

## The interaction of explicit and implicit contracts

Klaus M. Schmidt\*, Monika Schnitzer

*Wirtschaftspolitische Abteilung, University of Bonn, Adenauerallee 24, Bonn, D-53113, Germany*

Received 2 August 1994; accepted 18 August 1994

---

### Abstract

We analyze explicit and implicit contracts in a repeated principal–agent model with observable but only partially contractable actions of the agent. It is shown that the set of implementable actions may increase or decrease if additional actions become contractable.

*Keywords:* Implicit contracts; Repeated games

*JEL classification:* L14; C72; J33

---

### 1. Introduction

This paper analyses a repeated principal–agent relationship where all actions of the agent are perfectly observed by both parties but only some of them can be verified by the courts. This generalizes the work of MacLeod and Malcomson (1989) who analyzed implicit contracts in a set-up where all actions of the agent are non-verifiable. By characterizing the set of all actions that can be implemented with implicit and/or explicit contracts, we derive two main results. First, we show that the possibility of writing an explicit contract on some of the agent's actions may render an implicit contract on these actions infeasible. Thus, an explicit contract has to be written even if it is costly to do so. Second, the set of actions that can be implemented may increase or decrease if it becomes possible to govern some of the agent's actions by an explicit contract. This implies that even if there is no cost to writing an explicit contract, the mere possibility of explicit contracts may be harmful because it may make it impossible to implement a desired action.

There is a large and growing literature on implicit contracts in repeated principal–agent models with applications to labor economics, the theory of the firm, the sovereign debt

\* Corresponding author.

problem, etc.<sup>1</sup> Very little attention has been paid, however, to the interaction of implicit and explicit contracts. A notable exception is a recent paper by Baker et al. (1994). They consider a repeated principal–agent relationship where in every period the agent takes only one action which is not observed by the principal. The action generates two noisy signals, one of which is ‘objective’, i.e. it can be verified by the courts and contracted upon explicitly, while the other one is ‘subjective’ and not verifiable. The authors show that if the objective signal is sufficiently close to perfect, then an implicit contract on the subjective signal is not feasible. They also show that explicit and implicit contracts may be complements. These results, which are derived for a different class of models, are consistent with, and complementary to, ours.

## 2. Implicit contracts

Consider a principal and an agent, both of whom are infinitely lived and engaged in a repeated relationship in a stationary environment. To fix ideas, think of the principal as a firm and of the agent as a worker, but other interpretations are possible (e.g. a buyer and a seller, the government and a regulated firm, etc.) In every period the agent chooses an action  $x$  out of some (possibly multi-dimensional) Euclidean space  $X$ , and the principal makes a monetary payment,  $m \in \mathbb{R}_0^+$ , to the agent. We will interpret  $x$  as a vector of different activities below. The payoffs in each period are given by  $U(x, m) = m - \psi(x)$  for the agent and  $V(x, m) = \pi(x) - m$  for the principal, where  $\psi(x)$  denotes the agent’s effort cost of taking action  $x$  and  $\pi(x)$  is the principal’s profit (net of all non-labor costs). If the agent does not work for the principal, both parties get their non-negative outside option utilities  $U_0$  and  $V_0$ , respectively. Thus, the total surplus generated by action  $x$  is given by  $S(x) = \pi(x) - \psi(x) - U_0 - V_0$ . In the repeated game, both players maximize the discounted sum of their stage game payoffs, where  $\delta < 1$  denotes the common discount factor. The following assumption says that there are gains from trade, but not if the agent takes his most preferred action.

*Assumption 1.* The functions  $\pi(x)$  and  $\psi(x)$  are bounded and  $\psi(x)$  is non-negative. There exists an  $x \in X$  such that  $S(x) > 0$ . Let  $\underline{x} \in \operatorname{argmin}_{x \in X} \psi(x)$ . Then  $\psi(\underline{x}) = 0$  and  $\pi(\underline{x}) < 0$ .

The timing of the game is as follows. In the beginning of each period the principal decides on whether to hire or to fire the agent, and the agent decides whether to work for the principal or quit. If the agent is employed he chooses  $x$ , which is observed by both parties. Thereafter the principal makes a payment  $m$  to the agent.

