Contracts as Threats: on a Rationale For Rewarding $A$ while Hoping For $B$*

Elisabetta Iossa‡ and Giancarlo Spagnolo‡

This version: 28 September, 2009  
First version: September 2008.

Abstract

In this paper we explore theoretically the relationship between explicit and implicit/relational contracting distinguishing between the ex-ante decision to sign an explicit contract and the ex-post decision whether to actually apply it. We show, among other things, that the relational efficient explicit contract tends to display overcontracting on tasks or qualitative requirements (A) that are verifiable but apparently of little use for the principal. The ex-post (non)implementation of such explicit contract can then be exchanged against the provision of noncontractible tasks (B) that are highly valuable for the principal. An empirical implication of the result, consistent with casual observation in procurement, is that penalties for infringements established by explicit contracts are seldom exercised, even though violations take place and are easy to monitor and verify.

*JEL Classification:* C73, D86, L14.

*Keywords:* Cooperation, implicit contracts, incomplete contracts, moral hazard, multi-tasking, relational contracts.

1 Introduction

The puzzles. Since Kerr’s (1974) amusing list of examples and Holmstrom and Milgrom’s (1991) pioneering analysis, it is well known that explicit contracts often

---

*For helpful comments we wish to thank Sumon Bhaumik, Mathias Dewatripont, Willie Fuchs, Christian Hellwig, Paul Grout, Bruno Jullien, Patrick Rey, Stephane Saussier, Pablo Spiller, an excellent anonymous referee and conference and seminars participants at ATOM Centre Paris I, Brunel University, CMPO University of Bristol, VTI-KTH Stockholm, the ECARES Conference on Infrastructure Procurement in Bruxelles and the ISNIE Conference in Berkley.

‡Dept of Economics and Finance, Brunel University; DEI, University of Rome Tor Vergata, CMPO and CEDI.

‡‡University of Rome ’Tor Vergata’, SITE - Stockholm School of Economics, EIEF and CEPR.
reward low value (if not dysfunctional) easy-to-verify tasks at the expense of more valuable but hard to monitor ones. It is probably less well known that often real world parties do not enforce explicitly contracted provisions (e.g. do not impose penalties for contractual infringements) even though, by doing so, they apparently forego substantial gains. The puzzle is therefore twofold: Why tasks of little value are often contractually specified? Why penalties for noncompliance are often not exercised?

Examples abound. In many organizations, employees are contractually required to arrive at work every day at a certain time and to stay until another time. When physical presence does not facilitate monitoring and the task can be as efficiently performed at home (think about creative non-team tasks) this is of little value for the organization. And many organizations then do not enforce these contractual provisions but allow instead flexible working hours. Similarly, in several European countries (in particular Germany, but also Italy and France) professors’ contracts establish a large number of teaching hours the researcher is required to undertake. However, universities do not always apply these provisions: individual teaching loads and administrative duties are often informally reduced, particularly for top researchers in top departments. In public procurement, penalties for noncompliance with contractual obligations are also often waived. For example, a recent inquiry found that penalties for low quality provision by contractors procuring goods and services to Italian public administrations were imposed for less that 5% of the ascertained violations.\footnote{Own calculations on data from the third party inspections commissioned by the Italian Public Procurement Agency (Consip) and coordinated by their team in charge of monitoring supply contracts (Team Monitoraggio delle Forniture). In particular, in the period 2005-2008 on a total of 4095 inspections a total of 1455 contractual infringement by the contractor were ascertained, but contractual remedies/penalties where only exercised in 64 of those cases, i.e. against about 4.42% of the infringements (with no significant difference across types of public administrations or geographical location, typically important in Italy because of differences in Social Capital). See Albano et al. (2008) for a general discussion (in Italian) of the possible reasons behind such low contract enforcement rate, besides the ‘benevolent’ one discussed in this paper.} Analogous anecdotal evidence exists for large procurement of complex services the UK, showing that the public sector does not always levy deductions for underperformance (HM Treasury, 2006). A classic example of how a rigorous application of the explicit contract can be seriously detrimental to the other party is that of “work-to-rule” practices, also called “white strikes”: for a long time trade unions in the US, Italy and Spain have been using the practice of applying the explicit contract literally during wage negotiations, blocking production all toghether.\footnote{See \url{http://libcom.org/organise/work-to-rule} for a nice up-to-date explanation of the practice from direct users, where one also reads the following: “Almost every job is covered by a maze of rules, regulations, standing orders, and so on, many of them completely unworkable and generally ignored. Workers often violate orders, resort to their own techniques of doing things, and disregard lines of authority simply to meet the goals of the company. There is often a tacit understanding, even by the managers whose job it is to enforce the rules, that these shortcuts must be taken in order to meet targets on time.”} At the level of the social contract, where prescriptive laws and regulations
can be seen as the explicit part of the contract, it is self evident that in many countries there exist plenty of inefficient laws and regulations that (fortunately) are normally not enforced.

**This paper.** This paper provides an ‘optimistic’ joint explanation for these puzzling observations by analyzing the interaction between explicit and relational contracts from a novel perspective, one emphasizing the difference between the ex ante decision to sign an explicit contract, and that of actually applying it ex post. While contract renegotiation has been subject to extensive inquiry, to our knowledge the possibility of signing an explicit contract and then simply ignoring its requirements ex post has been largely neglected until now.3

Our model shows how explicit contracts on the provision of costly but low value tasks may not be enforced in equilibrium, but only used as effective ‘threats’ that facilitate relational contracting on more valuable but non-contractible tasks.

Consider a principal-agent relationship. With overcontracting on tasks that are verifiable, costly to the agent but of little value for the principal, the principal can exchange discretionally the non-implementation of such explicit contract against the provision of non-contractible tasks that are much more valuable to the principal. The principal designs the explicit contract by requiring the agent to do more than what he actually needs him to do on the verifiable task. The principal further informally tells the agent what he really needs to do on the more valuable non-verifiable task. The principal and the agent then agree not to enforce the explicit contract on the low value (part of) the verifiable task, as long as the agent does what agreed upon on the non-verifiable task. In order to reduce his own incentives to deviate from the implicit agreement and enforce the contract on the verifiable task, the principal chooses a verifiable task that is costly to the agent to undertake but brings little benefits to himself, just sufficient to cover contract enforcement costs (otherwise the ‘threat’ would not be credible).

This reasoning offers a rationale for the puzzles discussed earlier: organizations often do not enforce contractual provisions on working hours for committed and collaborative employees, but they might apply them to punish/fire employees who behave opportunistically. Universities often accommodate teaching reductions for academics who excel in research or are particularly collaborative on other activities that are hard to specify in a contract or verify for a court. In case of opportunistic behavior however, the application of the explicit contract might be used as a sanction against the researcher. Parties involved in a long term procurement relationship may not enforce the explicit contract and levy penalties for infringements if that helps them to improve

---

3See Bolton and Dewatripont (2006) and MacLeod (2007) for encompassing surveys of explicit and relational contract theories respectively. One exception is Doornik (forthcoming), who considers the possibility of not enforcing previously signed explicit contracts when the costs of contract enforcement are large relative to its benefits.
collaboration on noncontractible aspects of their relationship, but they may do so if cooperation breaks down.\textsuperscript{4} Governments may not enforce some existing but inefficient laws or regulations, but may decide to do that in particular cases against 'troublemakers' (which of course open up to abuses, as discussed below).

Or analysis shows that overcontracting on low value contractible tasks to be ignored but used as a threat may have two distinct effects, both of which may help parties governing trade on valuable but non-contractible tasks.

The first one is a short-run effect, relevant even for one-shot or occasional transactions: explicit contracts on low value non-required contractible tasks make the stage-game sequential, giving the option to each party to react sooner to a deviation by the other on non-contractible task by calling for the application of the contract. This transforms simultaneous transactions into sequential ones and - when the contract enforcement is valueless for the principal but costly to the agent - discipline one party without increasing the temptation to defect by the other.

The second is a more familiar long-run effect: a long-term contract on inefficient actions may strengthen the future punishment phase that follows a defection, thereby helping sustain cooperation in long-term relationships.

An important aspect that emerges from the analysis is that a main advantage of overcontracting is that it make utilities non-transferable. For this reason, overcontracting dominates relational contracting where incentives are provided only through discretionary monetary transfers, as monetary transfers are transferable utility and therefore can discipline one party only at the cost of increasing the temptation to defect for the other. For this same reason, overcontracting has distributional implications: a principal with bargaining power can use it to implement the efficient relational contract without leaving any rent to the agent.

