# Costly Contracting in a Long-Term Relationship[*]

Pierpaolo Battigalli
Bocconi University
pierpaolo.battigalli@uni-bocconi.it

Giovanni Maggi
Princeton University
maggi@princeton.edu

November 2004

## Abstract

We examine a model of contracting where parties interact repeatedly and can contract at any point in time, but writing enforceable contracts is costly. A contract can describe contingencies and actions at a more or less detailed level, and the cost of writing a contract is proportional to the amount of detail. We consider both formal (externally enforced) and informal (self-enforcing) contracts. The presence of writing costs has important implications both for the optimal structure of formal contracts, particularly the tradeoff between contingent and spot contracts, and for the interaction between formal and informal contracting. Our model sheds light on these implications and generates a rich set of predictions about the determinants of the optimal mode of contracting.

*JEL* classification: D23, C73.

KEYWORDS: writing costs, contingent *vs* spot contracting, formal *vs* informal contracts.

# 1. Introduction

Contracting in a long-term relationship may come in a variety of modes. Contracts may be formal (i.e. externally enforced), informal (i.e. self-enforcing), or a combination of the two. Formal contracting in turn may be of the "spot" type, of the contingent type, or a mixture of the two.[1] Empirical examples of these different contracting modes abound. The classic paper by Macaulay (1962) and a number of subsequent studies have highlighted that formal and informal contracting often coexist in long-term relationships, such as the ones between firms and suppliers or between employers and employees (see for example Baker *et al.*, 1994, for a discussion of this literature). In the construction industry, where formal contracting is intensively used, the relationship between contractors and subcontractors often exhibits a mix of spot and contingent contracting. An important part of the contract concerns the project specifications, and this part is replaced project by project (spot contracting). But there is typically another part of the contract that applies to all projects, and in this part one can often find contingent clauses, for example clauses that specify circumstances under which a party can make changes to the project or terminate it.[2] One can also find examples of long-term relationships where contracts become gradually more contingent over time. Meihuizen and Wiggins (2000) study supply contracts in the U.S. natural gas industry and find that many of the observed contracts tend to become richer and to incorporate a growing number of contingent clauses over time.

In this paper we present a repeated principal-agent model in which the contracting mode is endogenously determined. A key feature of the model is that writing enforceable contracts is costly. We argue that writing costs have strong implications for the optimal mode of contracting, even if their magnitude is small. The model highlights that it may be optimal to use different

---

[1]By spot contracting we mean that the contract is noncontingent but gets modified as circumstances change. The difference between contingent and spot contracting is nicely explained by Williamson (1985, p.20), who states that the writing of contracts "...*can be done with a great deal of care, in which case a complex document is drafted in which numerous contingencies are recognized, and appropriate adaptations by the parties are stipulated and agreed to in advance. Or the document can be very incomplete, the gaps to be filled by the parties as the contingencies arise. Rather, therefore, than contemplate all conceivable bridge crossings in advance, which is a very ambitious undertaking, actual bridge-crossing choices are only addressed as events unfold.*"

[2]See Eccles (1981) for an analysis of the contractor/subcontractor relationships in the U.S. construction industry.

contracting modes for different aspects of the relationship, and yields rich predictions on the impact of fundamental parameters on the optimal mode of contracting.

Next we give a more detailed preview of the model and of the main results.

We consider a multi-task, principal-agent setting with verifiable contingencies and actions, where parties interact repeatedly and can write contracts at any point in time (this includes the possibility of spot contracting, as contracts can be written after observing the state of nature and before actions are taken). A contract can describe contingencies and actions at a more or less detailed level, and the cost of writing a contract is increasing in the amount of detail. In each period, parties can save on writing costs by modifying the previous contract rather than drafting a whole new contract.