MacLeod and Malcomson (1989) analyzed this principal–agent model under the assumption that the parties cannot write an explicit contract on  $x$  because this variable cannot be verified by the courts. In this case, an explicit contract can specify only a fixed wage  $w$ . Given a fixed wage payment the agent will always choose  $\underline{x}$ , so there are no gains from trade. However, the parties could agree on an implicit (or self-enforcing) contract. Without loss of generality we

<sup>1</sup> See, for example, Bull (1987), MacLeod and Malcomson (1989), Kreps (1990), and Eaton (1993).

can restrict attention to the following class of stationary implicit contracts.<sup>2</sup> The parties agree that the agent takes some action  $\hat{x}$  in every period, and that the principal pays a bonus  $\hat{b} \geq 0$  at the end of each period if the agent has taken  $\hat{x}$ . In addition, the principal pays a fixed wage  $\hat{w} \geq 0$ , which is specified in an explicit contract and can be enforced by the courts. Thus, under the terms of the agreement,  $m = \hat{b} + \hat{w}$ . If any party deviates, the firm will fire the agent and both parties receive their outside option utilities forever thereafter.<sup>3</sup> We say that an implicit contract *implements* the stationary allocation  $(\hat{x}, \hat{b}, \hat{w})$  if the strategies of the two parties prescribed by the implicit contract are part of a subgame perfect equilibrium, i.e. if sticking to the terms of the agreement is a mutually best response after history of the game and the contract is self-enforcing.

The following proposition, which is based on propositions 1 to 3 of MacLeod and Malcomson (1993), characterizes the set of stationary allocations that can be implemented with implicit contracts.

*Proposition 1 (MacLeod–Malcomson).*

(a) *There exists an implicit contract which induces the agent to take  $\hat{x} \in X$  in every period if and only if*

$$\psi(\hat{x}) - \psi(\underline{x}) \leq \frac{\delta}{1 - \delta} S(\hat{x}). \quad (1)$$

(b) *If  $\hat{x}$  can be implemented with an implicit contract, then this can be done by giving the agent any share  $\alpha \in [0, 1]$  of the total surplus  $S(\hat{x})$ .*

*Proof.* See MacLeod and Malcomson (1989).

### 3. Costly explicit contracts

The implicit agreement on the action  $\hat{x}$  and the bonus payment  $\hat{b}$  is enforced by the mutual threat to terminate the relationship as soon as one party deviates. This threat is credible if the principle believes that after a deviation the worker will always choose the least costly action  $\underline{x}$  and the agent believes that the principal will never pay a bonus again, i.e. if both parties lose all “trust” in the relationship.

Suppose, now, that  $x$  is verifiable and that it is possible to (costlessly) write an enforceable,

<sup>2</sup> MacLeod and Malcomson (1989) consider more general implicit contracts, in particular they allow for non-stationary environments, non-stationary contracts, and for the possibility of severance pay and bonding if the relationship is terminated. Their proposition 3 shows that if a stationary allocation can be implemented with an implicit contract, then it can be implemented with a contract of the class we consider.

<sup>3</sup> Note that if the employment relation is terminated both parties get their minmax payoffs. Thus, termination is the worst possible punishment, and it is easy to show that termination is, indeed, a subgame perfect equilibrium. Abreu (1988) has shown that the set of *all* subgame perfect equilibrium outcomes of a repeated game can be sustained by using the threat of the worst punishment continuation equilibrium to deter a deviation.

explicit contract on  $x$ . In this case it is clearly no problem to induce any  $\hat{x}$  with an explicit contract. Surprisingly, however, it is no longer possible to implement any  $\hat{x}$  with an implicit contract. The reason is that the threat of terminating the relationship if any party deviates is no longer credible if explicit contracts are feasible. To see this, suppose that one party deviates from the terms of the agreement. Instead of terminating the relationship the parties could renegotiate at the beginning of the next period and write an explicit contract implementing the same allocation as before. Since the explicit contract can be enforced by the courts, the breakdown of trust does not matter. Hence, a deviation from the implicit contract cannot be deterred.

To make this argument precise, suppose that if an implicit contract has been violated, the parties can renegotiate and write an explicit contract at the beginning of the next period. The negotiation game is not modeled explicitly. Instead, it is assumed that the explicit contract maximizes social surplus and that the bargaining power of the parties is such that the agent receives share  $\alpha \in [0, 1]$  of this surplus. Furthermore, we assume that there is some cost  $C(x) \geq 0$  to writing an explicit contract on  $x$ , e.g. the costs for a lawyer to write down the contract or (perhaps more importantly) the costs for setting up a monitoring technology that makes it possible to verify the action  $x$  of the agent to the court. We allow for the possibility that the cost to fix  $x$  in an enforceable contract varies with the action  $x$  (e.g. because some actions are more difficult to specify or to monitor than others). The following proposition shows that an implicit contract is self-enforcing only if the cost of writing an explicit contract is sufficiently high.

*Proposition 2. Suppose  $\hat{x}$  is verifiable by the courts and can be contracted upon at cost  $C(\hat{x})$ . Then there exists an implicit contract which induces the agent to take action  $\hat{x} \in X$  in every period if and only if*

$$\psi(\hat{x}) - \psi(\underline{x}) \leq \frac{\delta}{1 - \delta} [S(\hat{x}) - \max\{0, S(x^*) - (1 - \delta)C(x^*)\}], \quad (2)$$

where  $x^* \in \operatorname{argmax}_x [1/(1 - \delta)]S(x) - C(x)$ .