Of course, as for most other contractual commitment devices, the possibility of swift and cheap renegotiation tends to undermine the benefits of overcontracting. In fact, we show that the gain from overcontracting is smaller the lower the cost of renegotiating the contract. However, contrary to most other contractual commitment devices, fast and costless contract renegotiation at the end of the stage game does not completely undermine the benefit of overcontracting in long term relationships. Even when renegotiation is costs are zero, as long as it takes place at the end of the stage game (as in all models of relational contracting) it does not remove the ability of overcontracting of transforming simultaneous-move stage games into sequential-move ones where

\textsuperscript{4}The importance of the interaction between explicit and relational contracts emerges clearly in HM Treasury (2006): "The Government believes that the relationship between the public and private sector in a PFI project must always ultimately be contractual but should be overlaid with partnership working to ensure that operations are effective. In order to encourage this approach the Government will promote the development of a partnership agreement or shared vision document that sits outside of the actual PFI contract. This would not be legally binding but would set out the parameters of the public sector and private sector working relationship and spell out in some detail how the contract will be managed in practice."
defections can be observed and punished already in the period in which they take place.

Finally, our paper also offers a new explanation why penalties for contract infringement are often low, even though effective enforcement of contracted obligations would suggest to set them high (Becker 1968, Abreu 1988). In our model, increasing penalties for non-compliance by the agent reduces the agent’s incentive to defect but it raises the principal’s one. In order not to tighten the principal’s incentive constraint, with overcontracting contractual penalties for infringements are optimally bounded above by the cost of the contracted action for the agent.

Our results are general, but they are particularly relevant for large firms and public organizations subject to what we call ‘bureaucratic control’, i.e. formal accountability rules that do not allow their members to operate discretionary monetary payments that have no verifiable counterpart (like an explicit supply contract that justifies the payment). By limiting discretion bureaucratic control makes ‘standard’ relational contracting impossible (see e.g. Kelman, 1990). However, while checking the correspondence between contracts and payments for auditors is relatively easy, verifying whether the explicit contract has then been effectively executed, whether the promised level of quality of service has been met, and if not, whether contractual remedies have been effectively exercised, is much more difficult and rarely done. Bureaucratic control thus does not hinder relational contracting sustained through overcontracting and without using discretionary transfers. In fact, here an important additional value of signing but not applying explicit contracts arises, namely the use of the non-application of the contractual provisions as ‘method of payment’, i.e. as a reward for complying with the implicit agreement when direct payments cannot be used.

We wrote above that our is an ‘optimistic’ explanation for why explicit contracts are often not applied, since we focus on benevolent actors and the enforcement of ‘productive’ non-contractible tasks B. However, all our results can readily be re-interpreted from a less optimistic point of view, with the principal being a ‘non-benevolent’ agent of a large firm or public organization exploiting his discretion to extract B, reinterpreted as ‘Bribes’ or ‘private Benefits’, in exchange for not enforcing the explicit contract between his organization and outsiders. Similarly, a government in charge may abuse its ability to enforce normally ignored inefficient laws only in selected situations by using this discretionary power to illegally damage the opposition. As productive non-verifiable tasks, illegal gains must also be part of a self-enforcing implicit contract sustained by credible future threats, and the discretion on the ex-post execution of explicit contracts can be equally used to enforce legal/productive as well as illegal/corrupt exchanges.5

The structure of the paper. We review the related literature in Section 2. In Section 3, we consider a simple Prisoner’s Dilemma game where two parties must choose whether to cooperate or defect and highlight the value of explicit contracts as

5This is why Banfield (1975) suggested that, in order not to lose too much non-contractible quality through rigidities, some discretion and corruption should be optimally tolerated in bureaucracies.
threats when the game is one-shot and when it is repeated. In Section 4 we apply our ideas to a repeated Principal-Agent relationship in order to discuss aspects of real world relationships where contracts as threats may play an important role. Section 5 briefly concludes. All proofs are relegated to an Appendix.

2 Relation to the literature

Our work can be seen as a contribution to the economics of 'multi-tasking' and 'job design', sparked by the seminal work of Holmstrom and Milgrom (1991). That paper stressed that when some tasks are easy to monitor and contract upon whilst others are not, providing high powered incentives on the easily monitored tasks may damage the principal as it leads the agent to disregard the hard to monitor tasks that may be crucial. This formal result has been often related to Kerr's (1975) classic management science piece on "the folly of rewarding A when aiming for B". On the contrary, our results show that when one recognizes that signing and applying explicit contracts are distinct decisions and that exchanges are embedded in a long term relationship, high powered incentive contracting on valueless tasks or quality dimensions (A) may turn out to be the best instrument to obtain effort on other valuable but non-contractible tasks (B).

The static effect of over-contracting we uncover relate our model to the literature on contractual governance, and in particular to Williamson's (1983) too often forgotten idea of a third governance mode besides explicit contracts and long term relationships: transferring "hostages" between parties. In fact, the characteristics of a 'good hostage' are that of being valueless for one party, which can therefore credibly promise to return it or keep it depending on another party's behavior, much like those of 'good contracts as threats'. Accepting to sign an explicit contract that, if enforced, imposes costs to one party but produces no benefit to the other can therefore be seen as the first party accepting to become hostage of the second.

The dynamic part of the paper relates our work to various strands of literature focussing on relational contracts. Cooperative relationships sustained by "the shadow of the future" are a fundamental governance mechanism for most forms of economic interaction, and a rich literature reflects their importance. Macaulay (1963), Telser (1980) and Klein and Leffler (1981) stressed early that economic transactions are often episodes of a history of frequent exchanges where informal agreements sustained by future gains from cooperation govern crucial aspects of the relationship. The formal theory of implicit or relational contracts was developed by Bull (1987), MacLeod and Malcomson (1989, 1998) and Baker, Gibbons and Murphy (1994), and was considerably extended recently by Levin (2003), Rayo (2007), and Fuchs (2007).  

See also Calzolari and Spagnolo (2009) for an extension to multiple agents and competitive screening, and MacLeod (2007) for an elegant survey of this literature.
However, relational contracts are often embedded in a framework of explicit contracts on verifiable dimensions, and the complex interaction between explicit and relational contracts is not entirely understood. We know from Baker et al. (1994) and Schmidt and Schnitzer (1995) that effective explicit contracting may hinder relational contracts by improving parties’ fallback position when the relationships break down after defections. Baker et al. (1994) also showed that there are circumstances in which the presence of an explicit contract facilitates implicit ones, so that explicit and implicit contracts complement each other. Pearce and Stacchetti (1998) showed that a risk neutral principal will adjust the discretionary bonuses of the implicit agreement to further insure a risk averse agent, smoothing his income across periods.

More related to our work is the idea informally stressed by Masten (2000) that, to affect outside options during renegotiations, parties sometimes include contractual provisions that have little to do with what their actual intentions are. Also, Klein (1996, 2000) argued that parties may design explicit contracts with the aim of facilitating the relational contracts that complement them, in particular by limiting the parties’ temptations to defect. This last point has been recently formalized by Aghion, Dewatripont and Rey (2002) and Baker, Gibbons and Murphy (2006) in relation to the effects of the contractual allocation of decision rights on future cooperation.7 Even closer to our work is Bernheim and Whinston (1998), which also focuses on punishing defections, besides limiting them. Their model emphasizes that relational contracts need discretion to punish defections, so that creating ‘strategic ambiguity’ by limiting the extent of explicit contracting may be efficient from a relational perspective. They show that ‘undercontracting’, choosing not to contractually regulate some valuable contactible tasks, may be optimal since more discretion may facilitate punishment and thus cooperation on non-contractible tasks. Here, instead, we show that if we take into account the possibility to sign and then ignore explicit contracts, then ‘overcontracting’ - contractually requiring low value verifiable tasks - may become the best strategy to facilitate relational contracts on valuable but non-contractible tasks.

3 Main Ideas

In this section we make our points in the simplest possible setting. In the remaining sections we will extend the analysis to a principal-agent model.

Setup. Consider the game below (game B) where $D_i > C_i > N_i > S_i$, unless stated otherwise. Thus the game is a classic Prisoner’s Dilemma (for the sake of realism one can think of it as a bilateral specific investment problem with incomplete contracts). Since actions are non-contractible, the dominant strategy is to defect (not invest, say).

7 Arguments in this spirit have also been made in the less directly related literature analyzing how asset ownership interacts with relational contracting (see e.g. Baker et al. 2002, Halonen 2002).
Suppose now that players may agree, before playing $B$, to sign contracts on some verifiable actions other than those described in game $B$. In particular, suppose players can contract on some action $a_i$ that cost $a_i$ to player $i$ and produces no benefit for any player (say, digging holes in a desert). Moreover, suppose parties can choose to ignore such contracts if they wish. However, if any of the two players calls for the execution of an explicit contract, the other party must comply with it. Consider the following timing for this extended game.

**Timing 1**

*Step 1:* players sign some explicit contracts.

*Step 2:* game $B$ is played: players simultaneously choose whether to cooperate or defect.

*Step 3:* players observe the outcome of game $B$ and simultaneously choose whether to call for the application of any of the explicit contracts signed in step 1.