In the first part of the paper we focus on situations where the parties rely entirely on formal contracting. A simple way to capture these situations is to consider the finite-horizon version of the game, where there are no credible punishments to enforce obligations that are not specified in the formal contracts. We examine the optimal structure of formal contracts, and in particular the choice between contingent and spot contracts. At the intuitive level, it is not obvious whether the presence of writing costs should favor contingent or spot contracting: on the one hand, spot contracting avoids the cost of describing contingencies; on the other hand, spot contracts must describe the agent's *behavior* repeatedly, and this may push in favor of a contingent contract. The idea that transaction costs might favor contingent contracting has already been expressed informally, for example by Hart and Holmstrom (1987, p. 130), who write: "*if a relationship is repetitive, it may save on transaction costs to decide in advance what actions each party should take rather than to negotiate a succession of short term contracts.*" Our model allows us to explore this tradeoff more systematically.[3]

Absent writing costs, the model has no predictive power, because there is a plethora of optimal contracting plans, including a contingent contract, a sequence of spot contracts, and a host of intermediate solutions. But in the presence of (even very small) writing costs, the model

---

[3]In our model contingent contracting may be preferred to spot contracting because it is costly to describe actions. Another type of transaction cost that can generate a preference for contingent contracts is given by fixed recontracting costs. This point is made in the context of macroeconomic models by Gray (1976, 1978) and Dye (1985b).

yields a unique optimum. Each task is optimally handled in one of three ways: (i) a contingent clause, written once and for all; (ii) a sequence of spot clauses; or (iii) a noncontingent clause that is replaced at some point in time by a contingent clause (we refer to this as the "enrichment" approach). A contingent approach is optimal when the cost of describing contingencies is low relative to the cost of describing actions; a spot approach is optimal when this relative cost is high; and an enrichment approach may be optimal when this relative cost takes intermediate values.

If the enrichment approach is optimal for some of the tasks, the model predicts that the number of contingent clauses in the contract will increase over time. This is consistent with Meihuizen and Wiggins's (2000) observation that contracts in the natural gas industry become gradually more contingent over time. The interpretation of this phenomenon suggested by our model is that it may be efficient to introduce a new contingent clause in the contract when a new event occurs that was previously considered unlikely.[4]

Two other important determinants of the optimal contracting mode are the degree of uncertainty and the durability of the relationship. We find that tasks characterized by a higher degree of uncertainty are more likely to be regulated by contingent clauses, whereas lower-uncertainty tasks are more likely to be regulated by a spot approach. Similarly, the model predicts that the prevalence of spot contracting relative to contingent contracting should be higher when the external environment is more stable. On the other hand, we find that the relative prevalence of spot contracting is lower when the relationship is more durable. Interestingly, then, the stability of the relationship and the stability of the external environment have opposite effects on the choice between spot and contingent contracts.

Our model also offers interesting insights regarding the interaction of formal and informal contracting. As we mentioned at the outset, there is ample anecdotal evidence that these two modes of contracting often go hand in hand. The presence of writing costs provides a theoretical

---

[4]One might argue that the new contingent clauses may have been added because the parties became aware or previously unforeseen contingencies. But this cannot fully explain the increase in contingent clauses documented by Meihuizen and Wiggins. Around 1975, most of the gas supply contracts were amended to include a new clause that provided for renegotiation of the price in case of deregulation of the industry. Since the industry was regulated, the contracting parties were almost by definition aware of the possibility of deregulation. We are therefore inclined to think that this contingency was missing because it was considered very unlikely, and was later introduced because its likelihood was revised upwards.

explanation for the coexistence of formal and informal contracting. Intuitively, the advantage of a self-enforcing contract is that it can be communicated informally, rather than being written formally, and this saves on writing costs. On the other hand, if parties are not very patient, they will not be able to govern all tasks by self-enforcing contract. Hence there is a trade-off between formal and informal contracting.

Formally, we allow for the possibility of self-enforcing contracts by focusing on the infinite-horizon version of the game. We find that formal and informal contracting tend to be used jointly, with some tasks regulated formally and others regulated informally. In particular, tasks characterized (other things equal) by a lower cost for the agent are regulated by informal contracting, while higher-cost tasks are regulated by formal contracting. This is because the agent has a stronger incentive to cheat for higher-cost tasks. The presence of writing costs can thus contribute to explain the fact that long-term relationships are often managed by a combination of formal and informal contracting.

Formal and informal contracts coexist provided that (i) the discount factor is not too high, so that a fully informal contract cannot be sustained, and (ii) at least one task has a relatively low cost, so that at least one task can be regulated informally. Interestingly, this is true even if writing costs are very small. In fact, reducing writing costs may even *facilitate* informal contracting, as we point out below.