*Proof.* To characterize the set of subgame perfect equilibrium outcomes we have to find the worst possible punishment equilibrium that can be used to deter a deviation. Note that the parties cannot commit not to renegotiate and write an explicit contract. Two cases have to be distinguished. First, if there is no explicit contract generating a positive surplus, i.e.  $[1/(1 - \delta)]S(x^*) - C(x^*) \leq 0$ , then it is an equilibrium to terminate the relationship and we are back to Proposition 1. Second, if an explicit contract is profitable, i.e.  $[1/(1 - \delta)]S(x^*) - C(x^*) > 0$ , then the two parties will write an explicit contract on the surplus-maximizing action  $x^*$ , and divide the surplus according to the allocation of bargaining power such that the agent gets  $U^* = U_0 + \alpha[S(x^*) - (1 - \delta)C(x^*)]$  while the principal receives  $V^* = V_0 + (1 - \alpha)[S(x^*) - (1 - \delta)C(x^*)]$  on average in all future periods. Note that the surplus that can be generated per period by an implicit contract on  $\hat{x}$ , in addition to the surplus of an explicit contract on  $x^*$ , is given by  $\tilde{S}(\hat{x}) = S(\hat{x}) - U^* - V^* = S(\hat{x}) - S(x^*) + (1 - \delta)C(x^*)$ . The problem of characterizing the set of allocations that can be implemented with an implicit contract if an explicit contract is

feasible is equivalent to the problem of Proposition 1 where  $U_0$  and  $V_0$  have been replaced by the new outside option utilities  $U^*$  and  $V^*$ , and  $S(\hat{x})$  has been replaced by  $\bar{S}(\hat{x})$ . Q.E.D.

Thus, if condition (2) does not hold, the parties have to rely on an explicit contract even if it is costly to do so.

#### 4. Implicit and explicit contracts

Up to now we have assumed that the agent's action is either completely contractable or not contractable at all. A more typical case seems to be that some aspects of the agent's action can be contracted upon while other aspects cannot. In the firm–worker example it is often possible to write an explicit contract on the number of hours the worker spends in the factory or the number of items produced in some given period. On the other hand, it may not be possible to write a contract on the quality of the labor supplied, even if this is symmetric information between the worker and the firm. The main question we want to address in this paper is how the set of implicit contracts is affected if some (but not all) aspects of the agent's action can be contracted upon explicitly.

Suppose that  $x$  is a vector consisting of two components,  $x = (x_1, x_2)$ . Let us assume that  $x_1$  can be contracted upon costlessly, while writing an explicit, enforceable contract on  $x_2$  is infinitely costly. What can be achieved by an explicit contract on  $x_1$  in the absence of an implicit agreement on the choice of  $x_2$ ? Suppose an explicit contract forces the agent to take  $\hat{x}_1$  in every period. Let

$$x_2(\hat{x}_1) \in \arg \min_{x_2} \psi(\hat{x}_1, x_2), \tag{3}$$

and assume for simplicity that  $x_2(\hat{x}_1)$  is uniquely defined for all  $\hat{x}_1$ .<sup>4</sup> Choose  $\bar{x}_1$  such that

$$\bar{x}_1 \in \arg \max_{x_1} \pi(x_1, x_2(x_1)) - \psi(x_1, x_2(x_1)), \tag{4}$$

and let

$$\bar{S} = \max\{0, \pi(\bar{x}_1, x_2(\bar{x}_1)) - \psi(\bar{x}_1, x_2(\bar{x}_1)) - U_0 - V_0\} \tag{5}$$

be the maximal surplus that can be generated with an explicit contract in the absence of trust. Finally, let  $(\underline{x}_1, \underline{x}_2) \in \operatorname{argmin}_{x_1, x_2} \psi(x_1, x_2)$ . The following proposition characterizes the set of allocations that can be implemented with an implicit contract alone.

*Proposition 3. For any allocation of bargaining power,  $\alpha \in [0, 1]$ , there exists an implicit contract implementing the action  $\hat{x} = (\hat{x}_1, \hat{x}_2)$  if and only if*

$$\psi(\hat{x}_1, \hat{x}_2) - \psi(\underline{x}_1, \underline{x}_2) \leq \frac{\delta}{1 - \delta} [S(\hat{x}_1, \hat{x}_2) - \bar{S}]. \tag{6}$$

<sup>4</sup> If  $x_2(\hat{x}_1)$  is not uniquely defined, choose the  $x_2(x_1)$  which minimizes  $\pi(x_1, x_2(x_1))$ .