Graphically, the time line of this sequential game is therefore:

<table>
<thead>
<tr>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contracts on some actions $a_i$ are signed</td>
<td>Game $B$ is played</td>
<td>Outcome of $B$ is observed and decisions whether to apply/ignore contracts are taken simultaneously</td>
</tr>
</tbody>
</table>

**Benchmark.** Suppose that signed contracts cannot be renegotiated and consider the following strategy for player $i$.

**Strategies 1**

*In Step 1,* sign two contracts prescribing $a_i > D_i - C_i$ for each player $i$.

*In Step 2,* choose to cooperate.

*In Step 3,* if player $j$ cooperated in Step 2, ignore both explicit contracts. If player $j$ defected, call for the application of the explicit contract prescribing $a_j$ and ignore the one prescribing $a_i$.

By backward induction, it is easy to verify that bilateral cooperation is now a subgame perfect Nash equilibrium (although a 'weak' one because supported by player
i’s indifference on the application of the explicit contract prescribing $a_j$). If a party defects whilst the other cooperates, he will now obtain $D_i - a_i$ which by construction is lower than $C_i$, the payoff from cooperation. By writing an explicit contract on some verifiable task $A$ that has no value but is costly to undertake, and by choosing not to call for its application unless the other party deviates, the parties can sustain bilateral cooperation even in a one-shot game.

**RESULT 1.** Without renegotiation, writing and ignoring explicit contracts on independent actions that are valueless but costly to undertake allows to implement the cooperative outcome in static setting as a SPNE.

The explicit contract is used to construct a future credible punishment, ignored if a player cooperates, executed if he does not. It is of course crucial that implementing the explicit contract is not profitable for the party using it as a threat, as otherwise this party would always find it optimal ex post to call for its implementation, even after the opponent cooperated.

**Costly contract enforcement.** In reality calling for the execution of an explicit contracts entail costs, large costs in some countries. When contract enforcement is costly, in a standard set up (in the sense that only own material payoffs matter to players) a contract on a literally valueless action would not constitute a credible threat, as it would be always optimal to ignore it ex post and save contract enforcement costs.

Denote by $e > 0$ the contract enforcement cost, i.e. the cost of having an explicit contract implemented by the law enforcement system. Parties can now choose an action $a_i$ that generates benefit $\nu_i(a_i) < a_i$ for party $i$ and $\nu_j(a_i) = e$ to party $j$ (with $\nu_j(a_i) > e$ party $i$ would always call for the application of the contract, with $\nu_j(a_i) < e$ never and the threat would be empty) such that $\nu_i(a_i) + \nu_j(a_i) - a_i < 0$, so that the contract on action $a_i$ in itself is still inefficient. Now party $j$ can again credibly threaten to call for the application of the explicit contract prescribing action $a_i$ if $i$ deviates on $B$ and credibly promise not to do it when $i$ cooperates (and vice versa). Thus, we have the following.

**RESULT 2:** With positive contract enforcement costs, inefficient explicit contracts on valuable contractible tasks can be used as credible threats to elicit effort on more valuable non-contractible tasks. Contractible tasks used as threats will be more valuable the larger are contract enforcement costs.

This result suggests that in real world situations, in which contract enforcement costs are substantial, we should expect inefficient contracts on apparently valuable

---

8 The fact that in the last stage players are indifferent is not a major problem as we know that real players tend to prefer punishing if they have been cheated upon even when it is costly (Fehr and Gächter, 2000).

9 See e.g. Djankov et al. (2003) and http://www.doingbusiness.org/ExploreTopics/EnforcingContracts/
contractible tasks to be used as effective threats to implement even more valuable non-contractible tasks.

It is also worth noting that in some legal environments courts may not be willing to enforce a contract that is literally valueless for one party, although realizing or establishing this may be difficult. In principle this could make contracts as threats ineffective in the absence of enforcement costs. The presence of significant enforcement costs in reality removes even this potential problem, as then negative-gains-from-trade contracts on tasks delivering positive utility to the threatening party can (and should) be used as threats to obtain more valuable non-contractible tasks.

Remark 1 To simplify exposition and without loss of generality in the remainder of the paper we will normalize by setting \( e = 0 \). All the results will however apply to the real world case of \( e > 0 \), so the term ‘valueless task’ should be interpreted as ‘low value task that makes the threatening party indifferent’ between facing the costs of contract enforcement and save them by ignoring the contract.

Renegotiation. There are many situations in which explicit contracts cannot be renegotiated at any reasonable speed and cost, as assumed in our benchmark case above. The example of work-to-rule is a case in point: it is simply unconceivable that a collective labor contract involving unions and tenth of thousands of workers could be amended or renegotiated to prevent a work-to-rule strike in a plant aimed at punishing some form of misbehavior, say the cancellation of some informally promised benefits for workers, on the side of that plant’s management. Similarly, in public procurement contract renegotiation is often forbidden, strictly limited, or very costly.\(^{10}\)

However, there are also many other situations where renegotiation is possible at low cost and without too much delay. To take these situations into account, let us suppose here that signed contracts can be renegotiated instantaneously, and let \( z \) denote the renegotiation cost. Since implementing the explicit contract to punish defection is ex-post inefficient, the parties will renegotiate as long as the gain from renegotiation is larger than \( z \), that is if \( z \leq a_i \).

Consider first the case where the same two contracts described at Step 1 of the strategies above can be jointly and instantaneously renegotiated. Assuming 50:50 Nash bargaining the parties will obtain respectively

\[
R = \left\{ -a_i + \frac{1}{2} (a_i + a_j - 2z) ; -a_j + \frac{1}{2} (a_i + a_j - 2z) \right\},
\]

\(^{10}\)Even when contract renegotiation is feasible at low cost, it may need time, so parties could easily ensure that the explicit contracts are credible threats by establishing that the validity of the contracts expires in a time span shorter than that needed to renegotiate them.
where \(-a_i\) is the default payoff of party \(i\). Contracts as threats are then still effective as long as
\[
C_i \geq D_i - \frac{1}{2} (a_i - a_j + 2z),
\]
which in the symmetric case boils down to
\[
z \geq D_i - C_i.
\]

**RESULT 3.** When renegotiation is possible, contracts as threats remain effective as long as gains from defections are smaller than renegotiation costs, \(z \geq D_i - C_i\).

When contract renegotiation is possible, so that contractual obligations can be reciprocally waived by the parties, then only with positive renegotiation cost contracts as threats bite.

**Stipulated damages and non-transferable utilities.** Until now we have considered forcing contracts where the explicit contract has to be implemented if one party calls for its application. Suppose now that parties agree on some stipulated damages \(F_i\) for contract violation: if a party \(i\) does not comply with a contract requiring action \(a_i\), the other party is entitled to compensation \(F_i\). We let \(\mu F_i\) denote the net benefit for party \(j\) from receiving \(F_i\), where \(\mu \in [0, 1]\). After the application of the explicit contract is called for, the party can now choose between complying or paying \(F_i\). Suppose for simplicity that renegotiation costs are prohibitive. The a forth step should be added to the time line, as follows:

**Timing 2**

*Step 1:* players sign some explicit contracts \(a_i\) and \(F_i\).

*Step 2:* game \(B\) is played: players simultaneously choose whether to cooperate or defect.

*Step 3:* players observe the outcome of game \(B\) and simultaneously choose whether to call for the application of any of the explicit contracts signed in step 1.

*Step 4:* players choose between undertaking \(a_i\) or paying \(F_i\).

In this static framework, the presence of stipulated damages \((\mu F_i > 0)\) creates an incentives for parties to call for the application of the explicit contracts in the last stage of the game regardless of previous history. This pecuniary incentive destroys the value of contracts as threats not because they are not credibly implemented as a punishment, as in the case of costless and instantaneous joint renegotiation, but because they are always implemented, independently of the occurrence of a defection in game B.

**RESULT 4** With \(\mu F_i > 0\), contracts as threats are worthless in static settings.

It follows from Result 4 that in static settings when parties want to use contracts as threats and forcing contracts are not feasible, they will choose \(\mu = 0\), for example
by introducing a third party as the recipient of a fine. This highlights the importance of non-transferable utilities for contracts as threats and thus that contracts can be written that are costly to one party but bring no benefit to the other party.

**Repeated play.** We have shown that bilateral cooperation can be sustained in one-shot games by writing inefficient explicit contracts that prescribe apparently low value actions and then choose to ignore them unless one party misbehaves on other non-contractible dimensions of trade. This role of explicit contracts as threats is all the more valuable the more costly or slow is contract renegotiation relative to non-contractible trade, and the less transferable are payoffs generated by the explicit contract.

We now show that contracts as threats may be even more effective when the game is repeated. To this purpose, suppose now that the stage game described in above (Timing 2) is frequently and infinitely repeated, and that $\mu F_i > 0$, for example because the legal environment limits parties’s ability to write forcing contracts. Suppose also that contracting takes time, so that – as in all models of relational contracting – explicit contracts can be signed and renegotiated at cost $z$ at the beginning or the end of the period, but not within a period.\(^{11}\)

Consider the following strategies for player $i$.