When we examine the effects of exogenous parameter changes, we find that the relative importance of informal contracting tends to be higher when the potential surplus from the relationship is larger, when the degree of uncertainty is lower, and when writing costs are lower. This is a consequence of the fact that each of these parameter changes relaxes the incentive-compatibility constraints, and hence makes informal contracting easier to sustain.

Earlier in this introduction we talked about the tradeoff between spot and contingent contracting when only formal contracting is available. This tradeoff changes in an interesting way when informal contracting is available. The key point is that with spot contracting, writing costs must be incurred repeatedly over time, and this strains the incentive constraint, whereas with a contingent contract all writing costs are incurred upfront, and this does not place any burden on the incentive constraint. Thus contingent contracting has better incentive properties

4

than spot contracting. We conclude that the relative importance of contingent versus spot contracting should be higher in relationships where formal and informal contracting coexist, as compared with situations where contracting is purely formal.

Most of our analysis focuses on the case in which writing costs are relatively small, and more specifically, small enough that it is optimal to implement the first best outcome. In other words, we focus on the case in which contracting is "complete."[5] The emphasis on this case is useful for two reasons. First, all the main insights can be brought out with small writing costs, and the exposition is considerably simpler than in the case of large writing costs. Second, this helps to clarify that contractual incompleteness *per se* is not essential to our results. In the final part of the paper we show how to extend the analysis to the case of large writing costs. If writing costs are large, the main difference in results is that contracting may be incomplete. Contractual incompleteness can take two forms: rigidity and discretion. Other things equal, tasks characterized by low surplus are left to the agent's discretion, intermediate-surplus tasks are regulated by rigid rules, and high-surplus tasks are regulated in a first-best way (by formal or informal contracting). This parallels a result derived in Battigalli and Maggi (2002), which focuses on a static model of contracting with writing costs.[6]

The interaction between formal and informal contracts has been analyzed also by Baker *et al.* (1994) and Pearce and Stacchetti (1998).[7] These papers propose a different explanation for the combined use of the two types of contract. They consider a repeated principal-agent model where formal contracts can be based only on verifiable signals of the agent's action, whereas informal contracts can be based on unverifiable signals. In these models it may be optimal to combine a formal wage and an informal 'bonus'. Our model, on the other hand, explains why it may be efficient to regulate some tasks formally and some others informally (in our model there is no need for bonuses). Perhaps more importantly, the rationale for mixing formal and informal contracting in our model is not the presence of verifiable and non-verifiable signals,

---

[5]By "complete" contracting we mean simply that the first-best outcome is implemented, so this may include contingent formal, spot formal, and informal contracting.

[6]We should mention that there are other papers on complexity costs as a cause of contractual incompleteness in a static setting, for example Dye (1985a), Anderlini and Felli (1994, 1999), MacLeod (2000), Krasa and Williams (2001) and Al Najjar *et al.* (2002).

[7]There is also a vast literature on purely self-enforcing contracts. Bull (1987) and Mac Leod and Malcomson (1989), and Levin (2003) are prominent examples of this literature.

but the interaction between writing costs and self-enforcement constraints.

Our analysis also yields contrasting predictions on the interplay between formal and informal contracting. One key result in Baker *et al.* (1994) is that, if the imperfection in formal contracting is sufficiently small, informal contracts cannot be sustained. Therefore their model predicts that, if the formal-contracting system becomes more efficient, informal contracting may be undermined as a consequence. According to our analysis, even if formal contracting is close to perfect (i.e. if writing costs are close to zero), the optimum typically involves a mix of formal and informal contracting or a fully informal contract. Thus, we suggest that informal contracting need not disappear as the formal-contracting system becomes more efficient.[8]

Finally, our model yields interesting predictions on the choice between spot and contingent contracts. The models by Baker *et al.* (1994) and Pearce and Stacchetti (1998) are silent about this issue because the nature of the contractual imperfection is different. In these models the imperfection is given by the noisiness of verifiable signals about the agent's performance. This imperfection is exogenous and is present in every period, while in our model the imperfection (writing costs) is endogenous and need not be incurred in every period.

## 2. The contracting environment

We analyze a repeated multi-task principal-agent game where parties can contract in each period. Payoffs depend on actions and external contingencies, which are both verifiable in court but costly to describe in a written contract. Thus, the only contractual imperfection is the presence of writing costs. We model payoffs and writing costs in a similar way as in Battigalli and Maggi (2002). We start by describing the language used to write contracts.

