yields the following condition for the relational outsourcing contract to be self-enforcing (see appendix):

(6) $$|\Delta b - \tfrac{1}{2}\Delta Q| + |\Delta\beta - \tfrac{1}{2}\Delta P| \leq \delta(S^{RO} - U^{AO})$$

Like (3), (6) is both necessary and sufficient. The efficient relational outsourcing contract maximizes total surplus S^{RO} subject to (6).

4. Comparative analysis

We can now compare relational outsourcing with relational employment. First, observe that (3) and (6) underscore BGM's main proposition: The parties' temptation to renege on a given relational contract depends on asset ownership.[10] I will now show how asset and investment specificity affect the parties reneging temptations under different types of ownership. Define $Q_L - P_L$ as the level of asset specificity, and $\Delta Q - \Delta P$ as the level of investment specificity. Now, observe that in relational outsourcing the value of the upstream party's outside option, $U^{AO} \equiv P_L + \Delta P p(\mathbf{a}^{AO}) - c(\mathbf{a}^{AO})$, is part of the relational contract constraint. In relational employment, however, the outside option is equal to zero for any level of P_H and P_L. Hence, the levels of both asset specificity and investment specificity affect the relational outsourcing constraint, but not the relational employment constraint. In relational outsourcing the downstream party's temptation to renege is lower than in relational employment, since he cannot just take the good, but has to bargain a spot price with the upstream owner. On the other hand, the upstream party's temptation to renege is higher under relational outsourcing than under relational employment, because of his outside options. Under relational outsourcing, increased specificity will thus reduce the relative value of the upstream party's outside option, and thereby give scope for better relational contracts.

From (3) and (6) we observe that increasing incentive intensity, given by $\Delta b, \Delta\beta$, increases the total temptation to renege on a contract. Low bonuses may induce the upstream party to renege, while high bonuses may induce the downstream party to renege. Moreover, we observe that if U^{AO} is sufficiently low, then there is scope for higher-powered incentives in relational outsourcing than in relational employment. Hence, if the level of asset specificity is sufficiently high, which implies that $S^{RO} - U^{RO}$ is high, and high-powered incentives are desirable, then relational employment is inefficient compared to relational outsourcing. We gain intuition by thinking through an incentive for downstream to increase the specificity of an asset. If the upstream party possesses an asset that is highly valuable to a broad market, downstream may wish to acquire the upstream party's asset in order to avoid strategic behaviour. The problem then is that the downstream party's incentive to cheat on upstream

[10] Olsen (1996) has a related result, showing in a two-period model that the choice of renegotiating a contract depends on organizational form

increases, so upstream may call for lower-powered incentive schemes and higher fixed salaries. But if higher-powered incentives are desirable, he can make tailor-made investments in the asset in a manner that increases its internal, but not its external value. Then he can safely outsource the asset to upstream, achieving higher-powered incentives without running the risk of upstream opportunism.

I will now derive a formal result showing that relational outsourcing can be an efficient response to high levels of specificity. Assume that the two gradients of partial derivates $\frac{\partial q}{\partial a_i}(a^{FB}), \frac{\partial p}{\partial a_i}(a^{FB}),\ i = 1,2,\ldots,n$ are linearly independent (superscript FB denotes first-best). Then a first-best solution can only be achieved if $\Delta b = \Delta Q$ and $\Delta\beta = 0$.

Given (3), first-best can be achieved under a relational employment contract if

$$\delta \geq \frac{\Delta Q}{Q_L + \Delta Q q(\mathbf{a}^{FB}) - c(\mathbf{a}^{FB})} = \delta^{RE}.$$

Given (6), first-best can be achieved under a relational outsourcing contract if

$$\delta \geq \frac{\frac{1}{2}(\Delta Q + \Delta P)}{Q_L + \Delta Q q(\mathbf{a}^{FB}) - c(\mathbf{a}^{FB}) - P_L - \Delta P p(\mathbf{a}^{AO}) + c(\mathbf{a}^{AO})} = \delta^{RO}.$$

Hence, to be able to implement first-best at equal or lower discount factors in the outsourcing contract than in the employment contract, we must have $\delta^{RO} \leq \delta^{RE}$, that is

$$(7) \qquad \left(Q_L + \Delta Q q(\mathbf{a}^{FB}) - c(\mathbf{a}^{FB})\right)\left(\frac{\Delta Q - \Delta P}{2\Delta Q}\right) - \left(P_L + \Delta P p(\mathbf{a}^{AO}) - c(\mathbf{a}^{AO})\right) \geq 0.$$

We can then state:

Proposition 1: *i) Assume that there is investment specificity, defined as* $\Delta Q > \Delta P$. *If asset specificity is sufficiently large, in the sense that* Q_L *is sufficiently high and/or* P_L *is sufficiently low, then there exist critical discount factors* $\delta^{RE} > \delta^{RO} > 0$ *such that for* $\delta > \delta^{RO}$ *relational outsourcing is first-best and thus at least as efficient as relational employment, and for* $\delta^{RE} > \delta > \delta^{RO}$ *relational outsourcing is strictly more efficient than relational employment. ii) Assume that there is no investment specificity* ($\Delta Q < \Delta P$). *Then for any level of asset specificity, i.e. for any level of* Q_L *and* P_L, *there exist critical discount factors* $\delta^{RO} > \delta^{RE} > 0$ *such that for* $\delta > \delta^{RE}$ *relational employment is first-best and thus at least as efficient as relational outsourcing, and for* $\delta^{RO} > \delta > \delta^{RE}$ *relational employment is strictly more efficient than relational outsourcing.*

Proof: Given $\Delta Q > \Delta P$, *the left hand side of (7) is strictly increasing in* Q_L *and strictly decreasing in* P_L. *Given* $\Delta Q < \Delta P$ *, (7) never holds.*

I will show, for specific functions, that (7) is also a valid condition in second-best solutions. That is, given (7), relational outsourcing is always an equally efficient or more efficient solution than relational employment. I assume, like BGM, that the upstream party can take two actions: $\mathbf{a} = (a_1, a_2)$, and that the production functions are linear and the cost function quadratic:

(8)
$$q(a_1,a_2) = q_1a_1 + q_2a_2$$
$$p(a_1,a_2) = p_1a_1 + p_2a_2$$
$$c(a_1,a_2) = \tfrac{1}{2}a_1^2 + \tfrac{1}{2}a_2^2$$

where $q_1, q_2, p_1, p_2 \geq 0$ and $q_1p_2 \neq q_2p_1$.

The first-best actions are then $a_1^{FB} = q_1\Delta Q$ and $a_2^{FB} = q_2\Delta Q$. In both the outsourcing contract and the employment contract, the upstream party chooses to maximize $s + b_L + (q_1a_1 + q_2a_2)\Delta b + \beta_L + (p_1a_1 + p_2a_2)\Delta\beta - \tfrac{1}{2}a_1^2 - \tfrac{1}{2}a_2^2$, so that $a_1 = q_1\Delta b + p_1\Delta\beta$

and $a_2 = q_2\Delta b + p_2\Delta\beta$. A first-best solution can then only be achieved if $\Delta b = \Delta Q$ and $\Delta\beta = 0$

In order to keep it simple, I assume that $q_2 = p_1 = 0$. The agent can take one action that affects Q and another action that affects P.

Given (8), (7) can be written

$$(7') \quad Q_L\left(\frac{\Delta Q - \Delta P}{2\Delta Q}\right) - P_L - \left(\tfrac{1}{2}p_2{}^2\Delta P^2 - \tfrac{1}{4}q_1{}^2\Delta Q(\Delta Q - \Delta P)\right) \geq 0.$$

Given (7'), first-best cannot be achieved if

$$(9) \quad \delta < \frac{\tfrac{1}{2}(\Delta Q + \Delta P)}{Q_L + \tfrac{1}{2}q_1{}^2\Delta Q^2 - P_L - \tfrac{1}{2}p_2{}^2\Delta P^2} = \delta^{RO}.$$

Proposition 2: Given (7') and (8), *relational outsourcing is at least as efficient as relational employment if* $\delta \geq \delta^{RO}$ *and strictly more efficient than relational employment if* $\delta < \delta^{RO}$.

Proof: *see appendix.*

The propositions suggest that outsourcing may be an efficient response to high levels of specificity. Note the relationship between asset specificity, investment specificity and governance in proposition 1. If there is no investment specificity, relational outsourcing is always an inefficient governance mechanism compared to relational employment. Moreover, if there is investment specificity, relational outsourcing is an efficient response to increased asset specificity.

The propositions help elucidate anecdotic empirical evidence and case studies showing that non-integration is highly compatible with asset/investment specificity. And further, that specificity can actually be beneficial for non-integrated solutions. The proposition may also cast some light on empirical studies questioning other aspects of "Williamsonian" explanations of integration and outsourcing. Anderson, Glenn and Sedatole (2000) make an interesting empirical study of the relationship between asset complexity and outsourcing decisions. Using data on 156 sourcing decisions for process tooling (dies) of a new car program, they found that attributes that according to transaction cost economics favoured "insourcing", favours outsourcing if the parties engage in relational contracting. In particular, they found that firms outsourced parts with high levels of complexity, and insourced simple parts with low levels of complexity. Also, parts with high levels of "design constraints" were more likely to be outsourced than parts with low design constraint levels. Problems of strategic behaviour from these relational-dependent external suppliers were relatively small, and field investigations suggested that the external suppliers were more responsive to incentives than internal suppliers.

5. Concluding remarks

The model in this paper identifies local non-monotonic relationships between asset specificity and vertical integration. In vertically integrated firms, there will always be some kind of complementarity between the assets, and the assets of an upstream party vertically integrated with its downstream buyer will always to a certain extent be specific to the downstream party's needs. If there is no asset specificity or investment specificity, we have a competitive market with no need for contractual incentive schemes. Then the "old rule" would apply, saying that the best manager of an asset is its owner. So if the upstream party is the one taking actions, he should also own the asset. But if we are in an economic environment with significant levels of specificity, as assumed in the model of this paper, then the relationship between asset specificity and vertical integration becomes more complex. To a certain extent, specificity may induce integration, as the downstream party wishes to avoid unfavourable strategic behaviour from upstream. But if the level of specificity is sufficiently high, the relative value of external trade is reduced, and the incentive for upstream to behave opportunistically is reduced as well. Integration may then be an inefficient governance solution compared with non-integration: If the parties can engage in relational contracts, and the surplus from external trade is relatively small compared with the surplus from trade in the specific relationship, the parties will be able to design higher-powered incentive schemes if upstream owns the asset than if downstream owns the asset.

APPENDIX

1. The conditions for honouring the relational employment contract

The upstream party's condition is given by

(1) $b_i + \beta_j + \delta U^{RE} \geq 0$

The downstream party's condition is given by

(2) $b_i + \beta_j \leq \delta D^{RE}$

Since $i=H,L$ and $j= H,L$, each of these two conditions contains four constraints. We see that the high quality realisation always imposes the binding constraint on the downstream party, while low quality realisation imposes the relevant constraint on the upstream party. The relevant constraints are then:

$$(b_L + \beta_L) + \delta U^{RE} \geq 0$$
$$(b_L + \Delta b + \beta_L + \Delta\beta) \leq \delta D^{RE}$$

Multiplying the upstream constraint by (-1) and adding the downstream constraint yields the following necessary and sufficient condition for honouring the relational employment contract:

(3) $|\Delta b| + |\Delta\beta| \leq \delta S^{RE}$

2. The conditions for honouring the relational outsourcing contract

The upstream party's condition is given by:

(4) $b_i + \beta_j + \delta U^{RO} \geq \frac{1}{2}(Q_i + P_j - v) + \delta U^{AO}$

The downstream party's condition is given by:

$$(5) \qquad \delta D^{RO} \geq b_i + \beta_j - \tfrac{1}{2}(Q_i + P_j - v)$$

It is now less obvious which constraints are binding. But there will always be two constraints at most that are binding. We see that it depends on the differences: $\tfrac{1}{2}\Delta Q - \Delta b$ and $\tfrac{1}{2}\Delta P - \Delta\beta$.

When $\tfrac{1}{2}\Delta Q > \Delta b$ and $\tfrac{1}{2}\Delta P > \Delta\beta$, the relevant constraints are:

$$\tfrac{1}{2}(Q_L + \Delta Q + P_L + \Delta P - v) + \delta U^{AO} \leq b_L + \Delta b + \beta_L + \Delta\beta + \delta U^{RO}$$

$$\tfrac{1}{2}(Q_L + P_L - v) + \delta\left(Q_L + \Delta Q q(a^{RO})\right) \geq b_L + \beta_L + \delta\left(b_L + \Delta b q(a^{RO}) + \beta_L + \Delta\beta p(a^{RO})\right)$$

When $\tfrac{1}{2}\Delta Q > \Delta b$ and $\tfrac{1}{2}\Delta P < \Delta\beta$, the relevant constraints are:

$$b_L + \Delta b + \beta_L + U^{RO} \geq \tfrac{1}{2}(Q_L + \Delta Q + P_L - v) + \delta U^{AO}$$

$$\delta D^{RO} \geq b_L + \beta_L + \Delta\beta_L - \tfrac{1}{2}(Q_L + P_L - v)$$

When $\tfrac{1}{2}\Delta Q < \Delta b$ and $\tfrac{1}{2}\Delta P > \Delta\beta$, the relevant constraints are:

$$b_L + \beta_L + \Delta\beta_H + U^{RO} \geq \tfrac{1}{2}(Q_L + \Delta Q + P_L - v) + \delta U^{AO}$$

$$\delta D^{RO} \geq b_L + \Delta b + \beta_L - \tfrac{1}{2}(Q_L + P_L - v)$$

When $\tfrac{1}{2}\Delta Q < \Delta b$ and $\tfrac{1}{2}\Delta P < \Delta\beta$, the relevant constraints are:

$$b_L + \beta_L + U^{RO} \geq \tfrac{1}{2}(Q_L + \Delta Q + P_L - v) + \delta U^{AO}$$

$$\delta D^{RO} \geq b_L + \Delta b + \beta_L + \Delta\beta - \tfrac{1}{2}(Q_L + P_L - v)$$

Multiplying the downstream party's constraints by (-1) and adding the upstream party's constraints yields a necessary and sufficient condition for each pair of constraints:

$$(6) \qquad \left|\Delta b - \tfrac{1}{2}\Delta Q\right| + \left|\Delta\beta - \tfrac{1}{2}\Delta P\right| \leq \delta\left(S^{RO} - U^{AO}\right)$$

3. The conditions for subgame perfect equilibria:

From Selten (1965), a Nash equilibrium is subgame perfect if the players' strategies constitute a Nash equilibrium in every subgame. In this game we have an infinite number of subgame divided into three categories: The games that start after trade governed by a relational contract, the games that start after trade governed by a spot contract, and the games that start after no trade between the parties / trade in the alternative market. The carrot and stick strategies constitute subgame perfect equilibrium if U (D), in case of D (U) deviation in period t, finds it optimal to trade under spot governance, S, in period t; refuses to trade with D (U), i.e trades in the alternative market, A, in period $t+1$; and returns to relational contracting, R, in period $t+2$. We can write this "punishment path" $(S_t, A_{t+1}, R_{t+2}, R_{t+3} ...)$. U (D)'s feasible set of trade actions depends on D (U)'s offer. At the end of each period, the players have taken the same action, but in terms of feasibility A dominates S which dominates R.

There are an infinite number of strategies specifying punishment paths that could constitute subgame perfect equilibria. With the strategies specified here, we can, however reduce the relevant paths to $(S_t, A_{t+1}, R_{t+2}, R_{t+3},...)$ and $(A_t, R_{t+1}, R_{t+2},...)$. Recall that when identifying the conditions for subgame perfection, it is commonly assumed that U (D) assumes that D (U) follows his initial strategy after deviations. Hence, it is not possible for U (D) after D (U) deviation in period t to postpone the trade in the alternative market, for instance to play $(S_t, S_{t+1}, A_{t+2}, R_{t+3}, R_{t+4} ...,)$ since according to his initial strategy, D (U) will play A in period $t+1$. Hence, in addition to the strategy specified path $(S_t, A_{t+1}, R_{t+2}, R_{t+3},...)$, we are left with the "competing" path $(A_t, R_{t+1}, R_{t+2},...)$. No path will include more than one period of trade in the alternative market since A yields the lowest surplus. Also note that in the model it is assumed that a player who reneges on the relational contract offers spot contracting instead of direct trade in the alternative market. Then if $(S_t, A_{t+1}, R_{t+2}, R_{t+3},...)$ dominates $(A_t, R_{t+1}, R_{t+2},...)$, deviation starting with spot contracting dominates direct trade in the alternative market.

Relational outsourcing: If the downstream party reneges, the upstream party's punishment path $(S_t, A_{t+1}, R_{t+2}, R_{t+3},...)$ dominates $(A_t, R_{t+1}, R_{t+2},...)$ if

(A.1) $\frac{1}{2}(Q_i + P_j - v) + \delta U^{AO} \geq P_j - v + \delta U^{RO}$

i.e. $v \geq 2\delta(U^{RO} - U^{AO}) - (Q_i - P_j)$.

If the upstream party reneges, the downstream party's punishment path $(S_t, A_{t+1}, R_{t+2}, R_{t+3} \ldots)$ dominates $(A_t, R_{t+1}, R_{t+2}, \ldots)$ if

(A.2) $Q_i - \frac{1}{2}(Q_i + P_j - v) \geq -v + \delta\ \mathrm{D}^{\mathrm{RO}}$

i.e. $v \geq \frac{2}{3}\delta\ \mathrm{D}^{\mathrm{RO}} - \frac{1}{3}(Q_i - P_j)$

For sufficiently high switching costs, the upstream party (downstream party) will play $(S_t, A_{t+1}, R_{t+2}, R_{t+3}, \ldots)$ in case of downstream party (upstream party) deviation. For the strategies to constitute subgame perfect equilibrium, (A.1) and (A.2) must hold for the critical discount factor that is necessary for (6) to hold with equality. Note that for sufficiently high levels of specificity (A.1) and (A.2) hold for $v=0$.

Relational employment: In relational employment, the upstream party cannot trade in the alternative market, but he can refuse to trade by not producing the good. But if the downstream party reneges on the contract in period t, the upstream party cannot refuse to trade with the downstream party in this period, since he has already realized Q_i. Hence the upstream party cannot play and is thus "forced" to follow the strategy-specified punishment path $(S_t, A_{t+1}, R_{t+2}, R_{t+3}, \ldots)$.

If the upstream party reneges in period t, the downstream party have no incentive to play$(A_t, R_{t+1}, R_{t+2}, \ldots)$ since in period t he can just take the realized Q_L. Hence, he follows the strategy-specified punishment path $(S_t, A_{t+1}, R_{t+2}, R_{t+3}, \ldots)$.

Conclusion: The strategies constitute a subgame perfect equilibrium if (3), (6), (A.1) and (A.2) hold.