**Strategies 2**

"Start signing two separate long-term contracts each prescribing action $a_i > D_i - C_i$ for player $i$ in each period and stipulated damages $F_i = a_i$.

In the cooperative phase, choose to cooperate and ignore the explicit contracts as long as everybody cooperates.

In the punishment phase, if player $j$ defects, call for the application of the explicit contract prescribing $a_j$ and stop cooperating forever after."

Total gains from renegotiation are then

$$\frac{\delta (a_i + a_j)}{1 - \delta} - 2\delta z,$$

where $\delta$ is the common intertemporal discount factor. With the strategies above cooperation is sustainable if the punishment is sufficiently harsh to deter all defections.

\(^{11}\)Allowing also renegotiation of strategies would complicate the set up without adding any insight. In our set up renegotiation-proofness of the punishment strategies regarding future play of game B can be easily achieved designing the punishment phase as in van Damme (1989). Moreover, since there are monetary transfers even Pareto perfection can be achieved (the strongest form of renegotiation-proofness) through strategies that are equivalent to those we analyze but require after a deviation a transfer from the deviating party to the damaged one at the end of the period when the deviation occurred or at the beginning of the next, followed by return to cooperation (see Levine 2003, Corollary 1).
Player $i$ will not defect if

$$\frac{C_i}{1 - \delta} \geq D_i - a_i + \mu F_j + \frac{\delta (N_i + \frac{-a_i + a_j}{2})}{1 - \delta} - \delta z.$$ 

Inspecting this constraint leads to the next result.

**RESULT 5.** When the game is repeated and contract renegotiation can only occur at the beginning or the end of each period, contracts as threats are valuable even when forcing contracts are not possible and $\mu F_i > 0$ or $z < D_i - C_i$. In particular, contracts facilitate cooperation as long as $\mu < 1$, or $z > 0$.

When the game is repeated, contracts may have two effects. The first one is generating sequentiality within the stage game, thereby allowing parties to punish deviations within the stage, the effect described in Results 1 and 2. The second one is due to the ability of contracts to strengthen the future punishment phase supporting cooperation in the repeated game. The first effect remains as long as $\mu < 1$, but as in the static settings (Result 4), it disappears when $\mu = 1$. The second effect remains as long as $z > 0$. It only disappears when $z = 0$, as in this case, contractual obligations can be reciprocally waived by the parties at no cost.

For example, consider the symmetric case in which $F_i = F_j = a_i = a_j$. Then, when $\mu = 1$, fines and actions cancel out in the short-term payoffs from deviation and the short-term effect of contracts as threats disappears. However, as long as $z > 0$, the long-term effect of contracts on the punishment phase remains. Conversely, when $z = 0$, renegotiation removes all effects of contracts on the punishment phase, but as long as $\mu < 1$, contracted utility is not perfectly transferable and the short-term effect remains.\(^\text{12}\)

### 4 A Principal-Agent Model

We now consider a stylized principal-agent model to discuss other aspects of real world relationships where contracts as threats may play an important role.

#### 4.1 Basic Setup

We consider a long term (infinite horizon), bilateral, repeated interaction between a principal and an agent. Time is discrete and both parties discount future payoffs through a common and strictly positive factor $\delta < 1$. Let $c_j(j)$ denote the agent’s

\(^{12}\)Note that in the asymmetric case, $a_i \neq a_j$, contracts may affect the punishment phase even when $z = 0$, by shifting the cost of the punishment phase across parties. This may further facilitate cooperation when monetary transfers are not allowed.
increasing and convex cost function of providing a certain level $j$ of effort in task (or good, service, investment of type) $J = A, B, C, \ldots$, and let $v_J(j)$ denote the weakly concave increasing value accruing to the principal when receiving intensity $j$ of task $J$ from the agent. Denote by $w_J$ the per-period surplus generated by the exchange of $J$ against a monetary transfer or price $p_j$ paid by the principal to the agent, 

$$w_J = v_J(j) - c_J(j),$$

and let $W_J = \sum_{\tau=1}^{\infty} \delta^\tau w_\tau$. Denote by $j_{FB}$ the first best level of task $J$ intensity. This maximizes $\frac{v_J(j)-c_J(j)}{1-\delta}$ leading to

$$v'_J(j_{FB}) = c_J(j_{FB}).$$

Both the principal and the agent are risk neutral and receive zero if they choose not to trade.\textsuperscript{13}

The relationship between the principal and the agent can be characterized by some verifiable ingredients, e.g. the terms of an explicit contract on a verifiable task, say task $A$ (and thus task intensity $a$, the corresponding per-period contracted price $p_A$, the fines for the agent’s non-performance $F^\alpha$ and for the principal’s non-performance $F^P$), and by non-verifiable ingredients, the self-enforcing part of the agreement on a non-verifiable task, say task $B$ (and thus task intensity $b$ and a per-period discretionary transfer $t_A$).

We start by assuming that the principal and the agent can commit to a long-term contract (or that, equivalently, renegotiation cost is very high). We relax this assumption in Section 4.5 where we study the effect of renegotiation and short-term contracting. For the sake of crispness we will continue neglecting contract enforcement costs, though all results apply to the realistic case of positive enforcement costs (according to Result 2).\textsuperscript{14}

The timing of the game between the principal and the agent can be either simultaneous or sequential. In case of sequential timing, without loss of generality, we assume that the principal moves first.

**TIMING(S)**

**period 0:** The principal and the agent may sign an explicit contract and/or agree on an implicit/relational contract.

**period 1:** An infinite repetition of (one of) the following stage game(s) takes place:

\textsuperscript{13}Linearity in income is a convenient simplification and, with no uncertainty in the model, the degree of risk aversion plays no role.

\textsuperscript{14}We should therefore continue to interpret the statement that "a the task is valueless to a party" to mean that "the expected value of calling for the enforcement of the contract is zero for that party" because the positive enforcement costs erode the positive value of the contracted task.
Stage Game 1 (simultaneous actions)

Step 1: The principal pays any explicitly contracted transfers.
Step 2: The principal and the agent simultaneously choose verifiable and non-verifiable actions (including exercising the fine $F^P$ in case the principal violated the explicit contract and implicitly contracted transfers, if any).
Step 3: The principal observes the agent’s actions and, if the agent violated the explicit contract, chooses whether or not to impose the fine $F^\alpha$.

Stage Game 2 (sequential)

Step 1: The principal pays any explicitly contracted and implicitly agreed upon transfer and chooses any other contractible or non-contractible actions.
Step 2: The agent observes the transfers and the actions chosen by the principal and chooses his own verifiable and non-verifiable actions.
Step 3: as in Stage Game 1.

4.2 Benchmark: Standard Relational Contracting

When a task $B$ is non-verifiable, explicit contracting is not possible at time 0 and a relational contract is the only potential alternative. The principal and the agent can agree on an implicit, relational contract requiring each period the agent to undertake task $B$ at level $b$ and the principal to operate a discretionary transfer $t_B$, with equilibrium per-period payoffs $\{t_B - c_B(b), v_B(b) - t_B\}$. Suppose the parties use grim strategies: if the agent deviates, the principal will withhold payment forever in the future and the agent will not cooperate; if the principal deviates, the agent will never exert effort again in the future and the principal will withhold payments. We refer to this type of relational contracting, based on the use of discretionary transfers, as "standard relational contracting" (ST).

Consider first the case of simultaneous actions, i.e. Stage Game 1. The principal and the agent will accept to participate to the contract if their respective expected payoff is non-negative, which gives the following participation constraints

$$IR^{ST}_P : V = \frac{v_B(b) - t_B}{1 - \delta} \geq 0$$

and

$$IR^{ST}_\alpha : U = \frac{t_B - c_B(b)}{1 - \delta} \geq 0.$$ 

The principal would then stick to the relational contract if defecting, by withholding the payment $t_B$ in the current period, is less profitable than sticking to the equilibrium strategy, enjoying $\frac{\delta(v_B(b) - t_B)}{1 - \delta}$ in all subsequent periods.
Consider now the relational incentive constraint of the agent (RIC-α) which is a necessary and sufficient condition for the agent not to defect. The agent would stick to the implicit contract if not exerting effort in the current period, saving \( c_B(b) \), is less profitable than sticking to the equilibrium strategy and enjoying the surplus \( t_B - c_B(b) \) in all subsequent periods. Thus we have

\[
RIC^{ST} - \alpha : U = \frac{t_B - c_B(b)}{1 - \delta} \geq t_B. \tag{3}
\]

When instead the timing of the exchange is sequential as in Stage Game 2, there are no gains from defection for the principal: as the principal moves first, deviating by not paying the transfer \( t_B \) is not profitable as the agent will not deliver \( b \). This implies that the only relevant relational incentive constraint is (RIC-α).