$\Pi^e = \{e_1, e_2, ..., e_N\}$ is a finite collection of *primitive sentences*, each of which describes an *elementary event* concerning the external environment. For example, $e_1$: "the passenger has a moustache", $e_2$: "the passenger's bag is red".

$\Pi^a = \{a_1, a_2, ..., a_N\}$ is a finite collection of primitive sentences describing *elementary ac-*

---

[8]In section 4 we will be more specific about the reasons for this divergence in results.

*tions* (behavioral events), for example, $a_1$ : "check the passenger's passport", $a_2$ : "search the passenger's bag".

With a slight abuse of terminology, we will use the notation $e_k$ (resp. $a_k$) to indicate both an elementary event (resp. action) and the primitive sentence that describes it. We assume that this language is the (only) common-knowledge language for the parties and the courts. This ensures that there are no problems of ambiguous interpretation of the contract.

A *state* is a complete description of the exogenous environment, represented by a valuation function $s : \Pi^e \rightarrow \{0, 1\}$, where $s(e_k) = 1$ means that primitive sentence $e_k$ is true at state $s$ and $s(e_k) = 0$ means that primitive sentence $e_k$ is false at state $s$.[9] In other words, $s(e_k)$ is a dummy variable that takes value one if elementary event $e_k$ occurs and zero otherwise, and a state is a realization of the vector of dummy variables $(s(e_1), s(e_2), ....)$.

Similarly, a *behavior* is a complete description of all elementary actions, represented by a valuation function $b : \Pi^a \rightarrow \{0, 1\}$, where $b(a_k) = 1$ means that elementary action $a_k$ is executed, and $b(a_k) = 0$ that $a_k$ is not executed.

We assume a very simple payoff structure. There is a one-to-one correspondence between elementary actions and elementary events. The principal wants action $a_k$ to be performed if and only if elementary event $e_k$ occurs. In our airport example, the principal wants the agent to check the passenger's passport if and only if the passenger has a moustache, and to search his bag if and only if the bag is red.

Principal and agent are risk neutral. The principal gets an incremental benefit of $\pi_k$ from "matching" $s(e_k)$ with $b(a_k)$, while he gets zero incremental benefit if there is a "mismatch". Formally, the principal's per-period payoff gross of writing costs is:

$$\pi(s, b, m) = \sum_{k=1}^{N} \pi_k I_k(s, b) - m \tag{2.1}$$

where $m$ is the payment to the agent and $I_k(s, b) = s(e_k)b(a_k) + (1 - s(e_k))(1 - b(a_k))$ is a dummy variable that takes value one if there is a match between event $e_k$ and action $a_k$, and zero if there is a mismatch.

---

[9]To simplify the exposition we describe the basic notation omitting time subscripts. We will introduce time subscripts later in this section, when we describe the dynamic aspects of the game.

The agent's interests are always in conflict with the principal's, in the sense that his preferred actions are always opposite the principal's preferred actions. Formally, the agent's one-period utility is:

$$U(s, b, m) = m - \sum_{k=1}^{N} d_k I_k(s, b). \tag{2.2}$$

We will often refer to the job of matching action $k$ with event $k$ as "task" $k$. The parameter $d_k$ thus captures the agent's disutility from performing task $k$. We let $\mathbf{N} = \{1, ..., N\}$ denote the set of tasks and we use bold capital letters to denote subsets of tasks, as in $\mathbf{K} \subseteq \mathbf{N}$.[10]

The parties' reservation payoff is zero. Assuming $0 < d_k < \pi_k$ for all $k$, the parties' joint surplus (gross of writing costs) is maximized when the agent performs all tasks $k \in \mathbf{N}$. We will refer to this as the *first best* outcome. Payoffs are common knowledge to the contracting parties, and the state and the parties' behavior are verifiable in court. Thus, there are no issues of moral hazard or adverse selection. We assume that preferences and realized payoff levels are not verifiable in court, and that the principal cannot "sell the activity" to the agent (i.e., the agent cannot be made the recipient of the revenue $\pi$).[11]