4. Proof of proposition 2

For notational simplicity:

$\Delta b = x$

$\Delta \beta = y$

$Q_L = Q$

$P_L = P$

$\Delta Q = z$

$\Delta P = w$

$q_1 = q$

$p_2 = p$

Given the functional forms specified in (8) and the assumption that $q_2 = p_1 = 0$, the surplus from a relational contract is given by

$$S(x,y) = Q + q^2 xz - \tfrac{1}{2}q^2 x^2 - \tfrac{1}{2}p^2 y^2$$

The outsourcing constraint is given by

$$\left|x - \tfrac{1}{2}z\right| + \left|y - \tfrac{1}{2}w\right| \le \delta(Q + q^2 xz - \tfrac{1}{2}q^2 x^2 - \tfrac{1}{2}p^2 y^2 - P - \tfrac{1}{2}p^2 w^2)$$

Geometry suggests that the solution is found in the area $\tfrac{1}{2}z < x$ and $\tfrac{1}{2}w > y$. The maximization problem can then be written

$$\underset{x,y}{Max}\ S(x,y)$$

subject to

$$x - \tfrac{1}{2}z + \tfrac{1}{2}w - y \le \delta(Q + q^2 xz - \tfrac{1}{2}q^2 x^2 - \tfrac{1}{2}p^2 y^2 - P - \tfrac{1}{2}p^2 w^2)$$

Solving for x and y, and then substituting them into the surplus function yields

$$S^{RO} = \frac{1}{2}\frac{q^2\left(\delta p^2(z+w)+\delta^2 p^2(2P+p^2w^2)-2\right)-2p^2}{\delta^2 q^2 p^2}$$
$$+\frac{\sqrt{q^2+p^2}\sqrt{p^2+q^2\left(\delta^2 p^2(q^2z^2-p^2w^2)+1+2\delta^2 p^2(Q-P)-\delta p^2(z+w)\right)}}{\delta^2 q^2 p^2}$$

Let us now look at the relational employment contract. The constraint is given by

$$|x|+|y| \leq \delta(Q+q^2xz-\tfrac{1}{2}q^2x^2-\tfrac{1}{2}p^2y^2)$$

Geometry suggests that $y=0$. Assuming that $x \geq 0$, the maximization problem can be written

$$\underset{x}{Max}\, S(x,0)$$

subject to $x \leq \delta(Q+q^2xz-\tfrac{1}{2}q^2x^2)$

Solving for x and then substituting into the surplus function yields:

$$S^{RE} = \frac{\delta q^2 z - 1 + \sqrt{\delta^2 q^4 z^2 - 2\delta q^2 z + 1 + 2\delta^2 q^2 Q}}{q^2\delta^2}$$

Will now show that when (7') and (9) hold, we have

$S^{RO} \geq S^{RE}$
i.e
$f(\delta, P, q, p, z, w, Q) \geq 0$

where

$$f(\delta, P, q, p, z, w, Q) = 2p^2q^2\delta^2(S^{RO}(P)-S^{RE}) =$$

$$-q^2\delta p^2 z + q^2\delta p^2 w + 2q^2\delta^2 p^2 P + q^2\delta^2 p^4 w^2 - 2q^2$$
$$+2\sqrt{(p^2+q^2)}\sqrt{q^2\left[\delta^2 p^2(z^2 q^2 - w^2 p^2) + 1 + 2\delta^2(Q-P)p^2 - \delta(w+z)p^2\right] + p^2}$$
$$-2p^2\sqrt{(\delta^2 q^4 z^2 - 2\delta q^2 z + 1 + 2\delta^2 q^2 Q)}$$

From (7') we have

$$P \le Q(\tfrac{z-w}{2z}) - (\tfrac{1}{2}p^2 w^2 + \tfrac{1}{4}zwq^2 - \tfrac{1}{4}q^2 z^2) = P_0$$

Will now show that

$f(\delta, P, q, p, z, w, Q) \ge f(\delta, P_0, q, p, z, w, Q) \ge 0$

for every δ for the special case $p = q = 1$.

For $p = q = 1$ we have

$P_0(w, z, Q) = Q(\tfrac{z-w}{2z}) - (\tfrac{1}{2}w^2 + \tfrac{1}{4}zw - \tfrac{1}{4}z^2)$

and

$f(\delta, P, 1, 1, z, w, Q) =$

$$-\delta z + \delta w + 2\delta^2 P + \delta^2 w^2 - 2 + 2\sqrt{2}\sqrt{\delta^2 z^2 - \delta^2 w^2 + 2 + 2\delta^2 Q - 2\delta^2 P - \delta w - \delta z)}$$
$$-2\sqrt{(\delta^2 z^2 - 2\delta z + 1 + 2\delta^2 Q)}$$

Note that $P_0 \ge 0$ requires $w \le z$. For $P = P_0(w, z, Q)$ we get

$f(\delta, P_0(w, z, Q), 1, 1, z, w, Q) =$

$$\left[\delta - \delta^2(Q\tfrac{1}{z} + \tfrac{1}{2}z)\right]w - \delta z + \delta^2(Q + \tfrac{1}{2}z^2) - 2$$
$$+2\sqrt{2}\sqrt{2 + \delta^2(Q + \tfrac{1}{2}z^2) - \delta z + \left[(Q\tfrac{1}{z} + \tfrac{1}{2}z)\delta^2 - \delta\right]w} - 2\sqrt{\delta^2(z^2 + 2Q) - 2\delta z + 1}$$

$$= A(\delta, z, Q)w + B(\delta, z, Q) - 2 + 2\sqrt{2}\sqrt{2 + B(\delta, z, Q) - A(\delta, z, Q)w} - 2\sqrt{C(\delta, z, Q) + 1}$$

where

$$A(\delta,z,Q)=\left[1-\delta(Q\tfrac{1}{z}+\tfrac{1}{2}z)\right]\delta$$
$$B(\delta,z,Q)=-zA(\delta,z,Q)$$
$$C(\delta,z,Q)=-2zA(\delta,z,Q)$$

Hence, it is shown that

$$f(\delta,P_0(w,z,Q),1,1,z,w,Q)$$
$$=A(\delta,z,Q)(w-z)-2+2\sqrt{2}\sqrt{2-A(\delta,z,Q)w+z)}-2\sqrt{1-2zA(\delta,z,Q)}$$

For second best solutions we have from (9) that

$$\delta<\frac{\tfrac{1}{2}z+\tfrac{1}{2}w}{Q+\tfrac{1}{2}q^2z^2-P_0(w,z,Q)-\tfrac{1}{2}p^2w^2}=2\frac{z}{2Q+z^2}=\delta_0(z,Q)$$

Hence

$A(\delta,z,Q)=\left[1-\delta(Q\tfrac{1}{z}+\tfrac{1}{2}z)\right]\delta=(1-\delta\tfrac{1}{\delta_0})\delta\in(0,A_m)$,

where $A_m=\frac{1}{4\delta_0}=\frac{1}{8z}(2Q+z^2)$

Must also have expressions inside roots nonnegative

$2-A(\delta,z,Q)(w+z)\geq 0$ i.e. $w+z\leq\frac{2}{A(\delta,z,Q)}$,

and

$1-2zA(\delta,z,Q)\geq 0$ i.e. $A(\delta,z,Q)\leq\frac{1}{2z}$

Hence, we must have

$A_m\leq\frac{1}{2z}$ i.e. $\frac{1}{4}(2Q+z^2)\leq 1$

Note that

$$\frac{\partial}{\partial w}\left[A(w-z)-2+2\sqrt{2}\sqrt{2-A(w+z)}\right]=-A\frac{-\sqrt{(2-Aw-Az)}+\sqrt{2}}{\sqrt{(2-Aw-Az)}}<0$$

Hence, the expression is minimal when w is maximal, i.e. for

$$w=\min\left\{z,\tfrac{2}{A(\delta,z\,Q)}-z\right\}=z$$

where last equality follows because $\frac{1}{A}\geq 2z$. This yields

$$f(\delta,P_0(w,z,Q),1,1,z,w,Q)\geq 2(-1+\sqrt{2}\sqrt{1+a}-\sqrt{a})\geq 0$$

where $a=1-2zA(\delta,z,Q)\in(0,1)$, and the last inequality follows because expression is decreasing in a on (0,1).

It remains to consider

$$\tfrac{\partial}{\partial P}f(\delta,P,1,1,z,w,Q)=-2\delta^2\frac{-\sqrt{(\delta^2z^2-\delta^2w^2+2+2\delta^2Q-2\delta^2P-\delta w-\delta z)}+\sqrt{2}}{\sqrt{(\delta^2z^2-\delta^2w^2+2+2\delta^2Q-2\delta^2P-\delta w-\delta z)}}$$

Expression inside root is

$$(\delta^2z^2-\delta^2w^2+2+2\delta^2Q-2\delta^2P-\delta w-\delta z)=2+\delta\left[\delta(z^2-w^2)+2\delta(Q-P)-(w+z)\right]$$

We must have

$$\delta<\frac{\frac{1}{2}z+\frac{1}{2}w}{Q+\frac{1}{2}q^2z^2-P-\frac{1}{2}p^2w^2}=\frac{\frac{1}{2}z+\frac{1}{2}w}{Q+\frac{1}{2}z^2-P-\frac{1}{2}w^2}$$

i.e.

$$\delta(Q+\tfrac{1}{2}z^2-P-\tfrac{1}{2}w^2)<\tfrac{1}{2}z+\tfrac{1}{2}w$$

It follows that expression inside root above is < 2, and hence that $\frac{\partial}{\partial P}f<0$.

Thus we have shown

$f(\delta,P,1,1,z,w,Q)\geq f(\delta,P_0,1,1,z,w,Q)\geq 0$

for $P\leq P_0$.

REFERENCES

Anderson, Shannon W., David Glenn and Karen L. Sedatole. 2000. "Sourcing Parts of Complex products: Evidence on Transactions Costs, High-Powered Incentives and Ex Post Opportunism." *Accounting, Organizations and Society,* 25: 723-49.

Baker, George, Robert Gibbons and Kevin J. Murphy. 2002. "Relational Contracts and the Theory of the Firm." *Quarterly Journal of Economics* 117: 39-94.

Cassadeus-Masanell, Ramon and Daniel F. Spulber. 2000. "The Fable of Fisher Body." *The Journal of Law and Economics* 43: 67-104.

Chiu, Y. Stephen. 1998. "Noncooperative Bargaining, Hostages and Optimal Asset Ownership." *American Economic Review* 88: 882-901.

Coase, Ronald H. 1937. "The Nature of the Firm." *Economica* 4: 386-405.

Coase, Ronald H. 1988. "The Nature of the Firm: Origin, Meaning, Influence." *Journal of Law, Economics and Organization* 4: 3-47.

Coase, Ronald H. 2000. "The Acquisition of Fisher Body by General Motors." *Journal of Law and Economics* 37: 15-31.

Freeland, Robert F. 2000. "Creating Hold-up through Vertical Integration: Fisher Body Revisited." *Journal of Law and Economics* 37: 33-66.

Gibbons, Robert. 1992. *A primer in game theory.* London: Harvester/Wheatsheaf.

Grossman, Sanford and Oliver Hart. 1986 "The Costs and Benefits of Ownership: A theory of Lateral and Vertical Integration." *Journal of Political Economy* 94: 691-719.

Halonen, Maija. 2001. "Reputation and Allocation of Ownership." *The Economic Journal* 112: 539-58.

Hart, Oliver and John Moore. 1990. "Property Rights and the Nature of the Firm." *Journal of Political Economy* 98: 1119 – 58.

Hart, Oliver. 1995. *Firm, Contracts and Financial Structure.* Oxford: Clarendon Press.

Holmström, Bengt and John Roberts. 1999. "The Boundaries of the Firm Revisited." *Journal of Economic Perspectives* 12: 73-94.

Joskow, Paul L. 1988. "Asset Specificity and the Structure of Vertical Relationships: Empirical Evidence." *Journal of Law, Economics, & Organization 4:* 95-117

Klein, Benjamin, Crawford Robert G. and Alchian, Armen A. 1978. "Vertical Integration, Appropriable Rents, and the Competitive contracting Process." *Journal of Law and Economics* 21: 297-326.

Klein, Benjamin. 1980. "Transaction Cost Determinants of Unfair Contractual Arrangements." *American Economic Review* 70: 356-62.

Klein, Benjamin and Keith Leffler. 1981. " The Role of Market Forces in Assuring Contractual Performance." *Journal of Political Economy* 89:615-41.

Klein, Benjamin. 2000. "Fisher-General Motors and the Nature of the Firm." *The Journal of Law and Economics* 37: 105-41.

MacLeod, Bentley and James Malcomson. 1989. "Implicit contracts, Incentive Compatibility, and Involuntary Unemployment." *Econometrica* 57: 447-80

Milgrom, Paul and John Roberts. 1992. *Economics, Organization and Management.* New Jersey: Prentice Hall International, Inc.

Myerson, Roger B. 1997. *Game Theory: Analysis of Conflict.* Harvard University Press

Norsok Reports. 1995. Information on Norsok. See www.olf.no or www.oed.no .

Olsen, Trond. 1996. "Agency Costs and the Limit of Integration", *Rand Journal of Economics* 27: 479-501

Osmundsen, Petter. 1999. "Cost Overruns from the Perspective of Contract and Incentive Theory." Attachment to *Norwegian Government Report 11:* 113-31.

Shelanski, Howard A. and Peter G. Klein. 1995. "Empirical Research in Transaction Cost Economics: A Review and Assessment." *The Journal of Law Economics, & Organization* 11: 335-61.

Selten, R. 1965. "Spieltheoretisce Behandlung eines Oligopol-modells mit Nachfragetrageheit" *Zeitschrift für Gesamte Staatswissenschaft* 121:301-24.

Williamson, Oliver E., 1983. "Credible Commitments: Using Hostages to Support Exchange."
American Economic Review 73: 519-540.

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism*. New York: The Free Press.

Williamson, Oliver E. 1991. "Comparative Economic Organization: The Analyses of Discrete Structural Alternatives." *Administrative Science Quarterly* 36: 269-96.

Essay

II

Human Capital and Risk Aversion in Relational Incentive Contracts

Ola Kvaløy*

Norwegian School of Economics and Business Administration

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, Petter Osmundsen and especially Trond E. Olsen for valuable comments and suggestions.

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

[1] See for instance Becker (1975), Williamson (1975), Cheung (1982), Teece (1986), Mailath and Postelwhite (1990), Liebskind (1996).

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 gives the conditions for when increased specificity enhance social surplus.

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 α 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*. Since the noise term, $x \in (x_L, x_H)$, does not fulfil the requirements of a normally distributed random variable, this linear incentive scheme cannot be optimal. Holmström and Milgrom (1987) showed that normally distributed noise terms are necessary for linear incentive contracts to be optimal. Prior to Holmström and Milgrom's seminal paper, the linear incentive contract was regarded second best. Mirlees (1974) showed that the best linear contract is inferior to various non-linear incentive contracts. In particular, a step function contract, where the agent earns w_H if $Y \geq \bar{Y}$, but w_L if $Y < \bar{Y}$, can approach both full incentives and full insurance. As w_L and $\bar{Y}$ approaches zero, the agent almost surely receives w_H, and yet have incentives from fear of w_L. Holmström and Milgrom 'rescued' the linear contracts by envision a sequence of actions/effort decisions influencing a corresponding sequence of outcomes, rather than a single effort decision influencing a single outcome (for instance one action and one outcome per day, over the course of a year). There are no connections across days and all past outcomes are observed before the next day's effort is chosen. Under certain assumptions, including exponential utility and normally distributed noise, it is optimal to repeat the same one-day contract every day, regardless of history. If the one-day output is binary (i.e. just two possible outcomes each day), then the aggregate wage payment for the year is a linear function of the aggregate output. With infinite number of periods within a year, the optimal scheme is linear in output if the worker can influence the output within these periods. Technically the worker must then control the drift of a Brownian motion.

But even if the model in this paper do not satisfy the normally distributed noise assumption, 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 (by imposing quite strong assumptions) but rather that its sequential-action model implicitly alerts us to gaming as a natural consequence of non-linear contracts." For example, the Mirrlees step contract would induce no effort once the worker's aggregate output to date passes $\bar{Y}$. More generally, if the incentive contract for the year is a non-linear function of the year's output, then the worker's incentives change from day to day, depending on the aggregate output to date. Linear incentive contracts have the advantage of preventing these kinds of dynamic moral hazard problems, or 'gaming 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

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

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 (see appendix).

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

$$TCE = e - C(e) - \tfrac{1}{2} r\beta^2 V.$$

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:

(1) $\quad \beta = \frac{\partial C}{\partial e}$

The employer now maximizes the total certainty equivalent by choice of β, subject to the incentive constraint. That is

$$\underset{\beta}{Max} \ (e - C(e) - \tfrac{1}{2} r\beta^2 V)$$

subject to (1)

Solving this for β yields

(2) $$\hat{\beta}=\frac{1}{1+rVC''},$$

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 (α,β), where α is a fixed salary to be paid ex ante the production of Y, and βY 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 (β) 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 γY. 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 (α,β). 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 γY, hereafter called a spot contract, forever after.[3] (Note that even if we now enter into the 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

$$(3) \qquad \beta(e^R+x)+\tfrac{\delta}{1-\delta}CE_w{}^R,$$

where superscript, R, denotes relational contract, δ denotes the discount factor and e^R maximizes the certain equivalent such that $CE_w{}^R=\underset{e}{Max}(\alpha+\beta e-C(e)-\tfrac{1}{2}r\beta^2V)$. If 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.

worker reneges on the contract, and calls for a renegotiation, the present value of his expected profit is given by

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

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

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

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

(6) $(1-\beta)(e^R + x) + \frac{\delta}{1-\delta} CE_e^R$,

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) $(1-\gamma)(e^R + x) + \frac{\delta}{1-\delta} CE_e{}^S$

where $CE_e{}^S = e^S - \gamma e^S$.

The employer will stick to original compensation package if

(8) $(1-\beta)(e^R + x) + \frac{\delta}{1-\delta} CE_e^R \quad \geq \quad (1-\gamma)(e^R + x) + \frac{\delta}{1-\delta} CE_e^S \quad \forall x$.

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

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

where $\Delta x = x_H - x_L$.

That is

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

The parties can choose the fixed salary, α, 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 β. 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* $\hat{\beta} \in (\beta_L, \beta_H)$, β_L *iff* $\hat{\beta} < \beta_L \leq \gamma$ *and* β_H *iff* $\hat{\beta} > \beta_H \geq \gamma$,

where $\beta_L | \beta_L \geq \hat{\beta}$ *is given by*

$$(10) \quad v(\gamma - \beta_L) = TCE^R - TCE^S, \qquad v = \Delta x \tfrac{1-\delta}{\delta}.$$

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

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

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* $\hat{\beta} \notin (\beta_L, \beta_H)$.

It is naturally most interesting to study the properties of relational incentive contracts when $\hat{\beta} \notin (\beta_L, \beta_H)$.[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

[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.

meet the worker's ex post outside opportunities. In order to earn profit he then has to reduce the fixed salary α. 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 $e(\beta, k)$ given by $\beta = \frac{\partial C}{\partial e}(e,k)$, we have $\frac{\partial^2 e}{\partial k \partial \beta} > 0$. That is, the incentive responsiveness $\frac{\partial e}{\partial \beta}$ increases with increasing k (see appendix for more details). We obtain

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* $\frac{\partial \hat{\beta}}{\partial r} = \frac{\partial \hat{\beta}}{\partial V} < 0$, $\frac{\partial \hat{\beta}}{\partial k} > 0$ *and* $\frac{\partial \beta_i}{\partial r} = \frac{\partial \beta_i}{\partial V} < 0$, $\frac{\partial \beta_i}{\partial k} > 0$, $i = H, L$.

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 $\hat{\beta} \notin (\beta_L, \beta_H)$, and contrary:

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

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 $\hat{\beta} < \beta_L \leq \gamma$ 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 γ. They can reduce γ by relationship specific investments, for instance in firm specific training programs. And they can increase γ by standardizing output or generalizing the skill of the worker. Of course, the parties must balance the gains from adjusting γ 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 β is in the region where the curved line lies above the chord, that is between β_L and $\gamma = \beta_H$ on the horizontal axis, where β_L is decided by the parameters.