The following Lemma is then obtained.

**Lemma 1 (Standard Relational Contracting)** With demand for a single non-verifiable task and no explicit contracting on other tasks, a relational contract to sustain positive levels of the non-contractible task is sustainable in equilibrium: (i) if actions are simultaneous and

\[
\Phi_{Sml}^{ST}(b, \delta) \equiv \delta^2 v_B(b) - c_B(b) \geq 0; \tag{4}
\]

or (ii) if actions are sequential and

\[
\Phi_{Seq}^{ST}(b, \delta) \equiv \delta v_B(b) - c_B(b) \geq 0 \tag{5}
\]

where, from (4), \( \Phi_{Seq}^{ST}(b) \geq \Phi_{Sml}^{ST}(b) \).

Proof: see the Appendix.

Through the use of discretionary transfers \( t_B \), relational contracting can help motivating the agent when explicit contracting is not possible. For any given \( \delta \), some positive level of task intensity \( b \) can be sustained through standard relational contracting, and for \( \delta \) sufficiently high the first best intensity \( b_{FB} \) is sustainable. For \( \delta \) binding, the set of effort level sustainable with sequential actions is strictly larger than with simultaneous actions, as it happens in a repeated ‘trust game’ if compared to a payoff-equivalent repeated simultaneous actions Prisoner’s Dilemma game.\(^{15}\)

\(^{15}\)In the sequential game we have assumed that first the principal pays \( t_B \) and then the agent chooses task intensity. As the only possible deviation is from the agent, giving all the surplus to the agent by setting \( t_B = v_B(b) \), minimizes the gain from defection. If we invert the sequence of moves and assume that first the agent chooses verifiable and non-verifiable actions and then the principal pays any explicitly and implicitly contracted transfers, then the relevant constraint becomes (RIC-P). The set of sustainable actions remains the same and it is now found by setting \( t_B = c_B(b) \) and thus giving all the surplus to the principal.
4.3 The General Value of Contracts as Threats: (over)contracting on A to Obtain B

Assuming for the moment that renegotiation is prohibitively costly, suppose that the parties write a long term explicit contract on a verifiable task $A$, that is valueless to the parties (we shall show later that this is indeed optimal), and use this contract to motivate the agent to provide the valuable but non-contractible task $B$.

The principal and the agent sign a long-term explicit contract $\Phi = \{a, p_A, F^\alpha, F^P\}$ on $A$ such that $v_A(a) = 0$ for any level of $A$. The transfer in this ‘indirect’ relational contract is then $p_A$, which compensates the agent for his provision of costly $b > 0$. The principal and the agent also implicitly agree on a relational contract prescribing $a = 0$ and $b > 0$ on the equilibrium path, sustained by the threat of enforcing the explicit contract in case a defection is observed. Thus, the principal’s grim trigger strategy is: if the agent deviates in the current period, the principal will implement the contract and demand task intensity $a$ on task $A$ forever after, imposing fines $F^\alpha$ whenever that task is not delivered. The agent’s grim strategy is: if the principal deviates in the current period, the agent will not exert $b$ ever again in the future. In the punishment phase, the explicit contract is enforced: the principal will (have to) pay $p_A$ to the agent or incur fine $F^P$ and the agent will exert $a$ or incur fine $F^\alpha$. With regard to the timing of the discretionary payment $t_B$ we focus on the simultaneous case (Stage game 1).

In this case the participation constraints of the principal and the agent are respectively given by

\[
IR^{OV} - P : \frac{v_B(b) - p_A - t_B}{1 - \delta} \geq 0,
\]
\[
IR^{OV} - \alpha : \frac{t_B + p_A - c_B(b)}{1 - \delta} \geq 0.
\]

Suppose now that $t_B \geq 0$: the discretionary transfer takes the form of a payment from the principal to the agent. A deviation by the principal then consists in not paying $t_B$ when $b$ was observed and in levying the fine $F^\alpha$ on the agent for not exerting $a$.\(^{16}\) After such a defection, the principal will have to pay $p_A$ forever after (optimal if $F^P \geq p_A$), in exchange for the valueless $a$, or otherwise pay the fine $F^P$ (optimal if $F^P < p_A$). Therefore, the relational incentive constraints of the principal is

\[
RIC^{OV} - P : \frac{v_B(b) - p_A - t_B}{1 - \delta} \geq v_B(b) - p_A + \mu F^\alpha - \frac{\delta \min[p_A, F^P]}{1 - \delta}.
\]

Conversely, the agent can profitably defect from the relational contract by not providing $b$, but then he will have to simultaneously provide $a$ (optimal if $F^\alpha \geq c_A(a)$) or pay

\(^{16}\)Deviating by not paying the transfer $p_A(a_E)$ at the beginning is not profitable, as then the agent will not perform and will exercise the penalty $F^P$. 
the fine $F^\alpha$ (optimal if $F^\alpha < c_A(a)$). Therefore, the relational incentive constraints of the agent is

$$RIC^{\alpha} - \alpha : \frac{t_B + p_A - c_B(b)}{1 - \delta} \geq t_B + p_A - \min [F^\alpha, c_A(a)] + \delta p_A - \min [F^\alpha, c_A(a)]$$ \hspace{1cm} (7)

We then obtain the following Lemma:

**Lemma 2 (Optimality of Low Fines).** The set of sustainable relational contracts with explicit overcontracting is non-empty and it is maximized by setting the fines at the minimum level necessary to induce compliance with the explicit contract, i.e. $F^\alpha = c_A(a)$ and $F^P = p_A$.

Proof: see the Appendix.

Lemma 2 provides a novel rationale for fines to be bounded above. The value of fines is to induce the parties to comply with the explicit contract in the punishment phase. There is no need for the fines to penalize the parties by imposing a cost that goes beyond the cost of compliance, because non-compliance is perfectly detectable and there is nothing that fines can do which could not be done within the contract. Setting the fines higher than the cost of compliance for a party would reduce the feasibility of relational contracting by giving incentives to the other party to defect from the relational contract to cash the fine.

In the light of Lemma 2, the two relational constraints reduce to

$$R\check{I}C^{OV} - P : \frac{\delta v_B(b) - t_B}{1 - \delta} \geq \mu c_A(a),$$

and

$$R\check{I}C^{OV} - \alpha : c_A(a) \geq c_B(b) - \delta t_B.$$

Consider first the incentive constraint of the principal ($R\check{I}C$-P). By cooperating, the principal gains the surplus $v_B(b)$ from the valuable task in all future periods and gives up the gain $\mu F^\alpha = \mu c_A(a)$ from levying the fine in the current period. Consider now the incentive constraint of the agent ($R\check{I}C$-$\alpha$). By cooperating in the current period, the agent incurs the cost $c_B(b)$ in the current and in future periods but it secures $t_B$ in all future periods, whilst it saves on the cost $c_A(a)$ of undertaking the explicitly contracted task $A$ in the current and in future periods. These relational constraints are independent of $p_A$ since $p_A$ must be paid regardless of whether $b$ is performed.

Summing up the two constraints, we obtain

$$RIC^{OV} : \delta v_B(b) - c_B(b) - (1 - \delta) t_B + c_A(a)(1 - \mu(1 - \delta)) \geq 0$$ \hspace{1cm} (8)

Since the left hand side of this constraint decreases with $t_B$ and is not affected by $p_A$, the set of sustainable relational contracts is maximized at $t_B = 0$ with $p_A$ chosen so as to satisfy $IR^{OV}$. This leads us to the following Lemma.
Lemma 3 *(Sub-optimality of Discretionary Transfers).* With overcontracting, discretionary transfers are suboptimal: in any relational efficient equilibrium \( t_B = 0 \).

Proof: see the Appendix.

Before giving the intuition for the above Lemma, we state the following Proposition.

Proposition 1 *(Rewarding A to Obtain B)* (i) Explicit (over-)contracting on a valueless but verifiable task allows to sustain higher levels of the valuable and non-verifiable task than standard relational contracting under either sequential or simultaneous timing. The set of sustainable relational contracting is given by

\[
\Phi^{OV}(b, \delta) \equiv \delta v_B(b) - c_B(a)\mu (1 - \delta) \geq 0, \tag{9}
\]

with the highest implementable \( b, b^{OV}(\delta) \), obtained by choosing intensity of task \( A \) such that \( c_A(a^{OV}) = c_B(b^{OV}) \) and solving \( \Phi^{OV}(b^{OV}, \delta) = 0 \). (ii) The set of sustainable relational contracts under overcontracting is strictly greater than under standard relational contracting:

\[
\Phi^{OV}(b, \delta) > \Phi^{ST}_{Seq}(b, \delta) > \Phi^{ST}_{Sml}(b, \delta).
\]

Proof: see the Appendix.