Next we define a contract. A contract is a pair $(\mathbf{g}, m)$ where $\mathbf{g} = (g_k)_{k \in \mathbf{N}}$ is a set of $N$ clauses and $m$ is a transfer from the principal to the agent (wage). Each clause $g_k$ regulates a task. Given our simple matching structure between actions and events, we can restrict our attention to four types of clause: (i) a contingent clause, constraining the agent to do $a_k$ if and only if $e_k$ occurs, $C_k : [a_k \leftrightarrow e_k]$; (ii) a noncontingent positive clause, constraining the agent to do $a_k$ whatever happens, $R_k : [a_k]$; (iii) a noncontingent negative clause, constraining the agent to do *not* $a_k$ whatever happens, $\bar{R}_k : [\neg a_k]$; (iv) the empty clause, $D$ (for discretion), that imposes no constraint on the agent. For example, if $N = 3$, the set of clauses $(R_1, D, C_3)$

---

[10]We could consider more general payoff structures where the interests of principal and agent may not always be in conflict. For example, we could allow for the possibility that the agent incurs a cost $d_k$ if and only if he takes action $a_k$, and the principal gets a benefit $\pi_k$ if and only if the agent takes action $a_k$ when $e_k$ occurs. With more general payoffs the notation would be more complicated but nothing substantial would change in the results.

[11]If preferences were verifiable, the first-best outcome could trivially be achieved by a contract of the form "The agent's behavior must maximize the sum of the parties' utilities." On the other hand, if realized payoff levels were verifiable, the first-best outcome could be achieved by offering the agent a transfer that increases one-for-one with the principal's realized payoff level. And selling the activity to the agent would be equivalent to specifying a contingent transfer as in the previous point.

constrains the agent to do $a_1$ whatever happens and to do $a_3$ if and only if elementary event $e_3$ occurs, leaving the agent free with regard to task 2. Note that, since we include the empty clause among the possible clauses, there is no loss of generality in assuming that the number of clauses in the contract is $N$.[12]

The parties interact for $T$ periods (where $T$ may be infinite) and have common discount factor $\delta \in (0,1)$. The parameter $\delta$ can also be interpreted as capturing the stability of the relationship.[13] Within each period $t$, the timing is the following: the state of nature $s_t$ is observed; then the principal offers a contract $(\mathbf{g}_t, m_t)$ to the agent, incurring the associated writing costs; if the contract is accepted, the principal makes payment $m_t$ and then the agent acts (being constrained by the contract).

We assume that elementary events are independent of each other, and that each elementary event is governed by a Markov process. Let the transition probabilities for event $e_k$ be $\Pr[s_t(e_k) = 1 | s_{t-1}(e_k) = 1] = p_k$ and $\Pr[s_t(e_k) = 0 | s_{t-1}(e_k) = 0] = q_k$. Without loss of generality we define labels so that $p_k \geq q_k$ for all $k$. We assume that the process exhibits non-negative persistence, that is, $q_k \geq 1/2$ for all $k$. Note that this allows for the case of $i.i.d.$ process ($p_k = q_k = \frac{1}{2}$). Assuming nonnegative persistence simplifies the analysis without affecting the main qualitative results.[14] Finally we assume that the initial state is $s_1(e_k) = 1$ for all $k$; the extension to the more general case in which also the initial state is random is straightforward but tedious.

If at time $t$ the principal wants to offer a different contract than at time $t-1$, he can save on writing costs (whose exact nature will be specified below) by *modifying* the existing contract rather than writing a whole new contract. To capture this idea, we assume that the contract of period $t$ is obtained by amending some of the clauses of the contract in period $t-1$, which serves as a default. This means that it is understood that each contractual clause from the previous period is still effective in the current period proposal (and hence enforced if the

---

[12]We note that in this model there is nothing to gain from making wages contingent on the state or on the agent's behavior.

[13]The parameter $\delta$ can be interpreted as the composition of two parameters, $\delta = \rho\delta'$, where $\rho$ is the probability that the game will continue and $\delta'$ is the 'true' discount factor.

[14]We could consider a general stochastic matrix for each $s(e_k)$, but this would not add much to the qualitative insights of the model.

proposal is accepted) unless it is amended.