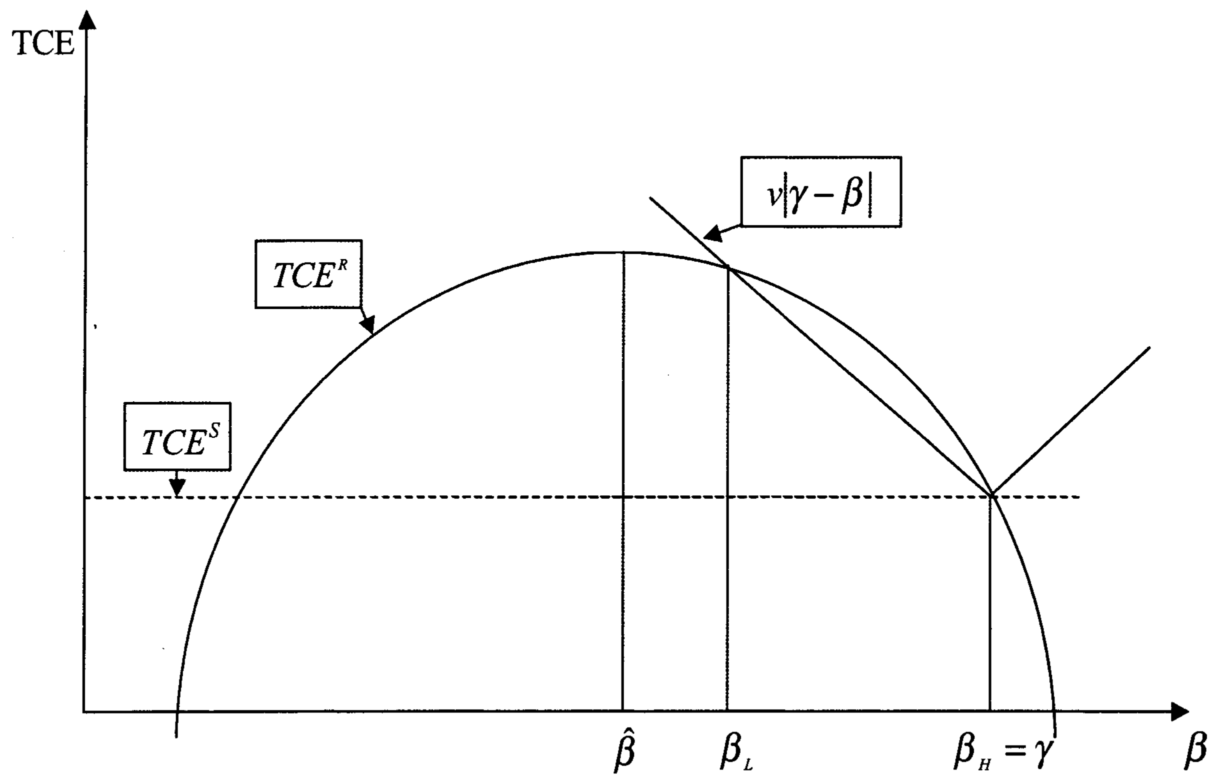


Figure 1

From figure 1 we see if $\hat{\beta} < \beta_L < \gamma$, a marginal increase in γ would reduce β_L (since the gradient of the chord is unaffected) and thus increase social surplus. An increase in γ increases the worker's short-term gain from deviating (given positive realizations of output Y), but it also a 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 γ, and the parties can only increase social surplus by lowering γ. 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:

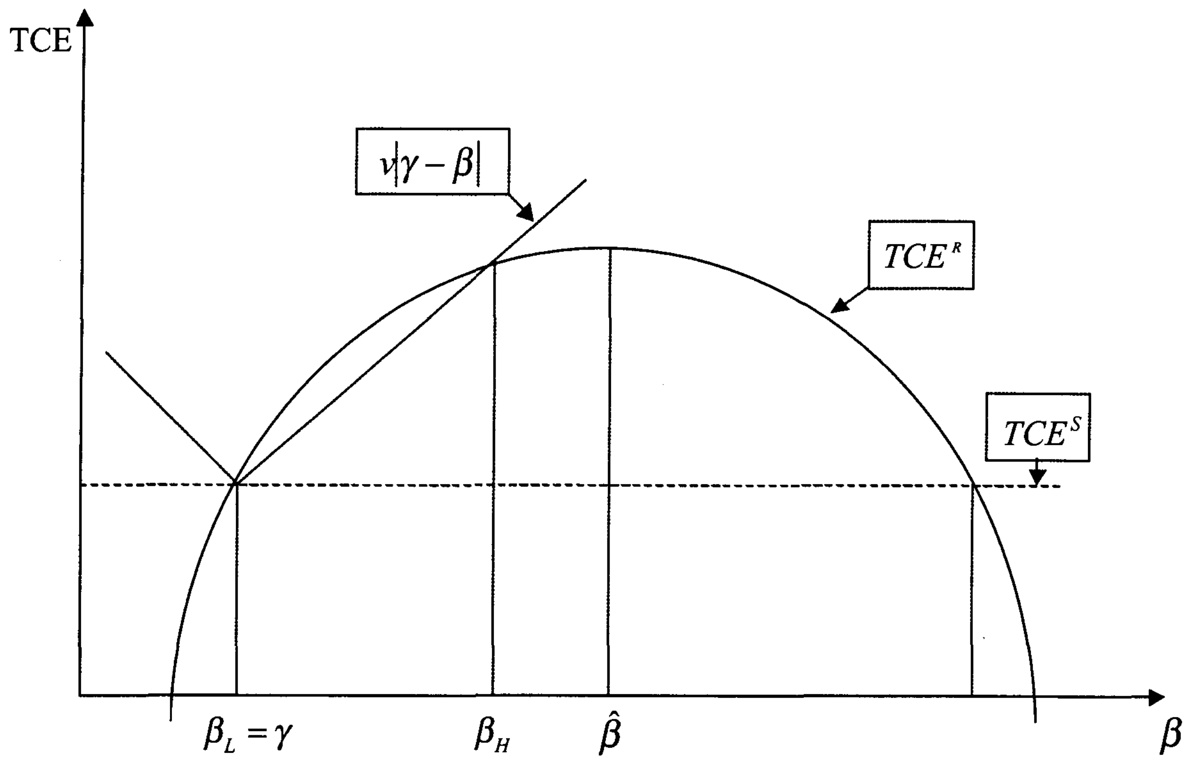


Figure 2

Figure 2 shows (9) when $\gamma < \beta_H < \hat{\beta}$. Now we see that a marginal *decrease* in γ would increase β_H, and thus increase social surplus. Here, an increase in γ 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 γ. 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 (see Holmström, Roberts, 1998), and lead to more efficient incentive schemes (Kvaløy, 2003). 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* $\frac{\partial TCE}{\partial \gamma} > 0$ *and* $\delta < \bar{\delta}$ *so that* $\frac{\partial TCE}{\partial \gamma} < 0$. *If* $\gamma < \beta_H < \hat{\beta}$, *then there exist a discount factor* $\delta > \tilde{\delta}$ *so that* $\frac{\partial TCE}{\partial \gamma} < 0$, *and* $\delta < \tilde{\delta}$ *so that* $\frac{\partial TCE}{\partial \gamma} > 0$.

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.[5] 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.[6] 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 $\hat{\beta} = \gamma = \beta^S$, then there is no need for a relational contract to implement it. Hence, the parties will not form a firm. If $\hat{\beta} \neq \beta^S$, then there exists a gain from engaging in a relational contract. If $\hat{\beta}$ is close to β^S, 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:

Proposition 5: *Let Φ 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

[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.

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 the worker's certainty equivalent $CE_w = \alpha + \beta e - C(e) - \frac{1}{2} r \beta^2 V$

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 $\overline{w} = E[w]$. Let us first leave out personal cost $C(e)$. The certainty equivalent is then approximately

$$\overline{w} - \tfrac{1}{2} r(\overline{w}) Var(w) = \hat{w},$$

where $r(\overline{w}) = -u''(\overline{w}) / u'(\overline{w})$

Derivation (from Milgrom and Roberts, 1992):

According to Taylor's theorem, for any z we have

$$u(z) = u(\overline{w}) + (z - \overline{w}) u'(\overline{w}) + \tfrac{1}{2}(z - \overline{w})^2 u''(\overline{w}) + R(z),$$

where $R(z) = u'''(\hat{z})(z - \overline{w})^3 / 6$ for some $\hat{z} \in [\overline{w}, z]$. This last term is assumed to be small and thus negligible. Hence, we can write approximately

$$u(z) \approx u(\overline{w}) + (z - \overline{w}) u'(\overline{w}) + \tfrac{1}{2}(z - \overline{w})^2 u''(\overline{w}).$$

Substituting w for z and computing the expectation, we find, approximately

$$E[u(w)] \approx u(\overline{w}) + E[w-\overline{w}]u'(\overline{w}) + \tfrac{1}{2}E\left[(w-\overline{w})^2\right]u''(\overline{w}).$$

Since $E[w-\overline{w}] = E[w] - \overline{w} = \overline{w} - \overline{w} = 0$, we can write

$$\text{(A.1)} \quad E[u(w)] \approx u(\overline{w}) + \tfrac{1}{2}E\left[(w-\overline{w})^2\right]u''(\overline{w}).$$

The certainty equivalent $\hat{w}$ is expected to be close to $\overline{w}$, so its utility is approximated differently, also using Taylor's theorem,

$$\text{(A.2)} \quad u(\hat{w}) = u(\overline{w}) + (\hat{w} - \overline{w})u'(\overline{w}) + \dot{R}(\hat{w})$$

where $\dot{R}(\hat{w}) = \frac{1}{2}u''(\hat{z})(\hat{w} - \overline{w})^2$ for some $\hat{z} \in [\overline{w}, \hat{w}]$. If we apply the approximation only when $\hat{w} - \overline{w}$ is small, the remainder term is again negligible. Since $\hat{w}$ is a certainty equivalent, we have $u(\hat{w}) = E[u(w)]$. So combining (A.1) and (A.2) yields

$$(\hat{w} - \overline{w})u'(\overline{w}) \approx \tfrac{1}{2}E\left[(w-\overline{w})^2\right]u''(\overline{w}).$$

This can be expressed in the form

$$\hat{w} - \overline{w} \approx \tfrac{1}{2}\left[u''(\overline{w})/u'(\overline{w})\right] E\left[(w-\overline{w})^2\right] = -\tfrac{1}{2}r(\overline{w})Var(w),$$

which establishes $\hat{w} = \overline{w} - \frac{1}{2}r(\overline{w})Var(w)$.

Subtracting personal cost, and insert for $\overline{w} = E[w] = \alpha + \beta e$, we obtain the worker's certainty equivalent

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

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

(A.3) $\beta(e^R + x_H) + \tfrac{\delta}{1-\delta} CE_w{}^R \quad \ge \quad \gamma(e^R + x_H) + \tfrac{\delta}{1-\delta} CE_w{}^S$

(A.4) $(1-\beta)(e^R + x_L) + \tfrac{\delta}{1-\delta} CE_e^R \quad \ge \quad (1-\gamma)(e^R + x_L) + \tfrac{\delta}{1-\delta} CE_e{}^S$

A necessary condition for the relational contract to hold is that the sum of (A.3) and (A.4) holds. This yields

(A.5) $(\gamma - \beta)(x_H - x_L) \le \tfrac{\delta}{1-\delta}(TCE^R - TCE^S)$.

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

(A.6) $\beta(e^R + x_L) + \tfrac{\delta}{1-\delta} CE_w{}^R \quad \ge \quad \gamma(e^R + x_L) + \tfrac{\delta}{1-\delta} CE_w{}^S$

(A.7) $(1-\beta)(e^R + x_H) + \tfrac{\delta}{1-\delta} CE_e^R \quad \ge \quad (1-\gamma)(e^R + x_H) + \tfrac{\delta}{1-\delta} CE_e{}^S$,

and the sum of (A.6) and (A.7) yields

(A.8) $(\beta - \gamma)(x_H - x_L) \le \tfrac{\delta}{1-\delta}(TCE^R - TCE^S)$.

Since (A.5) is relevant for $\beta \le \gamma$ and (A.8) is relevant for $\beta \ge \gamma$ we can write these to restrictions in one expression using absolutes:

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

where $\Delta x = x_H - x_L$. As noted, the parties can choose the fixed salary, α, to make the condition sufficient.

3. The measure of incentive responsiveness

For $e(\beta, k)$ given by $\beta = \frac{\partial C}{\partial e}(e,k)$, we have $\frac{\partial^2 e}{\partial k \partial \beta} > 0$. That is, the incentive responsiveness $\frac{\partial e}{\partial \beta}$ increases with increasing k. This holds if the cost function satisfies $\frac{\partial^2 C}{\partial e \partial k}\frac{\partial^3 C}{\partial e^3} - \frac{\partial^2 C}{\partial e^2}\frac{\partial^3 C}{\partial e^2 \partial k} > 0$. (Example, the condition holds for a cost function of the form $C(e,k) = A(k)e^n, n \ge 2$, where $A'(k) < 0$). With this condition, the gain from a marginal increase in β increases with the level of incentive responsiveness. That is $\frac{\partial TCE}{\partial k \partial \beta} = (1-\beta)\frac{\partial^2 e}{\partial k \partial \beta} > 0$ for $\beta < 1$.

4. Proof proposition 2 and 3

When $\beta \in (\beta_L, \beta_H)$ the optimal β 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 β_L is optimal iff $\hat{\beta} < \beta_L \le \gamma$

where β_L is given by

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

. When $\hat{\beta} < \beta_L < \gamma$ we must have (for simplicity I exclude functional arguments):

(A.9) $-\frac{\partial TCE^R}{\partial \beta}\Big|_{\beta=\beta_L} < v$

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

Differentiating (10) with respect to r yields

(A.10) $(-v - \frac{\partial TCE^R}{\partial \beta})\frac{\partial \beta_L}{\partial r} = \frac{\partial TCE^R}{\partial r} - \frac{\partial TCE^S}{\partial r}$

From (A.9), the bracket on the left hand side is negative, and the difference on the right hand side is positive since $\beta_L < \gamma$ and $\frac{\partial^2 TCE}{\partial r \partial \beta} < 0$. This yields $\frac{\partial \beta_L}{\partial r} < 0$, which also implies $\frac{\partial \beta_L}{\partial V} < 0$.

Differentiating (10) with respect to k yields

(A.11) $(-v - \frac{\partial TCE^R}{\partial \beta})\frac{\partial \beta_L}{\partial k} = \frac{\partial TCE^R}{\partial k} - \frac{\partial TCE^S}{\partial k}$

From (A.9), 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 $\frac{\partial^2 TCE}{\partial k \partial \beta} > 0$. This yields $\frac{\partial \beta_L}{\partial k} > 0$.

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

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

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

(A.12) $\frac{\partial TCE^R}{\partial \beta}\Big|_{\beta=\beta_H} < v$

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

Differentiating (11) with respect to r yields

(A.13) $(v - \frac{\partial TCE^R}{\partial \beta})\frac{\partial \beta_H}{\partial r} = \frac{\partial TCE^R}{\partial r} - \frac{\partial TCE^S}{\partial r}$

From (A.12), the bracket on the left hand side is positive, and the difference on the right hand side is negative since $\beta_H > \gamma$ and $\frac{\partial^2 TCE}{\partial r \partial \beta} < 0$. This yields $\frac{\partial \beta_H}{\partial r} < 0$, which also implies $\frac{\partial \beta_H}{\partial V} < 0$.

Differentiating (11) with respect to k yields

(A.14) $(v - \frac{\partial TCE^R}{\partial \beta})\frac{\partial \beta_H}{\partial k} = \frac{\partial TCE^R}{\partial k} - \frac{\partial TCE^S}{\partial k}$

From (A.12), 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 $\frac{\partial^2 TCE}{\partial k \partial \beta} > 0$. This yields $\frac{\partial \beta_H}{\partial k} > 0$.

Proposition 3 can be verified by differentiating (A.10), (A.11) (A.13) and (A.14) with respect to v noting that $\frac{\partial v}{\partial \delta} < 0$.

REFERENCES

Abreu, Dilip. 1988. " On the Theory of Infinitely Repeated Games with Discounting." *Econometrica* 56:383-96.

Asch, Beth. 1990. "Do Incentives Matter? The Case of Navy Recruiters." *Industrial and Labor Relations Review* 43:89-106.

Baker, George, Robert Gibbons and Kevin J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts." *Quarterly Journal of Economics* 109: 1125-56.

Baker, George, Robert Gibbons and Kevin J. Murphy. 2001. "Bringing the Market Inside the Firm." *American Economic Review Papers and Proceedings* 91: 212-18.

Baker, George, Robert Gibbons and Kevin J. Murphy. 2002. "Relational Contracts and the Theory of the Firm." *Quarterly Journal of Economics 117: 39-94.*

Becker, Gary S. 1975. *"Human Capital."* Chicago IL: University of Chicago Press.

Blair, Margaret, and Mark J. Roe. 1999. *"Employee and Corporate Governance."* The Brooking Institution.

Bull, Clive. 1987. "The existence of Self-Enforcing Implicit Contracts." *Quarterly Journal of Economics* 102: 147-59.

Cappelli, Peter. 2000. "Market-Mediated Employment: The Historical Context." In M. Blair and T.A. Kochan, eds: *The New Relationship. Human Capital in the American Corporation.* The Brooking Institution Press.

Cheung, Steven. 1982. "Property Rights in Trade Secrets." *Economic Inquiry* 20: 40-53

Coase, Ronald H. 1937. "The Nature of the Firm." *Economica* 4: 386-405.

Fehr, Ernst and Gacther, Simon. 2000. "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economics Perspective* 14: 159-81.

Fudenberg, D., B. Holmstrom and P. Milgrom, 1990, "Short-Term Contracts and Long Term Agency Relationships." *Journal of Economic Theory* 51: 1-31.

Fudenberg, Drew and Jean Tirole. 1991. *Game Theory*. Cambridge, MA: MIT Press.

Gibbons, Robert. 2002. "Incentives Between Firms (and Within)." Forthcoming *Management Science.*

Grossman, Sanford and Oliver Hart. 1986 "The Costs and Benefits of Ownership: A theory of Lateral and Vertical Integration." *Journal of Political Economy* 94: 691-719.

Hart, Oliver. 1995. *Firm, Contracts and Financial Structure.* Oxford: Clarendon Press.

Hart, Oliver. 2001. "Norms and the Theory of the Firm." *University of Pennsylvania Law Review* 149: 1701-15.

Healy, Paul. 1985. "The Effect of Bonus Schemes on Accounting Decisions." *Journal of Accounting and Economics* 7: 85-107.

Holmström, Bengt and Paul Milgrom. 1987. "Aggregation and Linearity in the Provision of Intertemporal Incentives." *Econometrica* 55: 303-28.

Holmström, Bengt and Paul Milgrom. 1991. Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership and Job Design." *Journal of Law, Economics and Organization* 7: 24-52.

Holmström, Bengt and John Roberts. 1999. "The Boundaries of the Firm Revisited." *Journal of Economic Perspectives* 12: 73-94.

Klein, Benjamin, Crawford Robert G. and Alchian, Armen A. 1978. "Vertical Integration, Appropriable Rents, and the Competitive contracting Process." *Journal of Law and Economics* 21: 297-326.

Klein, Benjamin and Keith Leffler. 1981. " The Role of Market Forces in Assuring Contractual Performance. " *Journal of Political Economy* 89:615-41.

Kreps, David M., 1990. "Corporate Culture and Economic Theory." In J. Alt and K. Shepsle, eds. *Perspectives on Positive Political Economy.* Cambridge University Press.

Kvaløy, Ola. 2003. "Asset Specificity and Vertical Integration. In *"Relational Incentive Contracts and Allocation of Ownership Rights."* Submitted for the degree of doctor oecon. Norwegian School of Economics and Business Administration.

Levin, Jonathan. 2003. "Relational Incentive Contracts." American Economic Review 93: 835-57.

Liebskind, Julia P. 1996. "Knowledge, Strategy and the Theory of the Firm." *Strategic Management Journal* 17: 93-107.

MacLeod, Bentley and James Malcomson. 1989. "Implicit contracts, Incentive Compatibility, and Involuntary Unemployment" *Econometrica* 57: 447-80.

Mailath, George J., and Andrew Postlewaite. 1990. "Workers versus Firms: Bargaining over a Firm's Value." *Review of Economic Studies* 57: 369-80.

Milgrom, Paul and John Roberts. 1992. *Economics, Organisation and Management.* New Jersey: Prentice Hall International, Inc.

Mirrlees, James. 1974. "Notes on Welfare Economics, Information, and Uncertainty." In M. Balch, D. McFadden, and S. Wu (eds.) *Essays on Economic Behaviour Under Uncertainty*, Amsterdam: North-Holland.

O'Connor, Marleen A. 1993. "The Human Capital Era: Reconceptualizing Corporate Law to Facilitate Labor- Management Cooperation " *Cornell Law Review* 78:899-965.

Oyer, Paul. 1998. "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality." *Quarterly Journal of Economics* 113:149-85.

Radner, R., 1981. "Monitoring Cooperative Agreements in a Repeated Principal-Agent Relationship." *Econometrica* 49: 1127-48.

Rajan, Raghuram R., and Luigi Zingales. 2001. "The Firm as a Dedicated Hierarchy."

Quarterly Journal of Economics 116: 805-51.

Rebitzer, John B., and L.J Taylor. 1997. "When knowledge is an asset: Explaining the Organizational Structure of Large Law Firms." *MIT Working Paper.*

Rogerson, W., 1985. "Repeated Moral Hazard." *Econometrica* 53: 69-76.

Rubinstein, Ariel.1982. Perfect Equilibrium in a Bargaining Model." *Econometrica* 50: 97-109.

Shapiro, Carl and Joseph E. Stiglitz. 1984. "Equilibrium Unemployment as a Worker Discipline Device." *American Economic Review* 74:433-44.

Simon, Herbert. 1951. "A formal theory of the employment relationship." *Econometrica* 19: 293:305.

Teece, David. 1986. "Profiting from Technological Innovation" *Research Policy* 15: 286-305.

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism*. New York: The Free Press.