Compared to standard relational contracting, overcontracting generates two effects. First, it introduces the agent’s obligation to comply with the explicit contract (or pay the fine) and it gives the principal the ability to levy the fine after observing whether the agent complied with the implicit contract. By allowing the principal to react immediately to a deviation by the agent, overcontracting generates a sequentiality that weakly reduces the agent’s gains from defection (strictly when the stage game is simultaneous). Through an appropriate choice of the cost of compliance, this effect can completely remove the incentives of the agent to deviate. In particular, by choosing intensity of task \( A \) such that \( c_A(a^{OV}) = c_B(b^{OV}) \), the agent is made indifferent between complying with the implicit contract, exerting \( b^{OV} \), and deviating, exerting \( a^{OV} \) (or pay the fine \( F^\alpha = c_A(a^{OV}) \)). Other things equal, overcontracting can therefore do at least as well as standard relational contracting with sequential timing where the agent moves first and only the principal’s incentive constraint matters. In fact, when \( \mu = 0 \), overcontracting also completely removes the incentives of the principal to defect as he obtains no benefit from levying the fine \( F^\alpha \). The incentive to defect of the principal increases with \( \mu \). This also suggests that forcing contracts, were feasible, maximize the benefit of overcontracting.

Second, overcontracting affects payoffs in the punishment phase starting in the period after a defection takes place. By requiring parties to comply with an explicit contract, overcontracting strengthens the punishment phase. This effect explains why overcontracting can do also strictly better than standard relational contracting with sequential timing.
The positive non-pecuniary effect of overcontracting as an 'inefficient threat' (the second effect described above), is related to the positive effect of 'money burning' in Clarke-Groves-Vickrey mechanisms and relational contracting models with subjective evaluation, like McLeod (2003), Levine (2003), and Fuchs (2007); and is reminding of the positive effect of inefficient ownership structures on investment in long term relationships with 'hold up' problems, stressed by Halonen (2002) with respect to joint ownership.

We have so far assumed that the verifiable task $A$, contracted upon to induce effort on task $B$, is valueless to the principal. The Proposition below shows however that when explicit contracts are only used as threats, it is indeed optimal to write the contract on tasks that are valueless to the principal, as this helps to enlarge the set of sustainable relational contracts.

**Proposition 2** The set of sustainable task intensities for the non-contractible task $B$ that is achievable through overcontracting on a verifiable task $A$ is maximized when $A$ is valueless to the principal.

Proof: see the Appendix.

Intuitively, if task $A$ was valuable to the principal, the principal would have more to gain from deviating by enforcing the explicit contract. His payoff in the punishment phase would therefore increase, thus reducing the gain from sticking to the relational contract.

4.4 Contracts as Threats and Bureaucratic Control

Explicit contracts that are not applied in equilibrium have additional value when the principal is a manager of a large organization subject to regulatory or governance rules. These rules often do not allow to operate discretionary payments $t_B$ that are not part of an explicit contract, say because it would be too hard then to prevent some dishonest agents present in the population to unduly appropriate part of the transfer. Let us name this common practice ‘Bureaucratic Control’ (BC), and define it by two assumptions very close to many real world situations. :

**BC1.** With BC, it is forbidden for the principal to operate ‘discretionary transfers’, i.e. transfers that are not justified by an explicit contract detailing what was that transfer paid for, and it is forbidden to renegotiate publicly awarded explicit contracts.

This type of control rule, used in most large hierarchical organizations to facilitate internal monitoring, excludes payments against any non-verifiable services/actions, like $t_B$. It only allows for monetary transfers, like $p_B$, prescribed by an explicit contract detailing for what kind of (contractible) good of service is exchanged against the monetary transfer.
BC2. Monitoring coincidence between payments and explicit contracts in the books is easy and cheap (e.g. through book audits) and it is always done; monitoring contract implementation, i.e. verifying whether the goods or services explicitly contracted upon have been effectively delivered at the quality levels the contract prescribed, is extremely costly/difficult and it is never done.

These assumptions bring to the limit, for the sake of crispness, the rather obvious observation that checking whether books are correct, i.e. that to each payment corresponds an explicit contract, is much cheaper and is done much more often in organizations than checking whether the explicit contracts have then been effectively managed/applied with the due care to obtain all what was paid for at the agreed quality level.

Suppose now that the principal is subject to BC, so that it cannot operate discretionary payments that are not part of an explicit contract. Standard relational contracting becomes impossible because the principal cannot reward the agent for his effort. However, parties under BC can circumvent the reduction of discretion through overcontracting.

**Proposition 3 (Bureaucratic Control).** With the principal subject to BC, \( t_B = 0 \) and standard relational contracts on task \( B \) are unfeasible: \( \Phi_{Seq}^{BC} = \Phi_{Small}^{BC} = \emptyset \). Explicit (over-)contracting on a valueless but verifiable task \( A \) allows to compensate the agent and sustain otherwise unfeasible relational contracts. The set of implementable task intensity satisfies \( \Phi^{OV}(b, \delta) \geq 0 \) as given in Proposition 1.

The proof is obvious and it is omitted.

This provides a further rationale as to why explicit contracts can be signed but not applied specific to situations where the principals are members of a large hierarchical organizations, like Civil Servants procuring for a Government or Procurement Managers of large firms. Overcontracting can be a means to sustain relational contracting when discretionary transfers are not allowed. Of course, all this relies on our choice to focus on ‘benevolent’ employees or civil servants. As mentioned in introduction, the same strategies could be used by non-benevolent employees to enforce implicit corrupt deals trading contract non-enforcements against bribes or private benefits B.

### 4.5 Renegotiation and Short Term Contracts

In the basic principal agent model we have assumed that the parties can commit to long-term explicit contracts. We now consider the possibility that explicit contracts can be renegotiated at low cost later than period 0, and in particular after a defection from a relational contract is observed. Clearly, this possibility has no effect under standard relational contracting, as no explicit contract is present. Consider therefore
the case of overcontracting on $a$ to sustain otherwise unfeasible levels of $b$. We will
discuss both long and short-term contracts. For simplicity we let $\mu = 1$.

In the long-term contracting case studied in the previous section, if renegotiation
costs are sufficiently small, after an implicit contract breaks up because of a defection it
is efficient for the principal and the agent to renegotiate the explicit contract to $a = 0$,
given that task $a$ is costly but valueless. Let $z > 0$ denote the cost of renegotiating
the explicit contract, and suppose that at the end (or at the beginning) of each stage
game an existing long term contract can be renegotiated, so that a fourth contract-
renegotiation step should be added to the timings described in Section 4.

Then, after a deviation is observed, at the renegotiation step the principal and
the agent will bargain to share the gain from not implementing the inefficient explicit
contract in the future net of the renegotiation cost:

$$c_A(a) - \frac{z}{2}.$$ 

When $z < \frac{c_B(a)}{1-\delta}$ at

$$c_A(a) = c_B(b),$$

we are in the situation of Section 4.3, so we focus here on $z < \frac{c_B(b)}{1-\delta}$.

Assuming $0 \leq z < \frac{c_A(a)}{1-\delta}$ and Nash bargaining in the renegotiation phase, the agent
obtains

$$\frac{\delta (p_A - c_A)}{1-\delta} + \frac{1}{2} \left( \frac{\delta c_A}{1-\delta} - z \right) = \frac{\delta (p_A - \frac{1}{2} c_A)}{1-\delta} - \frac{z}{2},$$

whilst the principal obtains

$$\frac{-\delta p_A}{1-\delta} + \frac{1}{2} \left( \frac{\delta c_A}{1-\delta} - z \right) = -\frac{\delta (p_A - \frac{1}{2} c_A)}{1-\delta} - \frac{z}{2}.$$ 

Therefore, the relational incentive constraints of the principal and the agent become
respectively (fines are unchanged)

$$RIC^OV - P: \frac{v_B(b) - p_A}{1-\delta} \geq v_B(b) - p_A + c_A(a) - \delta \frac{p_A - \frac{1}{2} c_A(a)}{1-\delta} - \frac{z}{2},$$

$$RIC^OV - \alpha: \frac{p_A - c_B(b)}{1-\delta} \geq p_A - c_A(a) + \delta \frac{p_A - \frac{1}{2} c_A(a)}{1-\delta} - \frac{z}{2}. \tag{10}$$ 

We then obtain the following Lemma.

**Lemma 4** With overcontracting and renegotiation cost of $z$, the set of sustainable
relational contracts is given by

$$\Phi^OV(b, \delta) = \delta v_B(b) - c_B(a)(1-\delta) \geq 0 \quad \{ \text{if} \quad \frac{c_B(b)}{1-\delta} < z, \}$$

$$\tilde{\Phi}^OV(b, \delta, z) \equiv \delta v_B(b) - c_B(b) + (1-\delta) z \geq 0 \quad \{ \text{if} \quad \frac{c_B(b)}{1-\delta} \geq z. \} \tag{11}$$

where $\tilde{\Phi}^OV(b, \delta, z)$ is increasing in $z$ with

$$\tilde{\Phi}^OV(b, \delta, z) \leq \Phi^OV(b, \delta).$$
Proof: see the Appendix.