More formally, the set of *effective* clauses at time $t$ is given by:

$$\mathbf{g}_t = f(\mathbf{g}_{t-1}, \mathbf{g}_t^A) = \left( (g_{k,t-1})_{k \in \mathbf{N} \setminus \mathbf{K}_t^A}, (\alpha_{k,t})_{k \in \mathbf{K}_t^A} \right),$$

where $\alpha_{k,t} \in \{C_k, R_k, \overline{R}_k, D\}$ $(\alpha_{k,t} \neq g_{k,t-1})$ is the amendment for task $k$ and $\mathbf{g}_t^A = (\alpha_{k,t})_{k \in \mathbf{K}_t^A}$ is the set of amendments.

We can now describe the costs of writing contracts. Writing primitive sentences is costly. We allow for a simple form of dynamic writing economies: writing a given primitive sentence $\xi \in \Pi^a \cup \Pi^e$ for the first time is (weakly) more costly than writing it the subsequent times. If $c(\xi)$ denotes the cost of writing $\xi$ for the first time, we assume that the cost of writing $\xi$ each subsequent time is $r \cdot c(\xi)$, where $r \in (0, 1]$. Specifying the wage and writing logical connectives (such as $\neg$ or $\rightarrow$) in the contract is costless.

In each period $t$ the principal incurs the costs of modifying the previous contract, that is, the writing costs of the amendments $\mathbf{g}_t^A$.[15] These costs (which are history-dependent, due to the recalling economies) can be derived using the assumptions above. Focusing on task $k$, there are only a few relevant possibilities that we need to consider:

(i) Writing a contingent clause $C_k$ at time $t = 1$ costs $c(a_k) + c(e_k)$;

(ii) Writing a noncontingent clause $(R_k$ or $\overline{R}_k)$ at time $t = 1$ costs $c(a_k)$.

(iii) Replacing clause $R_k$ with clause $\overline{R}_k$ or *viceversa* costs $r \cdot c(a_k)$, since this involves recalling an already-described action.

(iv) Replacing a noncontingent clause $(R_k$ or $\overline{R}_k)$ with a contingent clause $(C_k)$ costs $r \cdot c(a_k) + c(e_k)$, since this involves describing a new elementary event and recalling an already-described action.

(v) The empty clause $D$ of course involves no cost. Also, removing a clause (i.e. replacing it with the empty clause $D$) is costless.[16]

---

[15]Note that writing a clause at $t = 1$ can be seen as a modification of the empty clause $D$, which is the default clause at $t = 0$.

[16]Introducing a cost of removing clauses would make our notation heavier without changing our results.

We emphasize that in each period the contract is written *after* the state is observed. Thus, writing a contingent contract would not make sense if the parties interacted for just one period. But with repeated interaction, as we will see shortly, a contingent contract may be efficient.

A final note before we start with the analysis. In the literal interpretation of the game, the parties sign a one-period contract in each period. However, we could equivalently assume that contracts are open-ended, in the sense that the existing contract is automatically renewed unless modified. Thus, since a contingent contract is never modified, we can interpret it as a contract that is signed once and for all at $t = 1$.[17]

## 3. Formal contracting

In this section we focus on situations where parties rely entirely on formal contracting. A simple way to capture these situations is to assume that $T$ is finite, because in this case backward induction implies that the agent "shirks" on any task that is not covered by the formal contract, and hence informal contracting is ruled out.

To simplify computations we will consider the limit of the game as $T$ approaches infinity.[18] However, we note that the qualitative results of this section would be preserved for any game with at least three periods, therefore they can be interpreted as applying also to relationships that have a relatively short duration. The analysis of this section can also be interpreted as applying to relationships where the horizon is infinite but the parties for some reason cannot coordinate on more efficient equilibria, where informal contracts are sustained by the threat of reverting to a worse equilibrium.

Another reason for analyzing the finite-$T$ game is that it allows us to focus sharply on the tradeoffs between different modes of formal contracting (e.g. spot vs. contingent contracting)

---

[17]Note that we do not consider the possibility of multiperiod contracts (the current contract constrains players only for the current period), and the principal is not allowed to make payments in excess of the wage specified in the formal contract. At the end of section 4 we will argue that both of these restrictions are without loss of generality in this model. We will also argue that our results would not change if the payment $m$ were made after the agent acts.

[18]In our game the subgame perfect equilibrium of the finite-horizon game is generically unique, and if there are more than one they are payoff-equivalent. For this reason we speak of "the" equilibrium.

without the confounding effects due to the interaction between formal and informal contracting.