Essay

III

Team Incentives in Relational Contracts

Ola Kvaløy*

Norwegian School of Economics and Business Administration

Abstract: *Incentive schemes for teams are compared. I ask: under which conditions are relational incentive contracts based on joint performance evaluation, relative performance evaluation and independent performance evaluation self-enforceable. The framework of Che and Yoo (2001) on team incentives is combined with the framework of Baker, Gibbons and Murphy (2002) on relational contracts. In a repeated game between one principal and two agents, I find that incentives based on relative or independent performance are expected to dominate when the productivity of effort is high, while joint performance evaluation dominates when productivity is low. Incentives based on independent performance are more probable if the agents own critical assets or are able to collude.*

* I thank Frøystein Gjesdal, Petter Osmundsen, seminar participants at the Norwegian School of Economic and Business Administration, and especially Trond E. Olsen for valuable comments and suggestions.

1. Introduction

In the last decade we have seen a growth of group based incentive schemes such as profit sharing, gain sharing and employee ownership schemes to improve motivation and labour relations. It is recognized in private industry that by introducing group incentive compensation systems, it may be possible to induce workers to work both harder and more cooperatively, in a way that enhance their productivity (see Banker, Field, Schroeder and Sinha, 1996; Gerhardt, Minkoff and Olsen, 1995). Yet there has been paid little attention to group incentives in the economics literature. The main reason for this lack of theoretical interest is that it has been difficult to prove the efficiency of such schemes. Economists studying teams, beginning with Alchian and Demsetz (1972), have argued that group incentives are ineffective due to free-riding problems. If you award an equal bonus to each member of a team on the basis of the team's overall performance, it may give each member an incentive to shirk (free ride). It is not difficult to see that if the members of a team are interacting only once (for example in one project only) and each one knows that this is the one and only project of which they will work together, the free riding problem can easily occur. In such a static one period relationship an incentive scheme based on relative performance evaluation (RPE) is optimal (see Lazear and Rosen, 1981; Holmström, 1982; Mookherjee, 1984). An RPE scheme rewards team members that perform relatively better than their peers. That way an employer can make a worker's compensation independent from good or bad outside factors (common noise components). This lowers the cost of providing a given level of incentives. The RPE scheme also has the advantage that one only needs to detect relative performance, which can be easier than measuring absolute performance.

In recent years, however, economists have been able to show that joint performance evaluation (JPE) may come out as an optimal solution. A JPE scheme compensates the group members if the group as a whole performs well. Hence, a worker is rewarded if his peers perform well. The value of encouraging employees' cooperation is emphasized in works such as Holmström and Milgrom (1990), Itoh (1992, 1993), Macho-Stadler and Perez-Castrillo (1993) and particular Kandel and Lazear (1992) who emphasize the effect of peer pressure. In addition, the folk theorem of repeated games has for some years now provided a possible answer to the free rider critique of group incentives (see Radner, 1986; Weitzman and Kruse, 1990; and FitzRoy and Kraft, 1995). In a repeated setting where agents interact in an unknown number of periods, shirking from some desired, co-operative solution can be deterred by social sanctions including withholding co-operation in the future. This idea is most elegantly formalized by Che and Yoo (2001). They show how an implicit contract between the agents of a team can generate implicit incentives and thus make joint performance evaluation optimal.

But even if group incentives have gained popularity, we still see a high frequency of individual compensation schemes based on relative or independent performance evaluation (IPE). This is especially the case for white collar workers (Prendergast, 1999) and workers in higher levels of organizations (see Appelbaum and Berg, 1999). In the present paper I combine the framework of Che and Yoo on team incentives with the framework of Baker, Gibbons and Murphy (BGM) on relational contracts[1] to show when relative performance evaluation and independent performance evaluation (IPE) is expected to dominate in repeated transactions. Like Che and Yoo I study a repeated game between one principal and two agents, but contrary to them, I model a self-enforcing relational contract between the principal and his agents. Contrary to Che and Yoo, I assume that the quality of the agents' output is non-verifiable, so that legal enforcement is impossible. Like Che and Yoo, I compare three types of incentive schemes: Relative performance evaluation, joint performance evaluation and independent performance evaluation. But instead of focusing on optimality conditions, I focus on enforceability conditions: under which conditions are the various incentive schemes implementable? Since the parties only can choose incentive schemes among the enforceable ones, the results of my model often differs from Che and Yoo's results.

In Section 2 and 3 I analyse an environment where collusion is impossible, and where the principal owns the assets so that the agents do not have the possibility to hold-up any values ex post production. The main result from these sections is that when productivity of effort

[1] 'Relational' contracts and 'implicit' contracts are used synonymously in the literature. MacLeod and Malcomson (1989), Baker, Gibbons and Murphy (1994) and Schmidt and Schnitzer (1995) used 'implicit', while Bull (1987) used both 'implicit' and 'relational'. Newer papers such as Baker, Gibbons and Murphy (2002) and Levin (2003) use 'relational', inspired by the legal literature, particularly MacNeil (1978). Since implicit contracts can be interpreted as vaguer than relational contracts (due to the antonym implicit versus explicit), I will in this paper use the term 'implicit' on the contract between the agents (like Che and Yoo), since it is most natural to think about this contract as a verbal informal agreement. But I will use 'relational' on the contract between the principal and his agents, since this most likely is a formally written wage contract.

increases, the enforceability of IPE and RPE rises relatively to the enforceability of JPE. In general, an increase in productivity raises the cost of deviation, and the discount factor is allowed to decrease without running the risk of deviation. But JPE is vulnerable to low discount factors. Since the efficiency of JPE is dependent on the possibility for repeated interaction among the agents, the necessary JPE incentives increase with lower discount factors. Hence, higher productivity weakens the enforceability of JPE relatively to RPE and IPE.

In Section 4 I briefly discuss the implications of collusion. In particular, I show that independent performance evaluation may turn out as optimal if the agents are able to collude. In Section 5 I consider the situation when the agents own the critical assets. I show that the possibilities for agents to renegotiate the terms of trade ex post the realizations of values, makes incentives based on non-independent performance evaluation harder to enforce. This opens for independent performance evaluation to dominate.

The empirical relevance of the model is discussed in Section 6. Section 7 concludes.

2. The model

First I will replicate the repeated setting in Che and Yoo's model. Consider an economic environment consisting of one principal and two identical agents who each period produce either high, Q_H, or low, Q_L, values for the principal. Their effort level can be either high or low, where high effort has a disutility cost of c and low effort is costless. The principal can only observe the realization of the agents' output, not the level of effort they choose. But the agents can observe each other's effort decisions. Their output depends on effort decisions as well as a common environmental shock. A favourable shock occurs with probability $\sigma \in (0,1)$, in which case both agents produce high values for the principal. If the shock is unfavourable, the probability for agent i of realizing Q_H is q_H if the agent's effort is high and q_L if the agent's effort is low, where $q_H > q_L$.

Agent i receives a fixed payment, α, prior to (ex ante) value realizations,[2] where α can be both positive and negative. A bonus wage vector $\mathbf{ß}^i \equiv (\beta^i{}_{HH}, \beta^i{}_{HL}, \beta^i{}_{LH}, \beta^i{}_{LL})$ where the subscripts denote respectively agent i and agent j's realization of Q_i, $(i = H, L)$, is to be paid ex post the realizations of values. It assumed that all parties are risk neutral, except that the agents are subject to limited liability: the principal cannot impose negative bonus wages. Limited liability may arise fom liquidity constraints or from laws that prohibit firms from exacting payments from workers.[3]

Let agent i and j choose efforts $k \in \{H, L\}$and $l \in \{H, L\}$ respectively. Agent i's expected wage is then

$$\begin{aligned}(1)\quad & \pi(k,l,\mathbf{ß}^i) \equiv \alpha + (\sigma + (1-\sigma)q_k q_l)\beta^i{}_{HH} \\ & +(1-\sigma)\left[q_k(1-q_l)\beta^i{}_{HL} + (1-q_k)q_l\beta^i{}_{LH} + (1-q_k)(1-q_l)\beta^i{}_{LL}\right]\end{aligned}$$

For each agent, a wage scheme exhibits joint performance evaluation if $(\beta_{HH}, \beta_{LH}) > (\beta_{HL}, \beta_{LL})$,[4] (I suppress superscript since the agents are symmetric). In this case $\pi(k, H; \mathbf{ß}) > \pi(k, L; \mathbf{ß})$ so an agent's work yields positive externalities to his partner. A wage scheme exhibits relative performance evaluation if $(\beta_{HH}, \beta_{LH}) < (\beta_{HL}, \beta_{LL})$. In this case $\pi(k, H; \mathbf{ß}) < \pi(k, L; \mathbf{ß})$ so an agent's work generate a negative externality on his partner. A wage scheme exhibits independent performance evaluation if $(\beta_{HH}, \beta_{LH}) = (\beta_{HL}, \beta_{LL})$which implies $\pi(k, H; \mathbf{ß}) = \pi(k, L; \mathbf{ß})$, so an agent's work has no impact on his partner.

It is assumed that high effort is sufficiently valuable to the principal that he always prefers to induce the agents to exert high effort. The principal's problem is then to minimize the wage payments subject to the constraints that the agents must be induced to yield high effort. In a repeated setting, the agents can exploit the fact that they are able to observe each other's effort decisions. In particular, they can play a repeated game where they both play high effort if the

[2] Che and Yoo do not include fixed payments in their model.
[3] Limited liability in terms of liquidity constraints does not conflict with the possibility of a negative fixed payment since the fixed payment is paid ex ante. Also, a law against exacting payment ex post can still permit voluntary payments from workers ex ante.
[4] The inequality means weak inequality of each component and strict inequality for at least one component.

other agent played high effort in the previous period. In order for such a strategy to constitute a subgame perfect equilibrium, we must have

(2) $\frac{1}{1-\delta}(\pi(H,H;\text{ß})-c) \geq \pi(L,H;\text{ß})+\frac{\delta}{1-\delta}\min\{\pi(L,L;\text{ß}),\pi(L,H;\text{ß})\}$,

where δ is the discount factor. The left hand side shows the expected present value of playing high effort, while the right hand side shows the expected wage from unilaterally playing low effort in one period and being subsequently punished by the worst possible equilibrium payoff. Hence, (2) says that, given the strategy to play high effort if the other agent played high effort in the previous period, an agent will play high effort as long the present value from playing high effort is greater than the present value from playing low effort. Note that (2) is a necessary but not sufficient condition. For (2) to be sufficient, the punishment path specified on the right hand side must also be self-enforcing.

Observe that in a JPE scheme, $\pi(L,H;\text{ß}) > \pi(L,L;\text{ß})$. Thus the right hand side of (2) becomes $\pi(L,H;\text{ß})+\frac{\delta}{1-\delta}\pi(L,L;\text{ß})$. In an RPE scheme, however, $\pi(L,L;\text{ß}) > \pi(L,H;\text{ß})$, so the right hand side of (2) is reduced to $\frac{1}{1-\delta}\pi(L,H;\text{ß})$ which makes (2) coinciding with the static incentive constraint (see Che and Yoo). Hence, we see that repeated interaction between the agents can increase the punishment of playing low effort in a JPE scheme, but not in an RPE scheme. The intuition is straightforward: in the JPE scheme, low effort from agent *i* does not only imply a reduced chance for him to realize high values, but it also implies that his peer plays low effort and thus lower the chance of realizing high values. This is costly since a JPE scheme promises highest wage if *both* realize high values. Hence, the repeated interaction yields both direct and *implicit* incentives to yield high effort.

Now, for the principal to choose the most efficient wage scheme, he must solve

(3) $\min_{\text{ß}\geq 0} \pi(H,H;\text{ß})$, subject to (2).

This is a relaxed program since there also exists low-effort strategies that constitute subgame perfect equlibria.

The following lemma characterizes the solution to (3):

Lemma: *Define* $\hat{\delta}(\sigma) \equiv \frac{\sigma}{(1-\sigma)q_H q_L}$. *If* $\delta \in \left[\hat{\delta}(\sigma), 1\right]$, *then a JPE scheme* $\mathbf{ß}^{JPE} \equiv (\beta_{HH}^{JPE},\ 0,\ 0,\ 0)$ *where* $\beta_{HH}^{JPE} \equiv \dfrac{c}{(1-\sigma)(q_H + \delta q_L)\Delta q}$ *where* $\Delta q = q_H - q_L$ *solves (3). If* $\delta \in \left[0, \hat{\delta}(\sigma)\right]$, *then the RPE scheme* $\mathbf{ß}^{RPE} \equiv (0, \beta_{HL}^{RPE}, 0, 0)$ *where* $\beta_{HL}^{RPE} \equiv \dfrac{c}{(1-\sigma)(1-q_H)\Delta q}$ *solves (3).*

Proof: *See appendix.*

The lemma suggests that an extreme form of JPE is optimal for sufficiently high discount factors, while (under the no collusion assumption) an extreme form of RPE is optimal on sufficiently low discount factors. Intuitively, there must be a sufficiently high discount factor if an agent should have interests in assuring future high-effort from his peer, hence JPE is only optimal on high discount factors.[5]

It can be shown that $\mathbf{ß}^{JPE}$ makes the worst sustainable punishment -low effort from both workers (L,L)- self-enforcing. This makes high effort from both agents (H,H) a subgame perfect equilibrium (Che and Yoo call this a 'team equilibrium'). Hence, the incentive constraint given by (2) is sufficient when $\mathbf{ß}^{JPE}$ solves (3). When $\mathbf{ß}^{RPE}$ solves (3), (2) is sufficient if the discount factors are sufficiently high (see Che and Yoo for details).

The JPE scheme, $\mathbf{ß}^{JPE}$, in contrast to $\mathbf{ß}^{RPE}$, has the virtue of being collusion proof since each agent's work confers positive externalities to their peer, but $\mathbf{ß}^{RPE}$ is susceptible to collusion since both agents can jointly be better off by playing low effort. There may, however, exist institutional constraints to the possibilities of engaging in collusion. Of course, there are no technical constraints to collusion since the agents can observe each other's actions. But it is not unrealistic to assume that in an industrious corporate culture it is easier for workers to

[5] Not surprisingly then, the RPE scheme is optimal in the static version. In this sense Che and Yoo complements Holmström (1982) and Mookheerjee (1984). The optimality depends on the assumed specification of the common shock.

sustain a high effort culture than to initiate a low effort culture. A worker that initiates low effort collusion may risk great personal costs in loss of prestige and respect from his peers. Hence, before we proceed to the collusion problem, I will simply assume that there is a social cost associated with initiating low effort collusion that exceeds the benefits from such collusion.

2.1 Relational contract between principal and agents

Unlike Che and Yoo I will now assume that the value realizations are not verifiable to a third party. Hence, the contract between the agents and the principal must therefore be self-enforcing, and thus 'relational' by definition. I consider a *multilateral* relational contract, which implies that any deviation by the principal triggers low effort from both agents. The principal honours the contract only if *both* agents honoured the contract in the previous period. The agents honour the contract only if the principal honoured the contract with *both* agents in the previous period. A natural explanation for this multilateral feature is that the agents interpret a unilateral contact breach (i.e. the principal deviates from the contract with only one the agents) as evidence that the principal is not trustworthy (see Bewley, 1999).

The contract is self-enforcing if the present value of honouring is greater than the present value of reneging. Ex post realizations of values, the principal can renege on the contract by refusing to pay the promised wage, while the agents can renege by refusing to accept the promised wage. In this section I consider a *relational employment contract*, to use the terms of BGM. This means that the principal owns the critical assets. Asset ownership conveys ownership to the values produced, hence, the principal can ex post take the values even if he is not paying bonus wages.

The parties play trigger strategies. Like BGM, I assume that if one of the parties renege on the relational employment contract, the other insist on *spot employment* forever after. Spot employment implies that the agents exert low effort q_L, but receives zero wage (neither bonus wage nor fixed payment).

The condition for the optimal JPE contract, β^{JPE}, to be self-enforcing is (see appendix)

(4) $2\beta_{HH}^{JPE} \leq \frac{2\delta}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$,

where $\Delta Q = Q_H - Q_L$. The condition for the optimal RPE contract, $\mathbf{B}^{RPE}$, to be self-enforcing is (see appendix)

(5) $\beta_{HL}^{RPE} \leq \frac{2\delta}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$.

The expression in the square brackets shows the productivity of an agent's effort. Hence, the right hand side of (4) and (5) shows the present value of *both* agents yielding high effort. We clearly see the difference between the self-enforcing conditions of $\mathbf{B}^{JPE}$ and $\mathbf{B}^{RPE}$. While the principal 'risks' paying both agents in the JPE scheme, he only risks paying one if his agents in the RPE scheme. Levin (2002) argues that multilateral relational contracts makes relative performance evaluation favourable, since the principal only has to satisfy the sum of individual constraints (not every separate incentive constraint) and is thus committed to reward only one of his agents. But Levin does not allow for implicit contracting between the agents. In the present model, this possibility can make the necessary JPE wage, β_{HH}^{JPE}, lower-powered and thus easier to implement. Hence, there is a trade-off between enforcing a double set of small JPE wages and a single, but larger RPE wage.

3. Comparative analysis

We can formally compare the self-enforcing conditions of $\mathbf{B}^{JPE}$ and $\mathbf{B}^{RPE}$. Solving (4) for δ and inserting for β_{HH}^{JPE} yields

(6) $\delta \geq \frac{1}{2Aq_L}\left[-1 - Aq_L + \sqrt{(1 + 2Aq_H + A^2 q_H^2 + 4Aq_L)}\right] = \delta^{JPE}$,

where $A = \left[\frac{(1-\sigma)\Delta q \Delta Q}{c} - 1\right](1-\sigma)\Delta q$.

Solving (5) for δ and inserting for β_{HL}^{RPE} yields

(7) $\delta \geq \frac{1}{1+2A(1-q_H)} = \delta^{RPE}$.

Expression (6) and (7) show the critical discount factors for the optimal JPE contract and the optimal RPE contract, respectively, to be self-enforcing. If $\delta^{RPE} < \hat{\delta}(\sigma) < \delta < \delta^{JPE}$, then $\mathbf{B}^{JPE}$ is optimal, but not enforceable, while $\mathbf{B}^{RPE}$ is enforceable. The principal must then either choose a second best JPE contract, or $\mathbf{B}^{RPE}$ (a second best JPE contract implies a less extreme JPE contract where agent i can be paid even if (H,H) is not realized, that is $\beta_{LH}^{JPE} > \beta_{LL}^{JPE}$ and/or $\beta_{HH}^{JPE} \geq \beta_{HL}^{JPE} > 0$ (strict inequality if $\beta_{LH}^{JPE} = \beta_{LL}^{JPE}$), where $\beta_{HH}^{JPE} < \frac{c}{(1-\sigma)(q_H+\delta q_L)\Delta q}$). Since a second best JPE contract is more costly than the optimal JPE contract, $\mathbf{B}^{JPE}$, this will increase the critical discount factor, $\tilde{\delta}$, for when an incentive scheme that exhibits JPE is chosen. Hence, if $\delta^{RPE} < \hat{\delta}(\sigma) < \delta^{JPE}$, then there exist discount factors $\hat{\delta}(\sigma) < \delta < \tilde{\delta}$, where $\mathbf{B}^{RPE}$ is chosen.

If $\delta^{JPE} < \delta < \hat{\delta}(\sigma) < \delta^{RPE}$, then $\mathbf{B}^{RPE}$ is optimal, but not enforceable, while $\mathbf{B}^{JPE}$ is enforceable. The principal must then either choose a second best RPE contract or $\mathbf{B}^{JPE}$ (a second best RPE contract implies a less extreme RPE contract where agent i can be paid even if (H,L) is not realized, that is $\beta_{LH}^{RPE} < \beta_{LL}^{RPE}$ and/or $\beta_{HL}^{RPE} \geq \beta_{HH}^{RPE} > 0$ (strict inequality if $\beta_{LH}^{RPE} = \beta_{LL}^{RPE}$), where $\beta_{HL}^{RPE} < \frac{c}{(1-\sigma)(1-q_H)\Delta q}$). Since a second best RPE contract is more costly than the optimal RPE contract, $\mathbf{B}^{RPE}$, this will decrease the critical discount factor $\tilde{\delta}$ for when an incentive scheme that exhibits JPE is chosen. Hence, if $\delta^{JPE} < \hat{\delta}(\sigma) < \delta^{RPE}$, then there exist discount factors $\tilde{\delta} < \delta < \hat{\delta}(\sigma)$ where $\mathbf{B}^{JPE}$ is chosen.

I will now show that increasing levels of productivity increases the enforceability of $\mathbf{B}^{RPE}$ relatively to $\mathbf{B}^{JPE}$. In other words, for sufficiently high levels of productivity, $\delta^{RPE} < \delta^{JPE}$. We can rewrite (4) and (5) as

(4') $\frac{1}{q_H+\delta q_L} \leq \frac{\delta}{1-\delta} A$

(5') $\frac{1}{2(1-q_H)} \leq \frac{\delta}{1-\delta} A$,

where $A = \left[\frac{(1-\sigma)\Delta q \Delta Q}{c} - 1 \right] (1-\sigma)\Delta q$ is a proxy for productivity.