*Proof.* Again we have to look for the worst possible punishment equilibrium. The punishment cannot be worse than the termination of the relationship. However, terminating need not be credible if an explicit contract on  $x_1$  can be written. Two cases have to be distinguished. If  $\bar{S} = 0$ , then no positive surplus can be generated by fixing  $x_1$  in an explicit contract if there is no trust and the agent chooses  $x_2$  non-cooperatively. In this case, the threat of terminating the relationship if any party deviates is credible. On the other hand, if  $\bar{S} > 0$ , then the parties will not terminate the relationship but renegotiate and realize  $\bar{S}$  with an explicit contract. The remainder of the proof follows the proof of Proposition 2. Q.E.D.

A comparison of Propositions 1 and 3 shows that if  $x_1$  is sufficiently important in the sense that fixing  $x_1$  generates a sizeable surplus even if  $x_2$  is chosen non-cooperatively by the agent, then the set of allocations that can be implemented with an implicit contract alone may be considerably reduced as compared with a situation where no explicit contracts are feasible.

However, if an explicit contract on  $x_1$  is feasible, the more interesting question is which allocations can be implemented with a combination of an explicit and an implicit contract. The set of implementable allocations is characterized by Proposition 4, which is the main result of the paper:

*Proposition 4.* For any allocation of bargaining power,  $\alpha \in [0, 1]$ , there exists a combination of an explicit contract on  $x_1$  and an implicit contract on  $x_2$  implementing the action  $\hat{x} = (\hat{x}_1, \hat{x}_2)$  if and only if

$$\psi(\hat{x}_1, \hat{x}_2) - \min_{x_2} \psi(\hat{x}_1, x_2) \leq \frac{\delta}{1 - \delta} [S(\hat{x}_1, \hat{x}_2) - \bar{S}]. \quad (7)$$

*Proof.* Suppose the parties write an explicit contract on  $\hat{x}_1$  and agree implicitly on  $\hat{x}_2$  at the beginning of the relationship. Since  $\hat{x}_1$  can be enforced by the court, the most profitable deviation of the agent is given by  $\underline{x}_2(\hat{x}_1) \in \operatorname{argmin}_{x_2} \psi(\hat{x}_1, x_2)$ . The worst possible punishment that can be used to deter a deviation is the same as in the proof of Proposition 3. Q.E.D.

Obviously, the set of implementable allocations is enlarged as compared with Proposition 3. By fixing  $\hat{x}_1$  in an explicit contract from the beginning, the agent can only deviate by choosing  $x_2$  non-cooperatively, which reduces the potential gain from a deviation. It is more interesting to compare Proposition 4 with Proposition 1. If  $x_1$  becomes costlessly contractable, then both sides of inequality (1) are reduced. The left-hand side becomes smaller because a deviation is less attractive for the agent if  $x_1$  has been fixed. On the other hand, the right-hand side may also become smaller because the worst possible punishment in case of a deviation may be reduced if an explicit contract can be written after a deviation (depending on whether  $\bar{S} > 0$  or  $\bar{S} = 0$ ). Thus, the effect of the contractability of  $x_1$  on the set of implementable allocations is ambiguous. If fixing  $x_1$  does not affect the agent's gain from a deviation very much but reduces the worst possible punishment, then it can happen that an action  $(\hat{x}_1, \hat{x}_2)$  which is implementable if only implicit contracts are feasible cannot be implemented if  $x_1$  can be controlled by an explicit contract. On the other hand, it is also possible that explicit and implicit contracts are complements. If controlling  $x_1$  does not affect the worst possible punishment equilibrium

( $\bar{S} = 0$ ) but reduces the agent's gain from a deviation, then it may be possible to implement an action  $(\hat{x}_1, \hat{x}_2)$  with a combination of an implicit and an explicit contract that could not have been implemented with an implicit contract alone.

### **Acknowledgements**

We wish to thank Bengt Holmström who stimulated this research several years ago by telling an anecdote about the insufficiency of implicit contracts when explicit contracts are available. We also wish to thank Bentley MacLeod for helpful comments and discussion. Financial support by Deutsche Forschungsgemeinschaft, SFB 303 at the University of Bonn, is gratefully acknowledged.

### **References**

- Abreu, D., 1988, On the theory of infinitely repeated games with discounting, *Econometrica* 56, 383–396.
- Baker, G., R. Gibbons and K. Murphy, 1994, Subjective performance measures in optimal incentive contracts, Working Paper, Harvard Business School, 94–154.
- Bull, C., 1987, The existence of self-enforcing implicit contracts, *Quarterly Journal of Economics* 102, 147–159.
- Eaton, J., 1993, Sovereign debt: A primer, *The World Bank Economic Review* 7, 137–172.
- Kreps, D., 1990, Corporate culture and economic theory, in: J. Alt and K. Shepsle, eds., *Perspectives on positive political economy* (Cambridge University Press).
- MacLeod, B. and J. Malcomson, 1989, Implicit contracts, incentive compatibility, and involuntary unemployment, *Econometrica* 57, 447–480.