In our model, the limit of the sequence of subgame perfect equilibria as $T \to \infty$ is the Markov perfect equilibrium of the infinite-horizon game. A Markov perfect equilibrium (MPE) is a subgame perfect equilibrium (SPE) where strategies depend only on the payoff-relevant state variable (see Fudenberg and Tirole, 1991, Ch. 13). The payoff-relevant state variable has four components: the current state of the environment $s_t$, the set of clauses used in the previous period, $\mathbf{g}_{t-1}$, the set of tasks $\mathbf{M}_{t-1}^a$ for which the corresponding elementary action $(a_k)$ has been described in the past, and the set of tasks $\mathbf{M}_{t-1}^e$ for which the corresponding elementary event $(e_k)$ has been described in the past.[19] We will denote the overall state variable by $X_t = (s_t, \mathbf{g}_{t-1}, \mathbf{M}_{t-1}^a, \mathbf{M}_{t-1}^e)$. Notice that, since the agent's current choice has no impact on future states, Markov strategies do not depend on past actions of the agent. This implies that in the MPE the agent minimizes his disutility in every period taking the inefficient action for every task not covered by the contract.

In the MPE, the wage $m_t$ is set at the minimum level that induces the agent to accept the proposed contract. Since the determination of the wage is a trivial aspect of the analysis, we will focus on the set of clauses. Solving for the MPE boils down to maximizing the expected discounted value of the surplus net of writing costs.

To state the problem formally, define a *formal contracting policy* as a function of the form $X_t \mapsto \mathbf{g}_t^A = \psi(X_t)$. This function induces, for each $(t, \tau, X_t)$ $(t, \tau \geq 0)$, a random value for the surplus net of writing costs at date $t + \tau$, which we denote $\boldsymbol{\sigma}_{t+\tau}^{\psi} | X_t$. The problem can then be stated as

$$\forall X_t, \forall \tau \geq 0, \quad \max_{\psi} \mathrm{E} \left[ \sum_{\tau=0}^{\infty} \delta^{\tau} \boldsymbol{\sigma}_{t+\tau}^{\psi} | X_t \right] \tag{3.1}$$

The *optimal* contracting policy is the solution to problem (3.1).

Most of the interesting points can be brought out by considering strictly positive but relatively small writing costs. We assume that writing costs are sufficiently small that the optimal contracting policy implements the first best outcome, or in other words, "complete" contracting

---

[19]Formally, $\mathbf{M}_t^a$ is defined as follows: $\mathbf{M}_0^a = \emptyset$ and $\mathbf{M}_t^a = \mathbf{M}_{t-1}^a \cup \{k : g_{k,t} \neq D\}$, and $\mathbf{M}_t^e$ is defined as follows: $\mathbf{M}_0^e = \emptyset$ and $\mathbf{M}_t^e = \mathbf{M}_{t-1}^e \cup \{k : g_{k,t} = C_k\}$. We will sometimes refer to these as the "memory sets".

is optimal. A simple sufficient condition for this is:

$$\forall k, \ c(e_k) + c(a_k) < \pi_k - d_k \tag{C}$$

Later in the paper we will discuss how results change when writing costs are large, so that "incomplete" contracting may be optimal.

Given our assumptions, we can derive the optimal contracting policy by looking separately at each task $k$. For this reason, to keep the exposition simple we will drop the task subscript $k$ for the remainder of this section. We will reintroduce it when we examine informal contracting, as there will be nontrivial interactions among tasks.

It can be shown that there is no loss of generality in focusing on the following candidate *rules* for a given task:

(i) At time $t = 1$ a $C$ clause is written, and it is never modified. We will refer to this as a *contingent* rule, and denote it $\mathcal{C}$.

(ii) At time $t = 1$ a noncontingent clause $R$ is written, and it is amended every time the realization of $s_t$ changes. We refer to this as a *spot* rule, and denote it $\mathcal{S}$.

(iii) At $t = 1$ a clause $R$ is written, and it is permanently replaced by a clause $C$ the first time $s_t(e) = 0$ occurs. We refer to this as an *enrichment* rule, and denote it $\mathcal{E}$.

One can show that no other rule can be strictly optimal. If we dropped our assumption that the Markov process exhibits persistence, other rules could be strictly optimal, but the essence of the results would not change.[20] The following result characterizes the optimal rule for a given task.