Proposition 1: *There is a critical $A = \overline{A}$ such that $\delta^{JPE} = \delta^{RPE}$, that is $\frac{1}{2\overline{A}q_L}\left[-1 - \overline{A}q_L + \sqrt{(1 + 2\overline{A}q_H + \overline{A}^2 q_H^2 + 4\overline{A}q_L)} \right] = \frac{1}{1+2\overline{A}(1-q_H)}$. For given levels of q_H and q_L, if $A > \overline{A}$, then there exist discount factors $\delta^{RPE} \leq \delta < \delta^{JPE}$ where $\mathbf{B}^{RPE}$ is a self-enforcing incentive scheme, and $\mathbf{B}^{JPE}$ is not. If $A < \overline{A}$, then there exist discount factors $\delta^{JPE} \leq \delta < \delta^{RPE}$ where $\mathbf{B}^{JPE}$ is a self-enforcing incentive scheme, and $\mathbf{B}^{RPE}$ is not.*

Hence, if the productivity of effort increases through an increase in ΔQ, and/or a decrease in σ, and/or a decrease in *c,* we can move from a situation where there exists discount factors where $\mathbf{B}^{JPE}$ is enforceable and $\mathbf{B}^{RPE}$ is not, to a situation where $\mathbf{B}^{RPE}$ is enforceable and $\mathbf{B}^{JPE}$ is not. Hence, we obtain an empirical testable hypothesis: *when the productivity of effort is high, RPE is more common. When productivity of effort is low, JPE is more common.* The reasoning goes as follows: the higher the productivity, the easier it is for the principal to offer credible incentive schemes, hence the critical discount factors decrease. When the critical discount factors decrease, the necessary RPE wage, β_{HL}^{RPE} is not affected, but the necessary JPE wage β_{HH}^{JPE} increases since it is harder for the agents to enforce an implicit contract between them on low discount factors. Hence, higher productivity makes the enforceability of RPE rises *relatively* to JPE. Note that higher productivity does not mean that JPE is harder to enforce. High productivity implies however that there is scope for relational contracts even on low discount factors. But $\mathbf{B}^{JPE}$ is difficult to implement on low discount factors due to the implicit contract between the agents.

The proposition can be demonstrated graphically:

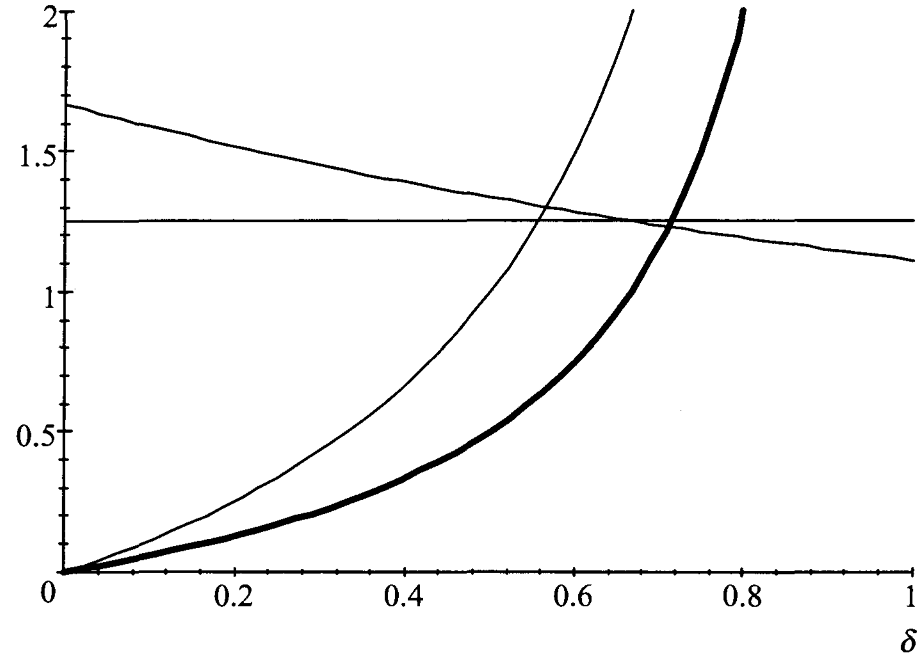


Figure 1

The downward-sloping curve corresponds to the left hand side of (4') (hereafter referred to as the 'JPE curve') and the horizontal curve the left hand side of (5') (hereafter referred to as the 'RPE curve') when $q_H = 0.6$ and $q_L = 0.3$. The thick upward-sloping curve corresponds to the right hand side of (4') and (5') when $A = 0.5$, while the thin upward-sloping curve, corresponds to the right hand side of (4') and (5') when $A = 1$. We see that when $A = 0.5$, $\delta^{JPE} < \delta^{RPE}$. Now, if productivity increases (A increases), given constant probabilities, the upward-sloping curve gets steeper and we move to a situation where $\delta^{RPE} < \delta^{JPE}$, as shown from the thin upward-sloping curve.

Note also that a decrease in q_L, given q_H, increases A (the upward sloping curve gets steeper), while the left hand side of (4') increases so the JPE curve shifts upwards. Hence, a decrease in q_L, given q_H, makes the enforceability of $\boldsymbol{\beta}^{RPE}$ increase *relatively* to the enforceability of $\boldsymbol{\beta}^{JPE}$. An increase in q_H, given q_L, also increases A, but now the JPE curve shifts downwards while the RPE curve shifts upwards. Hence, δ^{JPE} decreases, while δ^{RPE} decreases if $A(1-q_H)$ increases. Hence, an increase in q_H, given q_L, makes the enforceability of $\mathbf{ß}^{RPE}$ increase *relatively* to the enforceability $\mathbf{ß}^{JPE}$, given sufficiently high levels of ΔQ, and/or low levels of σ, and/or low levels of c.

If neither $\mathbf{ß}^{JPE}$ nor $\mathbf{ß}^{RPE}$ is enforceable, the principal must either choose a second best JPE or RPE scheme, or he can chose an incentive scheme based on independent performance evaluation (IPE). The optimal IPE scheme solves (3) subject to $(\beta_{HH}, \beta_{LH}) = (\beta_{HL}, \beta_{LL})$. The agents' incentive constraint is $(\sigma + (1-\sigma)q_H)\beta - c \geq (\sigma + (1-\sigma)q_L)\beta$. Solving for β yields $\beta \geq \frac{c}{(1-\sigma)\Delta q} = \beta_{HH}^{IPE} = \beta_{HL}^{IPE}$. On low realizations, the principal will pay zero. Hence the optimal IPE scheme is $\mathbf{ß}^{IPE} \equiv (\beta_{HH}^{IPE}, \beta_{HL}^{IPE}, 0, 0)$. The condition for the optimal IPE contract, $\mathbf{ß}^{IPE}$, to be self-enforcing is (see appendix)

(9) $$2\beta_{HH}^{IPE} \leq \tfrac{2}{1-\delta}\left[(1-\sigma)\Delta q\Delta Q - c\right],$$

which is equivalent to the JPE condition. In IPE, as in JPE, the principal 'risks' paying both agents. Solving (9) for δ and inserting for $\beta_{HH}^{IPE} = \beta_{HL}^{IPE}$ yields

(10) $$\delta \geq \tfrac{1}{1+A} = \delta^{IPE}.$$

We then have $\delta^{IPE} < \delta^{RPE}$ when $q_H > \frac{1}{2}$ and $\delta^{IPE} < \delta^{JPE}$ for sufficiently high productivity. We can write (9) as

(9') $$1 \leq \tfrac{\delta}{1-\delta} A$$

Proposition 2: There is a critical $A = \tilde{A}$ such that $\delta^{JPE} = \delta^{IPE}$, that is $\frac{1}{2\tilde{A}q_L}\left[-1 - \tilde{A}q_L + \sqrt{(1 + 2\tilde{A}q_H + \tilde{A}^2 q_H^2 + 4\tilde{A}q_L)}\right] = \frac{1}{1+\tilde{A}}$. *For given levels of* q_H *and* q_L, *if* $A > \tilde{A}$, *and* $q_H > \frac{1}{2}$, *then there exist discount factors* $\delta^{IPE} \leq \delta < \delta^{JPE}$ *and* $\delta^{IPE} \leq \delta < \delta^{RPE}$ *where* $\mathbf{ß}^{IPE}$ *is self-enforcing and* $\mathbf{ß}^{JPE}$ *and* $\mathbf{ß}^{RPE}$ *are not.*

The proposition shows that the enforceability of $ß^{IPE}$ relative to $ß^{JPE}$ increases with effort productivity, for the same reasons as $ß^{RPE}$ to $ß^{JPE}$. Moreover the enforceability of $ß^{IPE}$ relative to $ß^{RPE}$ increases on high probabilities for positive outcome: while the necessary IPE wage β_{Hj}^{IPE} is independent of q_H, the necessary RPE wage, β_{HL}^{RPE}, increases in q_H, given q_L.

We can demonstrate proposition 2 graphically:

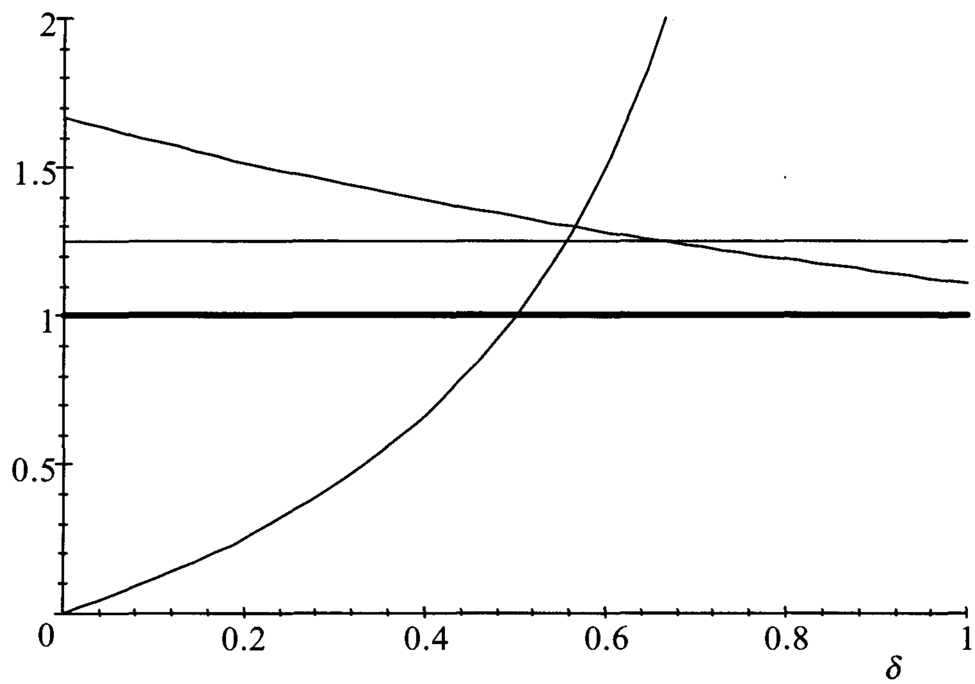


Figure 2

The parameters are the same as in Figure 1, with $A=1$. The thick horizontal line represents the 'IPE curve'(corresponding to the left hand side (9')), while the thin horizontal line is the 'RPE curve'. We see that for $q_H > \frac{1}{2}$ and a sufficiently high A, then $\delta^{IPE} < \delta^{RPE} < \delta^{JPE}$.

So far we have seen that high effort productivity strengthens the self-enforcing conditions of RPE and IPE schemes relatively to the optimal JPE scheme. The analysis reveals that the outcomes in Che and Yoo's model are sensitive to the assumption that the parties can write legally enforceable contracts. Once contracts have to rely on self-enforcement, the optimal choice of team incentives becomes more complicated, since the parties can only choose between self-enforcing incentive contracts. Some numerical examples elucidate this:

Assume that the discount factor is constant over time and has the value $\delta = 0.75$

Example 1

Parameters	**Estimates**
$c = 1$	$\delta^{IPE} = 0.74$
$\sigma = 0.05$	$\delta^{RPE} = 0.70$
$q_H = 0.4$	$\delta^{JPE} = 0.83$
$q_L = 0.2$	$\hat{\delta} = 0.66$
$\Delta Q = 15$	

β^{JPE} is optimal but not enforceable, while β^{RPE} is enforceable.

Example 2

Parameters	**Estimates**
$c = 1$	$\delta^{IPE} = 0.74$
$\sigma = 0.2$	$\delta^{RPE} = 0.83$
$q_H = 0.7$	$\delta^{JPE} = 0.74$
$q_L = 0.4$	$\hat{\delta} = 0.89$
$\Delta Q = 10$	

β^{RPE} is optimal, but not enforceable, while β^{JPE} is enforceable.

Example 3

Parameters	**Estimates**
c = 0.7	$\delta^{IPE} = 0.72$
$\sigma = 0.4$	$\delta^{RPE} = 0.81$
$q_H = 0.7$	$\delta^{JPE} = 0.77$
$q_L = 0.1$	$\hat{\delta} = 9.52$
$\Delta Q = 4$	

β^{RPE} is optimal, but neither β^{JPE} or β^{RPE} is enforceable, while β^{IPE} is enforceable.

4. Collusion

As noted, in contrast to IPE and JPE, RPE is exposed to collusion. If the principal offers β^{RPE}, then the agents, for δ close to 1, are better off if they both play low effort (L,L) than if they both play high effort (H,H). And for any value of δ, the agents can play a correlated randomisation of (L,H) and (H,L) and generate a higher joint payoff than (H,H). Hence, if there are no institutional constraints to collusion between the agents, then the cost of using RPE to generate (H,H) raises. It is complicated to analyse the optimal scheme in the region where the principal 'would have chosen' RPE if it was not for the collusion problem. In order to prevent collusion, the principal has to offer incentives which make high effort from both agents (H,H) the only subgame perfect equilibrium. A correlated randomisation strategy between (H,L) and (L,H) provides the 'hardest' equilibrium to prevent. It can be shown (see appendix) that the condition that makes this randomisation strategy a non equilibrium is given by[6]

(12) $$\beta_{HL} \geq \frac{(2-\delta)c}{(1-\sigma)(2-2q_H-\delta)\Delta q} = \beta_{HL}^{RPEC}.$$

[6] I thank Yeon-Koo Che for help with deducing this expression.

It will of course be more difficult to enforce RPE schemes if the agents can collude. But the basic implications from proposition 1 and 2 are not altered. If we compare β_{HL}^{RPE} to β_{HL}^{RPEC} we see that the first order effects of $\sigma, \Delta q, q_H$ and c is corresponding. Moreover, a decrease in δ decreases β_{HL}^{RPEC},[7] while it increases β_{HH}^{JPE}. Hence, the economics of the observations in the previous section is robust to the threat of collusion.

Two things are happening to optimality conditions when we allow for collusion. First, since RPE is now more expensive (that is $\beta_{HL}^{RPE} < \beta_{HL}^{RPEC}$ for $\delta > 0$), the region where β^{JPE} is optimal expands. Second, in contrast to the no collusion case, relative performance evaluation does not always dominate independent performance evaluation, since the expected cost per agent in RPE, $(1-\sigma)q_H(1-q_H)\beta_{HL}^{RPEC}$, exceeds the expected cost per agent in IPE, $(\sigma + (1-\sigma)q_H\beta_{HL}^{IPE}$, when $\delta > \dfrac{2\sigma(1-q_H)}{q_H^{2}(1-\sigma)+\sigma} = \underline{\delta}$. And since the expected cost per agent in IPE is lower than the expected cost per agent in JPE, $(\sigma + (1-\sigma)q_H q_H)\beta_{HH}^{JPE}$ when $\delta < \dfrac{\sigma(1-q_H)}{q_L(q_H(1-\sigma)+\sigma)} = \overline{\delta}$, there are levels of discount factors where IPE is optimal.

Proposition 3: *If the agents are able to collude, there exist discount factors $\overline{\delta} > \delta > \underline{\delta}$ where independent performance evaluation is optimal.*

Observe that the larger $q_H - q_L = \Delta q$, the larger is the region $(\underline{\delta}, \overline{\delta})$ where IPE is optimal. The reason is that JPE is costly when q_L is low and RPE is costly when q_H is large.

5. Asset ownership

So far I have considered the principal to be the owner of the assets involved. Since the principal was able to utilize the values that the agents created even after reneging on the bonus-payment, the assumption was made that the principal owned the critical assets involved in producing values. This section discusses the self-enforcing conditions of the incentive

[7] Observe that when $\delta = 0$, $\beta_{HL}^{RPE} = \beta_{HL}^{RPEC}$. As δ approaches $2-2q_H$, β_{HL}^{RPEC} approaches infinity. For $\delta \in (2-2q_H, 1)$, RPE can never succeed in preventing the randomization collusion.

schemes when the agents own critical assets so that they are able to renegotiate the terms of trade ex post the realization of values. I will argue that independent performance evaluation is more likely in this situation.

Assume that there are two assets involved in the production, and that each agent uses one asset to produce values for the principal. Asset ownership conveys ownership of the values produced. In the previous section it was implicitly assumed that the principal owned both assets ('principal-ownership'). In this section I assume that the agents own one asset each ('agent-ownership'). This does not necessarily mean that the agents are independent suppliers owning physical assets. It can also be interpreted as agents in an employment relationship, where the critical assets are human capital. The main difference from the previous sections is that the agents are able to renegotiate the terms of trade ex post the realization of values.

In agent-ownership, if the agents deviate from the relational contract by refusing to accept the promised bonus, they can "hold up" the values they have produced and renegotiate the price. Assume that there exist an alternative market for the output, and that the price in this market is either high, P_H, or low P_L, where $Q_H > Q_L > P_H > P_L$. The favourable shock, which occurs with probability $\sigma \in (0,1)$, makes both agents produce high values both for the principal and for the alternative market. I assume one-dimensional effort. This means that the agents cannot take actions that increase the probability of realizing high- alternative use values P_H, without also increasing the probability of realizing high values for the principal Q_H. Hence, if the shock is unfavourable, the probability of realizing P_H is q_H if the agent's effort is high and q_L if the agent's effort is low. The one dimensional effort constraint ensure that agent i's wage vector $\mathbf{\beta}^i \equiv (\beta^i{}_{HH}, \beta^i{}_{HL}, \beta^i{}_{LH}, \beta^i{}_{LL})$ still applies since there is then no point in letting wage depend on the realization of alternative use values Hence, the optimal schemes in principal-ownership, $\mathbf{\beta}^{JPE}, \mathbf{\beta}^{RPE}$ and $\mathbf{\beta}^{IPE}$, also applies in agent-ownership. Note that the one-dimensional effort constraint is most natural in the human capital interpretation of asset ownership, where the principal can decide the agent's behavioral pattern even though the agents own the critical assets.

In general, the principal's reneging temptations are weaker in agent-ownership than in principal-ownership, since he in agent-ownership cannot just take the goods, but has to bargain with the agents if he reneges. I assume that 50:50 Nash bargaining determines the price, that is $\frac{Q_i+P_j}{2}$. The agents' reneging temptations are stronger, however, since they can receive the Nash price if they renege. This 'outside opportunity' is especially binding with non-independent performance evaluation since the agents risk getting low wages from the relational contract, while high realization assure them a high Nash price if they renegotiate.

The condition for the optimal JPE contract, $\mathbf{ß}^{JPE}$, to be self-enforcing is (see appendix):

When $2\beta_{HH}^{JPE} > \Delta Q$

(13a) $2\beta_{HH}^{JPE} - \frac{1}{2}\Delta Q + \Delta P \leq \frac{2\delta}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$

When $2\beta_{HH}^{JPE} < \Delta Q$

(13b) $\frac{1}{2}\Delta Q + \Delta P \leq \frac{2\delta}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$

where $\Delta P = P_H - P_L$. Compared to (13a,b) to (4) shows that it is easier to enforce $\mathbf{ß}^{JPE}$ in agent-ownership than in principal-ownership only if $\Delta P < \frac{1}{2}\Delta Q$. A disadvantage with the optimal JPE scheme, $\mathbf{ß}^{JPE}$, in agent-ownership is that if only one of the agents realize high values, he receives no payments and is therefore tempted to renege; that is to renegotiate a Nash price with the principal. This 'non-independence' problem of agent-ownership becomes even more severe with the optimal RPE scheme $\mathbf{ß}^{RPE}$. The conditions for the RPE contract $\mathbf{ß}^{RPE}$ to be self-enforcing is (see appendix):

When $\beta_{HL}^{RPE} > \frac{1}{2}\Delta Q$

(14a) $\beta_{HL}^{RPE} + \frac{1}{2}\Delta Q + \Delta P \leq \frac{2\delta}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$

When $\beta_{HL}^{RPE} < \frac{1}{2}\Delta Q$

(14b) $\Delta Q + \Delta P \leq \frac{2\delta}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$

Comparing (14a,b) and to (5) shows that it is more difficult to enforce $\mathbf{ß}^{RPE}$ in the agent-ownership than in principal-ownership. Now, if *both* agents realize high values, they get zero, and both have high temptations to renege on the contract since the Nash price is high (due to the high output-realizations.)