By increasing the payoffs of the parties in the punishment phase, renegotiation reduces the sustainability of relational contracts with overcontracting. As payoffs raise more the lower is the renegotiation cost, the set $\Phi^{OV}(b, \delta, z)$ of sustainable relational contracts shrinks as $z$ decreases.

As for other inefficient contractual devices delivering strategic advantages, like strategic delegation contracts, the possibility of renegotiating at low costs tends to undermine the benefit of the arrangement (Dewatripont 1988). For example, in Halonen (2002) the positive effect of joint ownership is smaller the lower are renegotiation costs, and the effect vanishes all together when the renegotiation cost is zero. Similarly, Kocksen and Ok (2004) found recently additional strategic effects of contractual delegation, but only with costly renegotiation. These effects vanish when renegotiation becomes costless.

On the contrary, in our setup even in the degenerate case of $z = 0$ renegotiation does not completely eliminate the value of overcontracting. This is because of the first effect of overcontracting described above, namely the ability of overcontracting to create sequentiality when the original stage game is simultaneous. Indeed, by inspection, at $z = 0$, $\Phi^{OV}(z = 0, b) = \Phi^{ST}_{Seq}(z = 0, b)$ in (5). The following Proposition is then obtained.

**Proposition 4 (Overcontracting with Renegotiation).** Renegotiation reduces the set of relational contracts sustainable with overcontracting, the more the lower the renegotiation cost $z$. However: (i) As long as $z \neq 0$, overcontracting strictly dominates standard relational contracting, as $\Phi^{OV}_{\delta}(.) > \Phi^{Seq}(.) > \Phi^{Smt}(.)$. (ii) At $z = 0$, overcontracting weakly dominates standard relational contracting, and strictly dominates it when the timing of the stage game is simultaneous or if there is BC.

Proposition 4 stems from a combination of two factors. First, when renegotiation occurs, the agent bears a cost of deviating which is given by the cost of complying with the explicit contract. Second, such cost is endogenous: when renegotiation becomes less costly, the parties can design a more demanding explicit contract so as to leave the payoff of the agent unchanged. Thus, regardless of the cost of renegotiation, the agent can be given incentives not to deviate.

In our simple framework we also obtain this result when only one-period explicit contracts are available.

**Corollary 1** The set of sustainable relational contracts with overcontracting when only short term contracts are possible is the same as that with long-term overcontracting and costless renegotiation ($z = 0$). Hence, also when only short term contracts exist overcontracting weakly dominates standard relational contracting, and strictly dominates it when the timing of the stage game must be simultaneous or if there is BC.
4.6 Distributional Effects and Financing Constraints

In this section we make the extreme assumption that renegotiation is costless (or that only one-period explicit contracts can be signed), and show that even with sequential timing there might still be reasons for using overcontracting to sustain relational contracts.

**Distributional effects.** In standard relational contracting, an agent who gets no surplus gets no long-term benefit from continuing the contract, and will therefore deviate if the timing of the exchange cannot be properly adjusted by making the agent move first. Indeed, without overcontracting the condition \( t_B \geq c_B(b)/\delta \) is necessary for \((\text{RIC}^\text{ST})\) to be satisfied. Thus, under standard relational contracting, the agent must be given a rent of at least \((1 - \delta)c_B/\delta\). Under overcontracting, instead, an agent who gets no surplus can still get a long-term benefit from continuing the contract because there is a cost of deviating. In the case of no renegotiation, for example, when the agent deviates he triggers the implementation of the explicit contract, which costs him \(c_A(a)\) (or the correspondent fine \(F^\alpha\)). Since deviating saves the agent the amount \(c_B(b)\), choosing \(a\) such that \(c_A(a) = c_B(b)\) suffices to remove any incentive to deviate. We then obtain:

**Corollary 2** Even if \(z = 0\) or only short term contracts are available, if the timing of the exchange follows Stage Game 1 or 2, overcontracting on \(A\) allows the principal to implement any level of \(B\) that is feasible under standard relational contracting whilst retaining a greater share of the surplus. With overcontracting, no rent is needed to discipline the agent.

Proof: see the Appendix.

**Financing constraints.** This same ability of overcontracting to solve problems linked to the sequence of exchange has also important effects when there are financial constraints that prevent agents to implement the efficient sequence of the exchange.

Suppose, for example, that the principal has all bargaining power and therefore is the residual claimant of the future rent from the relationship. Suppose further that the task is a nonverifiable but costly investment, and that the agent does not have own money to finance such an investment and is financially constrained (cannot borrow it). Since the principal has the bargaining power, we know that it would be efficient that he required the agent to invest first, thus reversing the sequence of moves in Stage Game 2. However, if the agent is financially constrained, he will not be able to invest first and an inefficient timing will have to be chosen that will also cost the principal part of the obtainable rent, as shown by the Corollary above.
Overcontracting on a useless but verifiable task, even if $z = 0$ or only short term contracts are available, allows to solve the problem thanks to its distributional effects. It allows the principal to pay first $p_A$ to the agent, which allows the agent to finance the agreed upon investment $b$, without inducing any inefficiency and rent redistribution. We thus have:

**Corollary 3** Even if $z = 0$ or only short term contracts are available, overcontracting allows to sustain efficient relational contracts when financing constraints would not allow it otherwise.

## 5 Conclusions

We have shown that explicit contracts can not only be seen as safe "boundaries" within which relational contracts operate better, as suggested by Klein (2000), or as constraints on discretion that hinder relational contracting, as suggested by Bernheim and Whinston (1998), but also as credible "threats" that - not applied on the equilibrium path - actively help governing relationships. When this is the case, overcontracting may be observed, as explicit and not applied contracts on verifiable tasks with little value may enforce better relationships than standard relational contracts. This result could explain some cases in which we observe that explicit contracts regulate and reward apparently useless tasks and are not actually applied, and can be seen as the complement to the undercontracting result in Bernheim and Whinston (1998), where explicit contracting on valuable tasks is reduced to facilitate relational contracts. Both these results are based on externalities between contractible and non-contractible tasks, and therefore open the way to many new questions on optimal job and organizational design.

Also, as mentioned earlier, we have been mainly focussing on legal and productive non-contractible tasks $B$ (effort, or investments). However, illegal tasks, bribes and private benefits also cannot be explicitly contracted (since they are forbidden, they cannot be brought in front of a court), so that overcontracting may help to sustain corrupt relational exchanges exactly as it helps productive ones. The trade off between the positive and negative effects of overcontracting, and the optimal regulatory response to it in different legal and cultural environments appears therefore and important question to address in future work.
6 The Appendix

Proof of Lemma 1. (i) The set of feasible relational contracts with standard relational contracting and simultaneous timing maximizes \( b \) subject to the two participation constraints, \((IR^{ST-P})\) and \((IR^{ST-\alpha})\) and the two relational constraints, \((RIC^{ST-P})\) and \((RIC^{ST-\alpha})\). When \((RIC^{ST-P})\) and \((RIC^{ST-\alpha})\) are satisfied, \((IR^{ST-P})\) and \((IR^{ST-\alpha})\) are also satisfied, thus we can ignore \((IR^{ST-P})\) and \((IR^{ST-\alpha})\). Summing across \((RIC^{ST-P})\) and \((RIC^{ST-\alpha})\), we obtain a necessary condition for the relational incentive constraints to hold, the global relational incentive constraint, given by

\[
RIC^{ST} : \delta v_B(b) - c_B(b) - (1 - \delta) t_B \geq 0. \tag{12}
\]

Since the LHS of \( RIC^{ST} \) is decreasing in \( t_B \), the set of sustainable \( b \) minimizes \( t_B \) subject to \((RIC^{ST-P})\) and \((RIC^{ST-\alpha})\). This gives \( t_B = c_B(b) \) which is the value of \( t_B \) such that \((RIC^{ST-\alpha})\) is binding. Substituting for this value in \((RIC^{ST-P})\), we obtain expression (4).

(ii) With sequential timing following stage game 2, the set of sustainable relational contracts satisfies \((IR^{ST-P})\), \((IR^{ST-\alpha})\) and \((RIC^{ST-\alpha})\). However, when \((RIC^{ST-\alpha})\) is satisfied, \((IR^{ST-\alpha})\) is also satisfied. Thus we can ignore \((IR^{ST-\alpha})\). Noting that \((RIC^{ST-\alpha})\) is easier to satisfy the higher is \( t_B \), the set of sustainable \( b \) is maximized at the maximum \( t_B \) compatible with \((IR^{ST-P})\), which gives \( t_B = v_B(b) \). Substituting for \( t_B = v_B(b) \) in \((RIC^{ST-\alpha})\) we obtain expression (5).