This non-independent problem is eliminated with the β^{IPE} scheme. Since each agent gets paid according to their own output, independent from their peer's output, they do not risk receiving no payments within the contract while the Nash price is high. The condition for the IPE contract β^{IPE} to be self-enforcing is (see appendix):

$$(15) \quad \left|2\beta_{HH}^{IPE} - \Delta Q\right| + \Delta P \leq \tfrac{2\delta}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$$

Comparing (15) to (9) we see that it is easier to enforce β^{IPE} in agent-ownership than in principal-ownership when $\Delta Q > \Delta P$, hence if effort is more productive internally than externally (effort specificity).

A comparison of the constraints above suggests that the relative enforceability of β^{IPE} to β^{JPE} and β^{RPE} is strengthened if the agents own the assets. Since IPE, in contrast to JPE and RPE do not exhibit the 'non-independent problem' we would expect that *independent performance evaluation is relatively more common if agents own the critical assets (or have essential human capital) so that they are able to renegotiate the terms of trade ex post the realization of values, than if the principal owns the critical assets.*

6. Relevance

There are two important features with the model discussed in the previous section that decides its applicability. First, the agents can observe each other's actions. Consequently, the model is best applied on smaller organizations, or on subdivisions of larger organizations. Second, the relational contract is multilateral, which implies that the agents cooperatively decide whether or not to honour the contract. Hence, the model is best applied in environments or corporate cultures where worker-coordination is relatively easy.

Do we find empirical support for the model's predictions? A common wage/incentive structure of hierarchical organizations is that low-wage 'blue collar' workers in the bottom of a hierarchy enjoy some sort of group incentives, while 'white collar' workers higher up in the organization typically enjoy incentives based on individual or relative performance (see Prendergast, 1999 and Appelbaum and Berg, 1999). Baker, Gibbs and Holmström (1993), and Treble, van Gameren, Bridges and Barmby (2001), find that the salary range is much greater at the higher levels of an organization than at lower levels, which indicates a higher frequency of individual and relative performance based compensation structure on higher levels.

The model help explain these observations: Proposition 1 and 2 suggest that RPE or IPE are more common when the productivity of effort is high, while JPE is more common when productivity of effort is low. Since wages are expected to reflect the workers' productivity, we would therefore expect to see higher occurrence of IPE and RPE in higher levels of organizations where wage and productivity presumably are highest.

The higher frequency of IPE in highly paid jobs can also be explained by the non-independence problem of RPE and JPE discussed in Section 5. When the workers own the critical assets, or have essential human capital so that they are able to renegotiate the terms of trade ex post the realization of values, incentives based on IPE has the advantage that it balances the value of relational contracting with the value of independent ex post bargaining. With incentive schemes based on non-independent performance evaluation, there is a greater probability of getting low payments inside the relation and high payments outside the relation, which increases the possibility of contract breach. This theoretical result is not surprising. It helps explain the higher frequency of individual compensation packages and IPE incentives in human capital-intensive industries.

It is argued that RPE discourages cooperative work morale.[8] Human recourses managers often claim that salaries must be compressed to maintain internal harmony in a firm. If the

[8] It is also argued that RPE, in addition to discourage cooperative work morale, also encourages employees to adopt restricted work norms (see Baron and Kreps, 1999). RPE can also distort agents' incentives if they carry out multiple activities (multitasking). Gibbons and Murphy (1990), and more formally Baker (1992) show that if workers can take actions that effect the output of their peers, and in addition are able to "game" the compensation scheme to their benefit, RPE can distort incentives and thus make it less efficient.

difference between the winner's salary and the loser's salary is too great, morale suffers. Since the high positions in an organization tend to be dominated by individuals who have managed to fight the "corporate war", and thus presumably tend to be more aggressive and more willing to engage in sabotage, one should find it necessary to reduce the incentives for those workers to compete with each other (see for instance Lazear, 1989, 1998); in other words reduce: the use of relative performance evaluation. But the model in this paper shows that it is easier to implement incentives based on relative performance evaluation than joint performance evaluation in high ability/low trust environments (using the discount factor as a proxy for trust; see Hart 2001).

7. Conclusion

It can be efficient to reward a group on the basis of the group's joint performance if the problem of free riding can be deterred by mutual peer monitoring and social sanctions. But incentives based on relative performance evaluation and independent performance evaluation are still quite common even in industries were peer monitoring is possible. Individual compensation based on IPE and/or RPE is especially common in the higher levels of organizations. In this paper I have shown how the absence of legally enforceable contracts can explain this. In a model with self-enforcing relational contracts between principal and agents it is shown that we can expect a relatively higher frequency of incentive schemes based on RPE and IPE when the productivity of effort is high. Moreover it is shown that we can expect to see a relatively higher frequency of IPE schemes if agents own critical assets so that they can renegotiate the terms of trade ex post realization of values.

This paper has not fully characterized optimal solutions, but showed the important implications of the relational contract's enforceability constraints when dealing with team incentives. In formal studies of relational contracts it is important to bear in mind the subjective nature of the discount factor. It is not necessarily optimal for a principal to choose the optimal incentive scheme among the enforceable ones at a given date. If the discount factors vary over time, and there are costs associated with shifting from one scheme to another, it may be optimal to choose the incentive scheme with the lowest critical discount factor.

APPENDIX

1. Sketch to proof of Che and Yoo's Lemma

First we can write out the objective function:

$$\text{(A.1)}\quad \begin{aligned}&\pi(H,H,\mathbf{B})=(\sigma+(1-\sigma)q_Hq_H)\beta_{HH}\\&+(1-\sigma)q_H(1-q_H)(\beta_{HL}+\beta_{LH})+(1-\sigma)(1-q_H)^2\beta_{LL}\end{aligned}$$

As noted, in an RPE scheme, $\pi(L,L;\mathbf{B})>\pi(L,H;\mathbf{B})$, so the incentive constraint (2) is reduced to $\pi(H,H;\mathbf{B})-c\geq\pi(L,H;\mathbf{B})$. We can write this out:

$$\begin{aligned}&(\sigma+(1-\sigma)q_Hq_H)\beta_{HH}+(1-\sigma)[q_H(1-q_H)\beta_{HL}+(1-q_H)q_H\beta_{LH}+(1-q_H)(1-q_H)\beta_{LL}]-c\\&\geq(\sigma+(1-\sigma)q_Lq_H)\beta_{HH}+(1-\sigma)[q_L(1-q_H)\beta_{HL}+(1-q_L)q_H\beta_{LH}+(1-q_L)(1-q_H)\beta_{LL}]\end{aligned}$$

and simplify it to

$$\text{(A.2)}\quad q_H\beta_{HH}+(1-q_H)\beta_{HL}-q_H\beta_{LH}-(1-q_H)\beta_{LL}\geq\tfrac{c}{(1-\sigma)(q_H-q_L)}$$

The left hand side of the constraint is decreasing in β_{LL}, while the objective function is increasing in β_{LL}. Hence, it is optimal to set $\beta_{LL}=0$, which from $\pi(L,L;\mathbf{B})>\pi(L,H;\mathbf{B})$ implies that $\beta_{LH}=0$. Since both the objective function and the constraint are linear in $\mathbf{B}$, only β_{HH} or β_{HL} are strictly positive, which from$\pi(H,L;\mathbf{B})>\pi(H,H;\mathbf{B})$ implies $\beta_{HH}=0$. Hence , the optimal RPE scheme is an the extreme form $\mathbf{B}^{RPE}\equiv(0,\beta_{HL}^{RPE},0,0)$ where solving (A.2) for β_{HL} yields $\beta_{HL}^{RPE}\equiv\dfrac{c}{(1-\sigma)(1-q_H)\Delta q}$.

In a JPE scheme, $\pi(L,H;\mathbf{B})>\pi(L,L;\mathbf{B})$. Thus the incentive constraint (2) becomes $\frac{1}{1-\delta}\pi(H,H;\mathbf{B})-c\geq\pi(L,H;\mathbf{B})+\frac{\delta}{1-\delta}\pi(L,L;\mathbf{B})$. We can write this out:

$$\tfrac{1}{1-\delta}\{(\sigma+(1-\sigma)q_H q_H)\beta_{HH}+(1-\sigma)[q_H(1-q_H)\beta_{HL}+(1-q_H)q_H\beta_{LH}+(1-q_H)(1-q_H)\beta_{LL}]-c\}$$
$$\geq(\sigma+(1-\sigma)q_L q_H)\beta_{HH}+(1-\sigma)[q_L(1-q_H)\beta_{HL}+(1-q_L)q_H\beta_{LH}+(1-q_L)(1-q_H)\beta_{LL}]$$
$$+\tfrac{\delta}{1-\delta}\{(\sigma+(1-\sigma)q_L q_L)\beta_{HH}+(1-\sigma)[q_L(1-q_L)\beta_{HL}+(1-q_L)q_L\beta_{LH}+(1-q_L)(1-q_L)\beta_{LL}]\}$$

and simplify it to

$$\text{(A.3)}\quad \begin{aligned}&(q_H+\delta q_L)\beta_{HH}+(1-q_H-\delta q_L)\beta_{HL}+(\delta-q_H-\delta q_L)\beta_{LH}-(1-q_H+\delta(1-q_L))\beta_{LL}\\ &\geq\tfrac{c}{(1-\sigma)(q_H-q_L)}\end{aligned}$$

The left hand side of the constraint is decreasing in β_{LL}, while the objective function is increasing in β_{LL}. Hence, it is optimal to set $\beta_{LL}=0$. Observe that the coefficient of β_{HL} is weakly greater than that of β_{LH} in the left hand side of (A.3), but that their coefficients are the same in the objective function (A.1). Suppose $\beta_{LH}>0$. Then lowering β_{LH} and raising β_{HL} simultaneously so that the left hand side of (A.3) remains the same will reduce the value of the objective function. Hence, it is optimal to set $\beta_{LH}=0$. Since only β_{HH} or β_{HL} are strictly positive $\beta_{HH}=0$ from $\pi(H,L;\mathbf{B})<\pi(H,H;\mathbf{B})$. Hence , the optimal JPE scheme is an extreme form $\mathbf{B}^{JPE}\equiv(\beta_{HH},0,0,0)$ where solving (A.3) for β_{HH} yields $\beta_{HH}^{JPE}\equiv\dfrac{c}{(1-\sigma)(q_H+\delta q_L)\Delta q}$.

Now, $\mathbf{B}^{JPE}$ solves (3) if $\pi(H,H,\mathbf{B}^{JPE})<\pi(H,H,\mathbf{B}^{RPE})$. That is

$$(\sigma+(1-\sigma)q_H q_H)\frac{c}{(1-\sigma)(q_H+\delta q_L)\Delta q}<(1-\sigma)q_H(1-q_H)\frac{c}{(1-\sigma)(1-q_H)\Delta q}$$

Solving for δ yields

$$\delta>\tfrac{\sigma}{(1-\sigma)q_H q_L}\equiv\hat{\delta}(\sigma)$$

For further details see Che and Yoo (2001).

2. The conditions for self-enforcing relational contracts when the principal owns the assets

In JPE, the binding constraint for the principal is when both agents realize high values so that he has to pay wage β_{HH}^{JPE} to both agents. The binding constraint for the principal to honour the contract $\mathbf{B}^{JPE}$ is then given by

$$\text{(A.4)} \quad \begin{aligned} &-2\beta_{HH}^{JPE} + \tfrac{2\delta}{1-\delta}\left[Q_L + (\sigma + (1-\sigma)q_H)\Delta Q - (\sigma + (1-\sigma)q_H q_H)\beta_{HH}^{JPE} - \alpha\right] \\ &\geq \tfrac{2\delta}{1-\delta}\left[Q_L + (\sigma + (1-\sigma)q_L)\Delta Q\right] \end{aligned},$$

where the left hand side shows the expected present value from honouring and the right hand side shows the expected value from reneging. The square brackets show the principal's expected revenue per agent per period.

For the agents, the constraints bind for realization of zero bonus wages. Each agent will thus simply honour the contract if expected future wage exceeds zero:

$$\text{(A.5)} \quad \tfrac{\delta}{1-\delta}(\alpha + (\sigma + (1-\sigma)q_H q_H) \;\; \beta_{HH}^{JPE} - c) \geq 0$$

Combining (A.4) with (A.5) for both agents, that is multiplying (A.5) with 2 and add with (A.4) yields:

$$\text{(4)} \quad 2\beta_{HH}^{JPE} \leq \tfrac{2\delta}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$$

For (4) to be a sufficient condition for the relational contract to hold, either (A.4) or (A.5) must hold with equality. The fixed payment can always be chosen in a way that either (A.4) or (A.5) holds with equality. From the incentive constraint (2), we see that α must be negative for (A.5) to hold with equality. Note however; I have not included ex ante participation constraints in the model. But observe that if the agents have ex ante outside opportunities $\overline{w}$, then the principal must satisfy (A.5') $\tfrac{\delta}{1-\delta}(\alpha + (\sigma + (1-\sigma)q_H q_H) \;\; \beta_{HH}^{JPE} - c) \geq \overline{w}$ for both agents to honour the contract, implying that the fixed salary need not be negative for a (A.5) to hold with equality. As long as low effort is costless, the agents would not have incentives to *not participate* if the participation constraint is satisfied. Moreover, the agents cannot use ‘no participation’ as a threat in the relational contract if the principal then can contract with other agents at the same cost.

In RPE, the principal only has to pay wage to one of the agents, and this occurs only when one realizes high values and one realizes low values. The binding constraint for the principal to honour the contract $\mathbf{B}^{RPE}$ is thus

$$(A.6)\quad \begin{aligned}&-\beta_{HL}^{RPE}+\tfrac{2\delta}{1-\delta}\left[Q_L+(\sigma+(1-\sigma)q_H)\Delta Q-(1-\sigma)q_H(1-q_H)\beta_{HL}^{RPE}-\alpha\right]\\&\geq \tfrac{2\delta}{1-\delta}\left[Q_L+(\sigma+(1-\sigma)q_L)\Delta Q\right]\end{aligned}$$

The binding constraint for each agent is

$$(A.7)\quad \tfrac{\delta}{1-\delta}(\alpha+(1-\sigma)q_H(1-q_H)\ \ \beta_{HL}^{RPE}-c)\geq 0$$

Combining (A.6) and (A.7) for both agents yields

$$(5)\qquad \beta_{HL}^{RPE}\leq \tfrac{2\delta}{1-\delta}\left[(1-\sigma)\Delta q\Delta Q-c\right]$$

In IPE, the binding constraint for the principal to honour the contract $\mathbf{B}^{IPE}$ is given by

$$(A.8)\quad \begin{aligned}&-2\beta_{HH}^{IPE}+\tfrac{2\delta}{1-\delta}\left[Q_L+(\sigma+(1-\sigma)q_H)\Delta Q-(\sigma+(1-\sigma)q_H)\beta_{HH}^{IPE}-\alpha\right]\\&\geq \tfrac{2\delta}{1-\delta}\left[Q_L+(\sigma+(1-\sigma)q_L)\Delta Q\right]\end{aligned}$$

The binding constraint for each agent is

$$(A.9)\quad \beta_{HH}^{IPE}+\tfrac{\delta}{1-\delta}(\alpha+(\sigma+(1-\sigma)q_H)\ \ \beta_{HH}^{IPE}-c)\geq 0$$

Combining (A.8) and (A.9) for both agents yields

$$(9) \qquad 2\beta_{HH}^{JPE} \le \tfrac{2\delta}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$$

3. The conditions for self-enforcing relational contracts when the agents own the assets.

The principal will honour the contract $\mathbf{\beta}^{JPE}$ if

$$(A.10) \quad \begin{aligned} &-2\beta_{ij}^{JPE} + \tfrac{2\delta}{1-\delta}\left[Q_L + (\sigma + (1-\sigma)q_H)\Delta Q - (\sigma + (1-\sigma)q_H q_H)\beta_{HH}^{JPE} - \alpha\right] \\ &\ge -\tfrac{(Q_i + P_L)}{2} - \tfrac{(Q_j + P_L)}{2} + \tfrac{2\delta}{1-\delta}\left[Q_L + (\sigma + (1-\sigma)q_L)\Delta Q - \gamma\right] \end{aligned}$$

where $\gamma = \frac{Q_L + (\sigma + (1-\sigma)q_L)\Delta Q + P_L + (\sigma + (1-\sigma)q_L)\Delta P}{2}$ is the expected Nash bargaining price. To see which realizations bind, we must study (i) $2\beta_{ij}^{JPE} \le \frac{Q_i}{2} + \frac{Q_j}{2}$. Observe that (H,L) and (L,H) never bind since the right hand side of (i) is then higher than (L,L) while both (L,L), (H,L) and (L,H) yields zero wage. Hence, we must compare (ii) $2\beta_{HH}^{JPE} \le \frac{Q_H}{2} + \frac{Q_H}{2}$ with (iii) $2\beta_{LL}^{JPE} = 0 \le \frac{Q_L}{2} + \frac{Q_L}{2}$. We see that if the difference between the left hand side of (ii) and (iii) are smaller than the difference between the right hand side of (ii) and (iii), that is $2\beta_{HH}^{JPE} - 0 < Q_H - Q_L$, then (L,L) binds since (iii) are then weaker than (ii). Hence, when $2\beta_{HH}^{JPE} > \Delta Q$ (A.10) binds on high realizations of Q, and when $2\beta_{HH}^{JPE} < \Delta Q$, (A.10) binds on low realizations.

We see that the only difference from (A.4) where the principal has all the ex post bargaining power, is that the deviation payoff on the right hand side is smaller since he cannot just take the good, but has to pay for it after deviation. Note also that for a relational contract to exist, the expected Nash bargaining price cannot satisfy the incentive constraint, because then would either the principal, if it was sufficiently low, or the agents if it was sufficiently high, insist on spot trading, instead of long term contracts. Hence, if there exists a relational contract, the agent's will play low effort after deviation.

The binding constraint for agent i is when he realizes high output as agent j realizes low output. The optimal JPE scheme, β^{JPE}, then yields no wage to the agent, while Nash bargaining yields a higher deviation payoff. Hence, the pair of constraints that bind are

(A.11) $\beta_{HL}^{JPE} + \frac{\delta}{1-\delta}(\alpha + (\sigma + (1-\sigma)q_H q_H) \quad \beta_{HH}^{JPE} - c) \geq \frac{Q_H + P_H}{2} + \frac{1}{1-\delta}\gamma$

(A.12) $\beta_{HL}^{JPE} + \frac{\delta}{1-\delta}(\alpha + (\sigma + (1-\sigma)q_H q_H) \quad \beta_{HH}^{JPE} - c) \geq \frac{Q_L + P_H}{2} + \frac{1}{1-\delta}\gamma$

When $2\beta_{HH}^{JPE} > \Delta Q$, combining (A.10) (A.11) and (A.12) yields

(13a) $2\beta_{HH}^{JPE} - \frac{1}{2}\Delta Q + \Delta P \leq \frac{2\delta}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$

When $2\beta_{HH}^{JPE} < \Delta Q$, combining (A.10) (A.11) and (A.12) yields

(13b) $\frac{1}{2}\Delta Q + \Delta P \leq \frac{2\delta}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$

In RPE, the principal will honour the contract β^{RPE} if

$$\text{(A.13)} \quad \begin{aligned} &-\beta_{ij}^{RPE} + \tfrac{2\delta}{1-\delta}\left[Q_L + (\sigma + (1-\sigma)q_H)\Delta Q - (1-\sigma)q_H(1-q_H)\beta_{HL}^{RPE} - \alpha\right] \\ &\geq -\tfrac{(Q_i + P_L)}{2} - \tfrac{(Q_j + P_L)}{2} + \tfrac{2\delta}{1-\delta}\left[Q_L + (\sigma + (1-\sigma)q_L)\Delta Q - \gamma\right] \end{aligned}$$

Observe that (H,H) never binds, since this yields zero wage outlays if he honours, and high wage outlays (in terms of Nash prices) if he deviates. Comparing $\beta_{HL}^{RPE} \leq \frac{Q_H}{2} + \frac{Q_L}{2}$ to $\beta_{LL}^{RPE} = 0 \leq \frac{Q_L}{2} + \frac{Q_L}{2}$ shows that when $\beta_{HL}^{RPE} > \frac{1}{2}\Delta Q$, then (H,L) binds and when $\beta_{HL}^{RPE} < \frac{1}{2}\Delta Q$, then (L,L) binds.

The binding constraint for both agents is

(A.14) $\beta_{HH}^{RPE} + \frac{\delta}{1-\delta}(\alpha + (\sigma + (1-\sigma)q_H(1-q_H)) \quad \beta_{HL}^{RPE} - c) \geq \frac{Q_H + P_H}{2} + \frac{1}{1-\delta}\gamma$

Observe that the constraint binds when both agents realize high output. The optimal RPE scheme, $\mathbf{B}^{RPE}$, then yields no wage to the agents, while high realizations yield a higher deviation payoff.

When $\beta_{HL}^{RPE} > \frac{1}{2}\Delta Q$, combining (A.13) and (A.14) then yields

(14a) $\beta_{HL}^{RPE} + \frac{1}{2}\Delta Q + \Delta P \leq \frac{2\delta}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$

When $\beta_{HL}^{RPE} \leq \frac{1}{2}\Delta Q$, combining (A.13) and (A.14) then yields

(14b) $\Delta Q + \Delta P \leq \frac{2\delta}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$

In IPE, the principal will honour the contract $\mathbf{B}^{IPE}$ if

$$\text{(A.15)}\quad \begin{aligned} &-2\beta_{ij}^{IPE} + \tfrac{2\delta}{1-\delta}\left[Q_L + (\sigma + (1-\sigma)q_H)\Delta Q - (\sigma + (1-\sigma)q_H)\beta_{HH}^{IPE} - \alpha\right] \\ &\geq -\tfrac{(Q_i+P_L)}{2} - \tfrac{(Q_j+P_L)}{2} + \tfrac{2\delta}{1-\delta}\left[Q_L + (\sigma + (1-\sigma)q_L)\Delta Q - \gamma\right] \end{aligned}$$

Each agent will honour the contract if

(A.16) $\beta_{ij}^{IPE} + \frac{1}{1-\delta}(\alpha + (\sigma + (1-\sigma)q_H)\ \ \beta_{HH}^{IPE} - c) \geq \frac{Q_i + P_H}{2} + \frac{\delta}{1-\delta}\gamma$

We see that when $2\beta_{HH}^{JPE} > \Delta Q$, (A.15) binds on high realizations of Q, and (A.16) binds on low realizations. Combining (A.15) with (A.16) for both agents yields $2\beta_{HH}^{IPE} - \Delta Q + \Delta P \leq \frac{2}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$. When $2\beta_{HH}^{JPE} < \Delta Q$, (A.15) binds on low realization of Q, and (A.16) binds on high realizations. Combining (A.15) with (A.16) for both agents yields $-2\beta_{HH}^{IPE} + \Delta Q + \Delta P \leq \frac{2}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$. Using absolutes, we can reduce it to one condition:

(15) $\left|2\beta_{HH}^{IPE} - \Delta Q\right| + \Delta P \leq \frac{2\delta}{1-\delta}\left[(1-\sigma)\Delta q \Delta Q - c\right]$

4. Deducing (12)

For (H,H) to be an equilibrium, β_{HL} must be chosen so that the randomisation becomes a non-equilibrium. This means that the workers must find it profitable to deviate from the randomisation strategy. If the worker fulfils the strategy, his payoff is

$$\text{(A.17)}\quad \alpha+\beta_{HL}(1-\sigma)q_L(1-q_H)+\tfrac{\delta}{1-\delta}\tfrac{1}{2}(\alpha+\beta_{HL}(1-\sigma)(q_H(1-q_L)+q_L(1-q_H))-c)$$

A worker may only deviate from the randomisation strategy when he gets a turn to play low effort. The worker can then deviate by playing high effort, and this will be punished in the future by $(H,H)^{\infty}$. Hence the deviation payoff will be

$$\text{(A.18)}\quad \tfrac{1}{1-\delta}(\alpha+\beta_{HL}(1-\sigma)q_H(1-q_H)-c)$$

To make the randomisation a non-equilibrium, β_{HL} must be chosen so that (A.18) exceeds (A.17). That is

$$\text{(12)}\quad \beta_{HL}\geq\frac{(2-\delta)c}{(1-\sigma)(2-2q_H-\delta)\Delta q}=\beta_{HL}^{RPEC}$$

For a sufficiently high δ, (12) becomes negative; and there exists no β_{HL} that will induce (A.18) to exceed (A.17). In that case, RPE can never succeed in preventing the randomisation collusion. But for sufficiently low discount factors there exist critical δ values for the randomisation strategy to be non-equilibrium.

REFERENCES

Alchian, Armen A., and Demsetz, Harold. 1972. "Production, Information Costs, and Economic Organization." *American Economic Review* 62: 777-95.

Appelbaum, Eileen and Peter Berg. 1999. "High-Performance Work Systems: Giving workers a stake. "In M. Blair and T.A. Kochan, eds: *The New Relationship. Human Capital in the American Corporation*. The Brooking Institution Press

Baker, George. 1992. "Incentive Contracts and Performance Measurement." *Journal of Political Economy*, 100: 598-614.

Baker, George, Robert Gibbons and Kevin J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts." *Quarterly Journal of Economics* 109: 1125-56.

Baker, George, Robert Gibbons and Kevin J. Murphy. 2002. "Relational Contracts and the Theory of the Firm." *Quarterly Journal of Economics* 117: 39-94.

Baker, George, Michael Gibbs and Bengt Holmström. 1993. "Hierarchies and Compensation. A Case Study." *European Economic Review* 37: 366-78.

Banker, R.D., J.M Field, R.C. Schroeder and K.K Sinha. 1996. "Impact of Work Teams on Manufacturing Performance: A longitudinal Field Study." *Academy of Management Journal* 39: 867-890.

Baron, James N. and Kreps, David M. 1998. *Strategic Human Resources: Frameworks for General managers*. New York: Wiley.

Bewley, Truman. 1999. *Why Wages Don't Fall During a Recession*. Cambridge, MA: Harvard University Press.

Bull, Clive. 1987. "The existence of Self-Enforcing Implicit Contracts. " *Quarterly Journal of Economics*. 102: 147-59

Che, Yeon-Koo and Seung-Weon Yoo 2001. Optimal Incentives for Teams. *American Economic Review* 91: 525-40

FitzRoy, Felix R. and Kornelius Kraft. 1995. "On the choice of Incentives in Firms." *Journal of Economic Behavior and Organization* 26.: 145-60.

Gerhardt, B., H.B. Minkhoff and R.N. Olsen. 1995. "Employee Compensation: Theory, Practice and Evidence."In G.R. Ferris, S. D Rosen and D.T. Barnum (Eds.), *Handbook of Human Resource Management*: 528-47, Cambridge, MA: Blackwell.

Gibbons, Robert and Kevin J. Murphy. 1990. "Relative Performance Evaluation for Chief

Executive Officers." *Industrial Labor Relations* 43:30-52.

Hart, Oliver. 2001. "Norms and the Theory of the Firm." *University of Pennsylvania Law Review*149(6). 1701-15.

Holmström, Bengt. 1982. "Moral Hazard in Teams." *Bell Journal of Economics* 13: 324-40.

Holmström, Bengt and Paul Milgrom. 1990. "Regulating Trade Among Agents."*Journal of Institutional and Theoretical Economics* 146: 85-105.

Itoh, Hideshi. 1992. "Cooperation in Hierarchical Organizations: An Incentive Perspective." *Journal of Law Economics and Organization* 98: 321-325.

Itoh, Hideshi. 1993. "Coalition Incentives and Risk Sharing," *Journal of Economic Theory* 60: 410-27.

Kandel, Eugene and Edward P. Lazear. 1992. "Peer Pressure and Partnerships." *Journal of Political Economy* 100: 801-817.

Lazear, Edward P. and Sherwin Rosen. 1981. "Rank-Order Tournaments as Optimum Labor Contracts."*Journal of Political Economy* 89: 841-64.

Lazear, Edward P. 1989. "Pay Equality and Industrial Politics." *Journal of Political Economy* 97: 561-80.

Lazear, Edward P. 1998. *Personnel Economics.* The MIT Press

Levin, Jonathan. 2002. "Multilateral Contracting and the Employment Relationship."*The Quaterly Journal of Economics* 117: 1075-1103

Levin, Jonathan. 2003. "Relational Incentive Contracts." *American Economic Review 93: 835-57.*

Macho-Stadler, Ines and J. David Perez-Catrillo. 1995. "Moral Hazard with Several Agents: The Gains from Cooperation." *International Journal of Industrial Organization* 11: 73-100.

MacLeod, Bentley and James Malcomson. 1989. "Implicit contracts, Incentive Compatibility, and Involuntary Unemployment." *Econometrica 57: 447-80.*

Macneil, Ian. 1978. "Contracts: Adjustments of Long-Term Economic Relations under Classical, Neoclassical, and Relational Contract Law." *Nortwestern University Law Review* 252: 854-906

Mookherjee, Dilip 1984. Optimal Incentive Schemes with Many Agents." *Review of Economic Studies* 51: 433-46.

Prendergast, Canice. 1999. "The Provisions of Incentives in Firms," *Journal of Economic Literature* 37: 7-63.

Radner, Roy. 1986. "Repeated Partnership Games with Imperfect Monitoring and No Discounting." *Review of Economic Studies* 43: 43-51.

Schmidt, Klaus M. and Monika Schnitzer. 1995. "The interaction of Explicit and Implicit contracts." *Economics letters* 48: 193-99.

Treble, John, Edwin van Gameran, Sarah Bridges, Tim Barmby. 2001. "The internal Economics of the Firm: Further Evidence from Personnel Data" *Labor Economics* 8: 531-552.

Essay

IV

Norms Matter

Ola Kvaløy*

Norwegian School of Economics and Business Administration

Abstract: *Even though norms have been integrated in the formal theory of the firm, we have not seen a clear-cut relationship between norms and firm boundaries. This paper shows how norms actually can determine firm boundaries. In a repeated trust-game with symmetric information and endogenous verifiability of actions, I show why non-integration is expected to dominate in low-trust environments. In a static trust-game with asymmetric information and tit-for-tat types, I show why assets should be allocated to those with the best reputation of being trustworthy.*

* I thank Frøystein Gjesdal, Petter Osmundsen and especially Trond E. Olsen for valuable comments and suggestions.

1. Introduction

The role of norms has become an important issue in the formal theory of the firm. Norms such as trust and trustworthiness help explain transactions both between and within firms. A norm can be defined as "...a rule that is neither promulgated by an official source, such as a court or legislator, nor enforced by the threat of legal sanctions, yet it is regularly complied with..." (Posner, 1997). This in contrast to a law, which is a rule that is enforced legally. Even though norms cannot be legally enforced, rational agents obey norms if obedience confers private benefits, and reciprocally behaving agents obey norms if others obey them.

Some rules must rest upon both law and norms to be enforced. Rules formulated in a contract are the most obvious. There are certain features of a contract that can be legally enforced, but most contracts are incomplete in the sense that there are elements that cannot be verified, and there are contingencies that cannot be foreseen. Still, contracts are adhered to, even when the possibility of legal enforcement is small. Contracts are in these instances enforced by norms of trust and trustworthiness.

There are two main approaches to formalize trust. The most commonly used is based on the framework of infinitely repeated games. If agents transact repeatedly, they will honour trust to ensure future cooperative transactions. This approach makes trust a feasible option for rational agents. Honouring trust is a way to invest in a valuable reputation. A second approach is based on asymmetric information and reciprocity. If agents regard it as a possibility that other agents act reciprocally, trust and cooperation can be attained in a one shot transaction. [1]

Formalization of trust has added a great deal to our understanding of relational contracts, dynamic incentives and so forth, but as Hart (2001) demonstrates, it is hard to find clear-cut relationship between trust and optimal allocation of asset ownership. In the repeated game approach, where the discount factor is a proxy for trust, a move from a "low-trust"

---

[1] On reciprocity, see Fehr and Gachter (2000) for an overview. On trust in finitely repeated games with asymmetric information, see Kreps et. al, (1982).

environment to a "high-trust" environment only reveals that cooperation in general is more likely to occur both between integrated parties and non-integrated parties.[2]

In this paper I will show that the level of trust and the trust-structure in a relationship can determine optimal asset allocation. In Section 2 I take a repeated game approach, analysing a game between a buyer and a seller of inputs. I do not make the standard assumption that contract verifiability is exogenously given; instead I introduce endogenous probability of contract verification.[3] The more the parties invest in contract specifications, the higher is the probability of a third party (the court) verifying and thus legally enforcing the contract. If the discount factor is sufficiently low, transactions will only be implemented if the possibility for contract verification is sufficiently high. By introducing endogenous probability of legal contract enforcement, we get hold of a trade-off between 'direct transactions costs', first discussed by Coase (1937), and "strategic transaction costs", introduced by Klein, Crawford and Alchain (1978), Williamson (e.g. 1979, 1985), Grossman and Hart (1986) and Hart and Moore (1990): the lower the degree of relationship-specific under investment, the higher is the necessary investment in contract specification.

This substitutability between direct and strategic transaction costs makes it possible to study optimal asset allocation focusing on institutional differences in possibilities of legal contract enforcement. Rock and Wachter (2001) argue that the incentives to invest in a detailed contract are weaker when parties are under the same ownership than when parties are under separate ownership. The governance systems meant to protect the integrity of a transaction are very different within firms than between firms. Within firms the residual claimant decides the outcome of any contract breach. Two subdivisions of a firm will be hesitant to write a strict formal contract since the power to ensure the implementation of it is weak. Corporate law gives shareholders within certain limits the right to exercise this power, and legal courts are not intended to intervene in intra-firm disputes. Transactions between firms, however, are

[2] For instance, in dynamic versions of property rights models (see Baker, Gibbons and Murphy, 2002; Halonen, 2002) there are not clear-cut relationships between discount factors and optimal allocation of assets. (The property rights theory was developed by Grossman and Hart, 1986; and Hart and Moore, 1990).

[3] So called Costly State Verification Models (CSV) have focused on contract design problems where verification is an issue. But the CSV approach considers verification of exogenous state variables, not endogenous effort variables as here. See Krasa and Villamil (2000) for a generalization of the CSV models. Ishiguro (2002) analysis endogenous verifiability of endogenous effort variables in a static principal agent model, but do not consider the effect of endogenous verifiability in self-enforced relational contracts, as in this paper. Kvaløy (2003) develops this point.

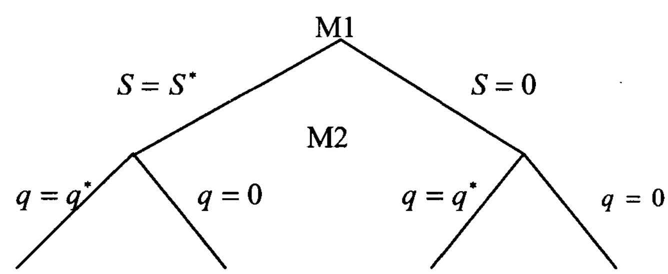
technically protected by legal courts. The problem is, of course, that it is hard to verify complex contracts. But parties who enjoy separate ownership are nevertheless more interested in specifying a detailed contract, since judges enforce contracts according to their terms, while shareholders exercise power in order to maximize shareholder value. Hence, the marginal effect of increased contract specifications on the probability of verifying and thus legally enforcing a contract is thought to be stronger between than within firms. In repeated relationships with contracts relying both on legal enforcement and self-enforcement, this institutional difference in possibilities for legal contract enforcement implies that non-integration dominates in low trust environments: if the discount factor is sufficiently low, the contract cost necessary to implement a contract is higher when the parties are integrated, than when the parties are non-integrated

In Section 3 I analyse the static version of the same game, but now with non-verifiable actions, reciprocal agents and asymmetric information. I show that if a trust game is played only once, the party with the best reputation of being trustworthy should own the asset. As residual claimant, the asset owner has the power to either honour or abuse trust, when contracts cannot be verified. If the asset owner is considered trustworthy, it relaxes the trusting party's participation constraint, and increases the surplus from the transaction. It is a well-known insight from the theory of reputations (see Kreps 1990, Tadelis, 1999, 2002) that if an agent gains a reputation of acting trustworthy and fair, it will be easier for other agents to transact with him. But this theory assumes rational agents and focuses on how reputation concerns can provide incentives for short-lived agents, and thus how a good reputation can become a tradable asset. The novelty in this paper is to demonstrate a clear-cut relationship between reputation and optimal asset allocation. Experimental evidence suggests that a good reputation increases the level of relationship specific investments even in one-shot transactions. Employers make generous job offers if they expect workers to behave in a reciprocal manner. Similarly, workers supply high effort if they expect employers to respond reciprocally (see Fehr, Gachter and Kirchsteiger, 1997). By applying the concept of sequential reciprocal equilibrium (SRE), developed by Dufwenberg and Kirchsteiger (2003), I show that if reciprocal (tit-for-tat) behaviour is the norm, then the parties' reputation of obeying this norm is decisive for optimal asset allocation, since the asset owner's reputation decides the surplus from the transaction.

2. Non-integration dominates in low-trust environments

Consider two managers, M1 and M2, where M2 uses an asset to produce a good that can be used in M1's production process. M1 pays M2 a salary *S*, in return M2 delivers a quality q. Payoff to M1 is $Y(q) - S$ where $Y(q)$ is the good's value added, and $Y'(q) > 0, Y''(q) \leq 0$. Payoff to M2 is $S - C(q)$ where $C(q)$ is the cost of producing quality q, and $C'(q) > 0, C''(q) > 0$. Total surplus from the transaction is then $\Pi(q) = Y(q) - C(q)$.

M1 and M2 agree upon a compensation S^* and a quality q^*. We can think of this relationship as a trust game:



M1's payoff:	$Y(q^*) - S^*$	$-S^*$	$Y(q^*)$	0
M2's payoff:	$S^* - C(q^*)$	S^*	$-C(q^*)$	0

Figure 1

M1 pays M2 a salary S^* ex ante, trusting that M2 will deliver on the agreed upon quality q^*.[4] M2 can honour that trust by delivering q^*, or he can abuse trust by not delivering a value adding quality, that is $q = 0$.[5] Assume first that the contract specifying the content of q^* is non-verifiable. Played only once, the Nash equilibrium is $[S = 0, q = 0]$, hence no transaction takes place. If the game is played repeatedly, and the discount factor is sufficiently high, then

[4] The trust structure in figure 1 is not unrealistic. Consider for instance a researcher doing a research report for an external company. It is not unusual that the researcher gets paid for the number of working hours reported. First when the researcher has got his payment, he delivers his report. Due to imperfect monitoring, the researcher can easily misinform number of hours worked, and the buyer risk paying for an invaluable report.

[5] Trust-abuse is equivalent to any $q \neq q^*$, but if M2 is to abuse trust we might as well model this as if he is able to play $q = 0$

trigger strategies constitute a subgame perfect Nash equilibrium where $\lfloor S = S^*, q = q^* \rfloor$ is played in every stage game.

In the repeated game, M2 will honour trust as long as $\frac{1}{1-\delta}(S - C(q)) \geq S$ where δ is the discount factor. That is

(1) $\delta \geq \frac{C(q)}{S}$, where $Y(q) \geq S \geq C(q)$.

If (1) holds, then transactions will take place, since M2 has incentives to honour trust, and M1 can anticipate this. Assume that M1 has the bargaining power to decide S. M1 then solves

$\underset{q,S}{Max}[Y(q) - S]$, subject to (1),

which yields the first order condition $Y'(q) = \frac{C'(q)}{\delta}$. Observe that first-best quality level, q^{FB}, is not achievable for discount factors lower than one. Due to the convex cost function, the parties will underinvest in quality, that is $q < q^{FB}$ when $\delta < 1$. We see that the lower the discount factor the lower the quality level. In other words: quality is a negative function of trust level.