Proof of Lemma 2. Summing up \((RIC^{OV-P})\) and \((RIC^{OV-P})\), we obtain

\[
RIC^{OV} : \frac{v_B(b) - c_B(b)}{1 - \delta} \geq v_B(b) + \mu F^\alpha - \frac{\delta \min[p_A, F^P]}{1 - \delta} - \frac{\min[F^\alpha, c_A(a)]}{1 - \delta} + \frac{\delta p_A}{1 - \delta}.
\]

By inspection, \( F^P \) should be maximized, as it appears with a negative sign on the RHS of the inequality; thus \( F^P = p_A \). Now suppose \( F^\alpha < c_A(a) \), then \( RIC^{OV} \) becomes

\[
\frac{v_B(b) - c_B(b)}{1 - \delta} \geq v_B(b) - \frac{1 - \mu (1 - \delta)}{1 - \delta} F^\alpha
\]

and as \( F^\alpha \) appears with a negative sign on RHS of \( RIC \), the set of sustainable contracts is found by setting \( F^\alpha = c_A(a) \).

Proof of Lemma 3. Suppose that \( t_B \leq 0 \). In this case a deviation by the principal only results in the levying of the fine; since \( t_B \) is non-positive, it is the agent who can withhold payment to the principal and not vice versa. The relational incentive constraints therefore become

\[
RIC^{OV-P} : \frac{v_B(b) - p_A - t_B}{1 - \delta} \geq v_B(b) - p_A + \mu F^\alpha - \frac{\delta \min[p_A, F^P]}{1 - \delta}.
\]
and

\[ RIC^OV - \alpha : \frac{t_B + p_A - c_B(b)}{1 - \delta} \geq p_A - \min [F^\alpha, c_A(a)] + \frac{\delta p_A - \min [F^\alpha, c_A(a)]}{1 - \delta}. \]

Summing up these two constraints (fines are unchanged), we have

\[ RIC^OV : \delta v_B(b) + (1 - \delta)t_B - c_B(b) + c_A(a)(1 - \mu(1 - \delta)) \geq 0 \]

where the LHS is increasing in \( t_B \), thus suggesting the optimality of \( t_B = 0 \).\[ \blacksquare \]

**Proof of Proposition 1.** The set of feasible relational contracts with overcontracting maximizes \( b \) subject to \((IR^OV-P), (IR^OV-\alpha), (RIC^OV-\alpha) \) and \((RIC^OV-P)\). Since \( p_A \) only affects \((IR^OV-P), (IR^OV-\alpha)\), then any \( p_A \in [c_B(b), v_B(b)] \) is feasible. Now let \( b^{OV}(\delta) \) denote the highest sustainable \( b \). We show that at \( b^{OV} \), both \((RIC^OV-\alpha)\) and \((RIC^OV-P)\) must be binding. Suppose by contradiction that \((RIC^OV-P)\) is binding whilst \((RIC^OV-\alpha)\) is slack. Then \( c_A(a) \) can be reduced so as to keep \((RIC^OV-\alpha)\) satisfied whilst loosening \((RIC^OV-P)\), making higher levels of \( b \) implementable: a contradiction.

Suppose next that \((RIC^OV-\alpha)\) is binding whilst \((RIC^OV-P)\) is slack. Then, we can increase \( b \) and \( a \) so as to leave \((RIC^OV-\alpha)\) binding and keep \((RIC^OV-P)\) satisfied. Setting both \((RIC^OV-\alpha)\) and \((RIC^OV-P)\) binding we obtain expression (9).

**Proof of Proposition 2.** Let \( v_A(a) \) denote the value of the verifiable task \( A \). The relational constraint of the principal modifies to

\[ RIC^OV - P : \frac{v_B(b) - p_A}{1 - \delta} \geq v_B(b) - p_A + \mu F^\alpha + \frac{v_A(a) - \delta \min [p_A, F^P]}{1 - \delta}. \]

whilst \( v_A(a) \) has no effect on \((RIC^OV-P), (IR^OV-P)\) and \((IR^OV-\alpha)\). Since the LHS of \((RIC^OV-P)\) is independent of \( v_A(a) \) whilst the RHS is increasing in \( v_A(a) \) and since the set of sustainable relational contracts is maximized when \((RIC^OV-P)\) binds, the result follows.\[ \blacksquare \]

**Proof of Lemma 4.** Suppose that \( \frac{\delta c_B(b^{OV})}{1 - \delta} \geq z \). The two relational constraints simplify to

\[ RIC^OV - P : \delta v_B(b) \geq c_A(a)[\mu(1 - \delta) + \frac{\delta}{2}] - (1 - \delta)\frac{z}{2}, \]

and

\[ RIC^OV - \alpha : c_A(a)(1 - \frac{\delta}{2}) \geq c_B(b) - (1 - \delta)\frac{z}{2}. \]

Following the same reasoning as in Proposition 1, the set of sustainable relational contracts is found by choosing \( a \) and \( b \) such that \((RIC^OV-\alpha)\) and \((RIC^OV-P)\) are binding.
This gives
\[ c_A(\hat{a}^{OV})(1 - \frac{\delta}{2}) = c_B(\hat{b}^{OV}) - (1 - \delta)\frac{z}{2}, \]
\[ \delta v_B(\hat{b}^{OV}) - c_B(\hat{b}^{OV})\frac{2(1 - \delta)\mu + \delta}{2 - \delta} + (1 - \delta)z\frac{1 + \mu(1 + \delta)}{2 - \delta} = 0. \]

with \( \frac{\delta}{1 - \delta}c_A(\hat{a}^{OV}) - z = 2\frac{\delta c_B(\hat{b}^{OV}) - z(1 - \delta)}{(1 - \delta)(2 - \delta)} > 0 \) for \( \frac{\delta c_B(\hat{b}^{OV})}{1 - \delta} \geq z \). Then the parties will prefer not to renegotiate the contract during the punishment phase and the analysis is equivalent to the case developed in Section 4.3. Thus the set of feasible relational contracts is then given by (9).

**Proof of Corollary 1.** Suppose now that the principal and the agent agree to a short term contract and to renegotiate it as long as parties do not defect. The relational incentive constraints of the principal and the agent become respectively (fines are unchanged)

\[ RIC_{Short}^{OV} - P : \frac{v_B(b) - p_A}{1 - \delta} \geq v_B(b) - p_A + c_A, \]

and

\[ RIC_{Short}^{OV} - \alpha : \frac{p_A - c_B(b)}{1 - \delta} \geq p_A - c_A(a). \]

Simplifying

\[ RIC_{Short}^{OV} - P : \delta(v_B(b) - p_A) - c_A(1 - \delta) \geq 0 \]

and

\[ RIC_{Short}^{OV} - \alpha : \delta p_A - c_B(b) + c_A(1 - \delta) \geq 0. \]

The set of relational contracts is found by setting \( c_B(b) = c_A = p_A \), yielding

\[ \Phi_{Short}^{OV}(b) = \delta v_B(b) - c_B(b) \geq 0, \]

which is equal to \( \bar{\Phi}^{OV}(b, \delta, z) \), at \( z = 0 \) where \( \bar{\Phi}^{OV}(b, \delta, z) \) is given by (11). This proves the first statement. The second one follows directly from the Proposition.

**Proof of Corollary 2.** (i) Consider the case of overcontracting with no renegotiation. We know from the Proof of Proposition 1 that the set of sustainable relational contracts is maximized by choosing \( a \) such that \( c_A(a) = c_B(b) \). The result then follows by noting that \( (RIC_{OV}^{\alpha}) \) in expression (7) is satisfied at \( p_A(b) = c_A(a) = c_B(b) \) and fines satisfying Lemma 2. (ii) Now consider the case of renegotiation. We know from the proof of Lemma 4 that the set of sustainable relational contracts is such that \( c_A(a)(1 - \frac{\delta}{2}) = c_P(b) - (1 - \delta)\frac{z}{2} \). The result then follows by noting that \( (RIC_{OV}^{\alpha}) \) in expression (10) is satisfied at \( p_A(b) = c_B(b) \) and \( c_A(a)(1 - \frac{\delta}{2}) = c_B(b) - (1 - \delta)\frac{z}{2} \).
References


Notes for presentation:

- the possibility of writing an explicit contract and then not enforcing it has been largely ignored by the literature on contract theory
  
  (see e.g. Bolton Dewatripont (2006) MIT Press; an exception is Doornik (JLEO 2007), http://jleo.oxfordjournals.org/cgi/content/abstract/ewn017)

- we are not arguing that this is the only or the main explanation for the puzzles, many other stories can be told, but our is clearly an important and intuitive one among them

- we are not talking about really ’valueless’ tasks, but of positive by low value tasks whose benefits for the principal are of the same order of magnitude of the cost of contract enforcement (which may be small or very large in reality, varying from country to country), so that the expected gains of enforcing the explicit contract are close to zero for that party but its effect as threat is large because of the large cost on the other...

- the static effect can be seen as a form of ’governance through hostages’ (Williamson 1983)