Assume now that the parties to a certain extent can rely on a third party (the court) to verify *q*. Let k denote the level of contract costs, where k includes the cost of all activities, excluding $C(q)$, necessary to implement the transaction. Let *v(k)* denote the probability that the true content of q^* will be verified in a court of law in case of contract breach, where $v(0) = 0$, $v'(k) > 0$, and $v''(k) < 0$. Verification implies legal enforcement. If the parties can write verifiable contracts specifying large penalties in case of contract breach, first-best is easily implemented. It is reasonable to assume, however, that contracted penalties are denied by the court if they are sufficiently large relative to the actual loss from contract breach. I follow Ishiguro (2002) in assuming that if a contract breach is verified, then the court applies expected damages as a breach remedy, so that the victim of a breach is as well off as if the contract were fully performed. Hence, verification implies that M2 is legally forced to deliver q^* to the cost of $C(q^*)$.

M2 will now honour trust if $\frac{1}{1-\delta}(S - C(q)) \geq S - v(k)C(q)$. Hence, transactions take place if

(2) $$\delta \geq \frac{C(q)(1-v(k))}{S-v(k)C(q)}.$$

M1 now solves

(3) $$\underset{q,k,S}{Max}\left[Y(q) - S - k\right] \text{ subject to (2).}$$

This yields the first order conditions

(4) $$Y'(q) = \frac{C'(q)}{\delta}(1-(1-\delta)v(k))$$

(5) $$\frac{C(q)}{\delta}(1-\delta)v'(k) \leq 1, \quad k \geq 0,$$

where (5) holds with complementary slackness. We observe that $q < q^{FB}$ when $v(k) < 1$ and $\delta < 1$. Observe that for $v(k) > 0$ and $v'(k)$ sufficiently high (in particular $v'(0)$ sufficiently high), M1 will choose $k > 0$ so that the quality level can be increased relatively to the situation where $v(k) = 0, \forall k$. The first order conditions elucidate a classic relationship between trust and transaction costs. Reduced trust level (reduced discount factor), reduces quality, ceteris paribus, but this can be compensated by increased investments in contract specifications. And correspondingly, the more costly it gets to write verifiable contracts (lower $v'(k), \forall k$), the lower the quality level, ceteris paribus, but this can be compensated if the parties are able to move to a higher trust-environment.

Interestingly, comparative static shows that the chosen quality level that the parties choose is not necessarily a positive function of the relationship's trust-level. If $v'(k)$ is not too large, however, the most intuitive result always holds: The higher trust level, the higher quality level, hence the less severe is the problem of relationship specific under investment, that is $\frac{\partial q}{\partial \delta} > 0$. But the analysis indicates that high levels of δ together with high levels of $v'(k)$ may yield $\frac{\partial q}{\partial \delta} < 0$ (see appendix). In this situation, higher trust level reduces quality and thus increases the problem of relationship specific under investments. Hence, a high trust level can prevent the agents from making profitable investments in a detailed verifiable contract, since they rather rely on relational contracting. The result is an interesting complement to Baker, Gibbons and Murphy (1994). They show that the presence of explicit (verifiable) contracts can make relational contracts infeasible. The model in this paper indicates that the feasibility of proper relational contracts can lead to an inefficient level of explicit detailed contracting.

We see that the model reveals a simple trade-off between quality investment and contract costs. The higher specified quality level, the higher is the contract costs necessary to enforce the transaction. We could say that the direct contract costs, k, corresponds to Coase's concept of transactions costs, while the efficiency-loss of not being able to implement first-best, $\Pi(q^{FB}) - \Pi(q^{*})$, corresponds to the strategic transaction costs discussed in different forms by Williamson et. al. Coase discusses the cost of discovering relevant prices, and the costs of "negotiating and concluding a separate contract for each exchange transaction...," but are virtually silent about strategic transaction costs. Williamson et. al. discuss problems of opportunism and under investment, but say little about direct costs of contracting. By introducing endogenous probability of legal contract enforcement, we get hold of a trade-off between direct transactions costs and strategic transactions costs:

Proposition 1: *Direct 'Coasian' transactions costs, k, and strategic 'Williamsonian' transactions costs, $\Pi(q^{FB}) - \Pi(q^{*})$, are substitutes.*

In principle, (3) can be solved for exact values of q and k. But in a competitive market, it is reasonable to believe that there exists a minimum necessary quality level. Assume now that M1 sells Y in a market that requires $q \geq \tilde{q}$. Moreover, assume that $\tilde{q}$ exceeds the quality level that solves (3). M1 must then solve (3) for $q = \tilde{q}$, that is

$$\underset{k,S}{Max}\left[Y(\tilde{q}) - S - k\right], \text{ subject to } \delta \geq \frac{C(\tilde{q})(1-v(k))}{S-v(k)C(\tilde{q})},$$

which yields

(6) $\frac{C(\tilde{q})}{\delta}(1-\delta)v'(k) \leq 1, \quad k \geq 0.$

We can then discuss the optimal contract cost, *k*, given quality level, $\tilde{q}$. If $\frac{C(\tilde{q})}{\delta}(1-\delta)v'(k) < 1$ then $k = 0$. This occurs if v'(0) is sufficiently low, or the discount factor is sufficiently high. But if $\frac{C(\tilde{q})}{\delta}(1-\delta)v'(k) = 1$ then M1 have incentives to invest $k > 0$ in specifying the content of $\tilde{q}$.

Consider now two types of ownership structure: a) Integration: M1 and M2 are agents at the same firm. A third party - the owner of the firm - has residual control right over physical assets, and is therefore supposed to solve any transactional disagreements between the agents. b) Non-integration: M2 is a supplier outside the control of the firm owning M1. In case of disagreements, the parties must rely on the court as a neutral third-party. As argued, it is reasonable to believe that the marginal effect of investment in contract specifications on the probability of verifying and legally enforcing the contract is greater when M1 and M2 enjoy separate ownership, than if M2 and M1 are agents at the same firm. That is $v_{NI}'(k) > v_I'(k)$ where subscript *NI* denotes non-integration and *I* denotes integration. Note that it need not be more complicated to specify the content of q in integrated solutions than in non-integrated solutions. But since the owner of the integrated company is prone to interfere in disputes between M1 and M2, the marginal effect of contract specifications on the probability of legal enforcement is reduced by the probability that the owner instead of the court will solve contractual disputes.

In order to insulate the effect of differences in verifiability, I consider the case where bargaining positions are independent from organizational form.[6] Then, from (6) we have

Proposition 2: *Assume* $v_{NI}'(k) > v_I'(k)$. *Let* $k_i, i = NI, I$ *denote the optimal contract cost, given quality level* $\tilde{q}$. *If* $\frac{C(\tilde{q})}{\delta}(1-\delta)v_i'(k) = 1$, $i = NI, I$ *then* $k_I > k_{NI} > 0$.

[6] Recall that bargaining power does not necessarily follow residual control rights. Even if M1 do not have residual control rights in any of the organizational forms discussed above, he can still have the bargaining power to decide level of S, for instance if he is a monopsonist.

Proposition 2 says that if the discount factor is sufficiently low, so that $k_i > 0$, then the contract cost is higher when the parties are integrated, than when the parties are non-integrated

Moreover, we know that if $Y(\tilde{q}) < \frac{C(\tilde{q})}{\delta}$, then $k > 0$. But if contract investments are not sufficiently profitable ($v'(k)$ too low), than the parties will rather cancel the whole transaction than invest in contract specifications. Hence, we obtain

Proposition 3: If $Y(\tilde{q}) < \frac{C(\tilde{q})}{\delta}$, $\frac{C(\tilde{q})}{\delta}(1-\delta)v_I{}'(0) < 1$, *and* $\frac{C(\tilde{q})}{\delta}(1-\delta)v_{NI}{}'(0) = 1$, *then transactions will only take place under non-integration*

Proposition 2 and 3 suggest that non-integration dominates in low-trust environments. If the discount factor is sufficiently low, the firm have incentives to outsource agents either because it is the only way to make transactions feasible, or because it reduces the transaction costs. This result is deduced from an economic environment where both legal enforcement capabilities and the level of trust matters. It demonstrates the importance of understanding the contract enforcement mechanisms[7] of various governance structures in order to determine optimal asset allocation.

3. The party with the best reputation should own the asset

Consider the same trust-game as in the previous section. Assume no third party owner, so that either M1 or M2 owns the asset. Assume that ownership of the asset conveys ownership of the good produced.[8] Moreover, assume here that the asset owner has the bargaining power to decide S. If M1 owns the asset, then M2 has the first move, allowing M1 to refuse paying the compensation S^* ex post the production of q^*. If M2 owns the asset, M1 has the first move, allowing M2 to receive S^* without producing q^*. The set up in figure 1 is thus one in which M2 owns the asset. Let us now assume reciprocal agents: the parties play tit for tat strategies, so that if the first mover trusts the asset owner, trust is honoured. Assume asymmetric

[7] See Torsvik (2000) for a discussion on the importance of contract enforcement mechanisms.
[8] This is not uncommon in the literature. See e.g. Baker, Gibbons and Murphy (2002).

information in the sense that the party not owning the asset cannot be certain whether the asset-owner is a tit-for-tat type. If M1 is the asset owner, M2 sees a possibility p_1 that M1 plays tit for tat. If M2 is the asset owner, M1 sees a possibility p_2 that M2 plays tit for tat. Hence, M*i* will as a first mover maximizes his profit given his beliefs about M*j*, while as a second mover honour trust if he is trusted. These strategies constitute a sequential reciprocal equilibrium (SRE) as defined by Dufwenberg and Kirchsteiger (2003).[9]

Assume that p_1 and p_2 are known to both parties. We can interpret these parameters as the parties' reputation for being trustworthy. A cooperative tit-for-tat strategy is considered to represent a fair, reciprocal behaviour. The greater p, the better is the reputation of fairness. To keep it simple, I here abstract from any possibility of verifying the contract, that is $v(k) = 0, \forall k$.

Let us first look at the case where M2 owns the asset. M1 will now enter the relationship if $p_2(Y(q) - S) + (1 - p_2)(-S) \geq 0$. That is

(7) $$Y(q) \geq \tfrac{1}{p_2} S.$$

One may now consider (7) as a constraint in M2's maximization problem:

$$\underset{q,S}{Max}[S - C(q)] \quad \text{subject to (7).}$$

This yields the first order condition

(8) $$Y'(q) = \tfrac{C'(q)}{p_2}.$$

We see that $p_2 < 1$ leads to under investment in quality: the participation constraint of M1 (given by (7)) makes it impossible to implement $q = q^{FB}$. The parties will thus agree upon a

[9] An SRE is a strategy profile such that at a given stage, each player makes choices that maximize his utility given his beliefs, and given that he follows his initial strategy, where his initial strategy is determined by his sensitivity towards reciprocity. In this paper the players are fully sensitive towards reciprocity, implying that if M*i* trusts M*j*, M*j* honours trust always.

quality level $q < q^{FB}$.[10] Compared to first-best we see that this efficiency loss is greater the lower p_2. If M2 has a bad reputation, it would be impossible for M1 to enter the relationship unless the size of the transaction (the levels of q and S) is modest.

Let us now turn to the case where M1 owns the asset. The trust game is then given in figure 2:

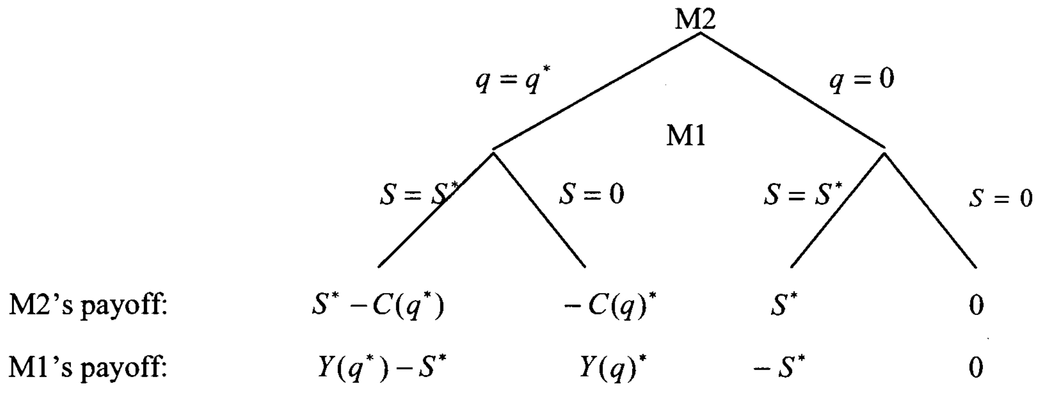


Figure 2

Now M2 enters the relationship if $p_1(S - C(q)) + (1 - p_1)(-C(q)) \geq 0$. That is

(9) $\quad S \geq \frac{C(q)}{p_1}$.

M1 now solves

$\underset{q,S}{Max}[Y(q) - S]$ subject to (9),

which yields

(10) $\quad Y'(q) = \frac{C'(q)}{p_1}$.

---

[10] Under investment is due to the assumptions of a convex cost function and a non-convex product function.

Now, $p_1 < 1$ leads to under investment in quality: the quality level must be reduced relative to first-best in order to make M2 enter the relationship.

Optimal asset allocation is now straightforward. Since social surplus from the transaction is given by $\Pi = Y(q^*(p_i)) - C(q^*(p_i)), (i = 1,2)$ where $\frac{\partial \Pi}{\partial p_i} > 0$, we obtain:

Proposition 4: *If* $p_1 > p_2$, *then it is optimal to let M1 own the asset. If* $p_1 < p_2$, *then it is optimal to let M2 own the asset.*

Proposition 4 suggests that the party with the best reputation of being trustworthy should own the asset. A difficulty with this approach is, of course, that it is hard to measure p_i, and that the players cannot easily communicate these estimations. On the other hand, p_i is likely to be a function of M*i*'s dependence of a good reputation. For instance, the players can operationalize p_i on the basis of the parties' number of alternative trading partners.

4. Conclusion

The purpose of this paper has been to show that norms should play a role in models of optimal asset allocation. In both the repeated game approach, and the reciprocal asymmetric information approach, norms matter. In the repeated game approach, it is shown that non-integration is expected to dominate in low trust environments. In the asymmetric information approach, it is shown that the party with the best reputation of being trustworthy should own the asset.

The repeated game approach in Section 2 shows a trade-off between direct transaction costs and strategic transaction costs. For a given level of quality on the transacted good, the efficient asset allocation is a way to reduce the direct cost of contracting that quality (transaction costs in Coases' sense). The asymmetric information approach in Section 3 abstract from direct transaction costs. There, asset allocation is a way to reduce the degree of relationship specific under investment (strategic transaction costs in Williamson et. al's sense). An issue to examine further is how norms affect the dominance of direct transaction costs versus strategic transaction costs. For instances; do inefficiencies from Coasian transaction costs dominate in low-trust environments with rational agents, while the problem of inefficient relationship-specific investments dominates in reciprocal corporate cultures with asymmetric information?

APPENDIX

Determining the sign of $\frac{\partial q}{\partial \delta}$ *:*

When (2) binds, we can insert for S in (3) and express the problem in the following form:

$$\underset{q,k}{Max}\, H(q,k,\delta) = Y(q) - C(q)\frac{1-(1-\delta)v(k)}{\delta} - k$$

The first order conditions (4) and (5) can then be written

$$H_q(q,k,\delta) = 0$$
$$H_k(q,k,\delta) = 0$$

where subscript denotes partial derivatives.

To find maximum, the second order (concavity) conditions must be satisfied (for simplicity I exclude functional arguments):

$$H_{qq} < 0$$
$$H_{kk} < 0$$
$$H_{qq}H_{kk} - (H_{qk})^2 > 0$$

By differentiating the first order conditions with respect to δ, we obtain the equation system

$$H_{qq}\frac{\partial q}{\partial \delta}+H_{qk}\frac{\partial k}{\partial \delta}=-H_{q\delta}$$
$$H_{qk}\frac{\partial q}{\partial \delta}+H_{kk}\frac{\partial k}{\partial \delta}=-H_{k\delta}$$

From Cramer's rule we can now solve for $\frac{\partial q}{\partial \delta}$:

(A.1) $$\frac{\partial q}{\partial \delta}=\frac{-H_{q\delta}H_{kk}+H_{k\delta}H_{qk}}{H_{qq}H_{kk}-(H_{qk})^2}$$

We have

$$H_{qq}=Y''(q)-C''(q)\tfrac{1-(1-\delta)v(k)}{\delta}<0$$
$$H_{kk}=C(q)\tfrac{(1-\delta)v''(k)}{\delta}<0$$
$$H_{qk}=C'(q)\tfrac{(1-\delta)v'(k)}{\delta}>0$$
$$H_{q\delta}=C'(q)\tfrac{1-v(k)}{\delta^2}>0$$
$$H_{k\delta}=C(q)\tfrac{1-v'(k)}{\delta^2}$$

For the second order conditions to be fulfilled, the denominator in (A.1) must be positive. Hence for $\frac{\partial q}{\partial \delta}>0$, we must have $-H_{q\delta}H_{kk}+H_{k\delta}H_{qk}>0$. For $v'(k)<1$, which implies $H_{k\delta}>0$, this is always true.

From the first order condition (5) we have

$$v'(k)=\tfrac{\delta}{1-\delta}\tfrac{1}{C(q)}$$

We see that in optimum $v'(k)>1$ for higher levels of δ. Hence, for high levels of δ and $v'(k)$, the numerator in (A.1) may turn negative, which implies, given that the second order conditions are satisfied, that $\frac{\partial q}{\partial \delta}<0$.

REFERENCES

Baker, George, Robert Gibbons and Kevin J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts." *Quarterly Journal of Economics* 109: 1125-56.

Baker, George, Robert Gibbons and Kevin J. Murphy. 2002. "Relational Contracts and the Theory of the Firm." *Quarterly Journal of Economics* 117: 39-94.

Coase, Ronald H. 1937. "The Nature of the Firm." *Economica* 4: 386-405.

Dufwenberg, Martin and George Kirchsteiger. 2003. "A Theory of Sequential Reciprocity." Forthcoming in *Games and Economic Behavior.*

Fehr, Ernst, Simon Gachter and Georg Kirchsteiger.1997. "Reciprocity as a Contract Enforcement Device: Experimental Evidence" *Econometrica* 65: 833-860.

Fehr, Ernst and Gacther, Simon. 2000. "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economics Perspective* 14: 159-81

Grossman, Sanford and Oliver Hart. 1986 "The Costs and Benefits of Ownership: A theory of Lateral and Vertical Integration." *Journal of Political Economy* 94: 691-719.

Halonen, Maija. 2002. "Reputation and Allocation of Ownership." *The Economic Journal.* 112: 539-58.

Hart, Oliver and John Moore. 1990. "Property Rights and the Nature of the Firm." *Journal of Political Economy* 98: 1119 – 58.

Hart, Oliver. 2001. "Norms and the Theory of the Firm." *University of Pennsylvania Law Review* 149: 1701-15.

Ishiguro, Shingo. 2002. "Endogenous Verifiability and Optimality in Agency". *Journal of Economic Theory* 105: 518-530.

Klein, Benjamin, Crawford Robert G. and Alchian, Armen A. 1978. "Vertical Integration, Appropriable Rents, and the Competitive contracting Process." *Journal of Law and Economics* 21: 297-326.

Krasa, S. and A.P. Villamil. 2000. "Optimal contracts when enforcement is a decision Variable." *Econometrica* 68: 119-134.

Kreps, David M., Paul Milgrom, John Roberts and Robert Wilson. 1982. "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma." *Journal of Economic Theory* 27: 245-52.

Kreps, David M., 1990. "Corporate Culture and Economic Theory." In J. Alt and K.

Shepsle, eds. *Perspectives on Positive Political Economy.* Cambridge University Press.

Kvaløy, Ola. 2003. "Endogenous Verifiability in Relational Contracts." *Unpublished manuscript.*

Posner, R.A. 1997. "Social norms and the Law: An Economic Approach." *American Economic Review.* 87: 365-9.

Rock, E. and M. Wachter. 2001. "Islands of Conscious Power: Law, Norms and the Self Governing Corporation." *University of Pennsylvania Law Review.* 149: 1619-77.

Tadelis, Steven. 1999. "What's in a Name? Reputation as a Tradeable Asset." *American Economic Review* 89:548-63.

Tadelis, Steven. 2002. "The Market for Reputations as an Incentive Mechanism" *Journal of Political Economy* 92:854-82.

Torsvik, Gaute. 2000. "Social Capital and Economic Development; A Plea for the Mechanisms." *Rationality and Society.* 12: 451-76.

Williamson, O. E. 1979."Transaction-Cost Economics: The Governance of Contractual Relations." *Journal of Law and Economics* 22: 233-61.

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism*. New York: The Free Press.