**Proposition 1.** *Let* $\gamma = \frac{c(e)}{c(a)}$. *Then, under assumption (C) the optimal rule is* $\mathcal{C}$ *if* $\gamma$ *is small,* $\mathcal{S}$ *if* $\gamma$ *is high and* $\mathcal{E}$ *if* $\gamma$ *is intermediate.* $\mathcal{E}$ *is strictly optimal for some value of* $\gamma$ *if and only if* $p > q$.

---

[20]Note that a *rule* as specified above does not completely determine a *policy*, because it specifies amendments only for reacheable states. For example, the rule saying that clause $C$ has to be written at time $t = 1$ and kept thereafter does not specify what to do after $C$ is removed. This partial description of the optimal policy is sufficient for our purposes.

**Proof.** Since all three rules ($\mathcal{C}$, $\mathcal{E}$, and $\mathcal{S}$) implement the first best, we only need to compare the PDV (present discounted value) of writing costs for each rule, denoted $\widehat{c}(\cdot)$:

$$\widehat{c}(\mathcal{C}) = c(a) + c(e),$$

$$\widehat{c}(\mathcal{E}) = c(a) + [r \cdot c(a) + c(e)] \sum_{t=2}^{\infty} \delta^{t-1} (p)^{t-2} (1-p) = c(a) + [r \cdot c(a) + c(e)] \frac{\delta(1-p)}{1-\delta p},$$

$$\widehat{c}(\mathcal{S}) = c(a) + r \cdot c(a) \sum_{t=2}^{\infty} \delta^{t-1} \Pr[s_t(e) \neq s_{t-1}(e)] = c(a) + r \cdot c(a) \frac{\delta(1-p)[1-\delta(2q-1)]}{(1-\delta)[1-\delta(p+q-1)]}.$$

It follows that (i) $\mathcal{C}$ is preferable to $\mathcal{S}$ iff (if and only if) $\gamma < \gamma^*_{\mathcal{C}/\mathcal{S}} = \frac{r\delta(1-p)[1-\delta(2q-1)]}{(1-\delta)[1-\delta(p+q-1)]}$; (ii) $\mathcal{E}$ is preferable to $\mathcal{C}$ iff $\gamma > \gamma^*_{\mathcal{C}/\mathcal{E}} = \frac{r\delta(1-p)}{1-\delta}$; (iii) $\mathcal{E}$ is preferable to $\mathcal{S}$ iff $\gamma < \gamma^*_{\mathcal{E}/\mathcal{S}} = \frac{r\delta(1-q)[1-\delta(2p-1)]}{(1-\delta)[1-\delta(p+q-1)]}$. Note that $\gamma^*_{\mathcal{C}/\mathcal{E}} \leq \gamma^*_{\mathcal{C}/\mathcal{S}} \leq \gamma^*_{\mathcal{E}/\mathcal{S}}$ iff $p \geq q$, which holds by assumption. Thus, rule $\mathcal{C}$ is optimal iff $\gamma \leq \gamma^*_{\mathcal{C}/\mathcal{E}}$, rule $\mathcal{S}$ is optimal iff $\gamma \geq \gamma^*_{\mathcal{E}/\mathcal{S}}$, and rule $\mathcal{E}$ is optimal iff $\gamma \in [\gamma^*_{\mathcal{C}/\mathcal{E}}, \gamma^*_{\mathcal{E}/\mathcal{S}}]$. $\mathcal{E}$ is strictly optimal for some value of $\gamma$ iff $\gamma^*_{\mathcal{C}/\mathcal{E}} < \gamma^*_{\mathcal{E}/\mathcal{S}}$, which in turn holds iff $p > q$. $\blacksquare$

This proposition highlights a key tradeoff involved in the choice of formal contracts. There are three contracting modes that can be optimal for a given task. The first is to write a contingent clause at the beginning of the relationship. This involves a higher upfront cost, but it does not involve any subsequent costs. The second is to write a rigid clause and modify it when needed; this is the "spot" mode, where the contract is adapted as the environment changes. The advantage of this contracting mode is that there is no need to describe contingencies, but only the agent's actions; the disadvantage is that writing costs must be incurred repeatedly over time. The third possibility (the "enrichment" rule) is one that does not fall squarely into either category of spot or contingent contracting: the contract initially specifies a rigid clause but this is replaced later on by a contingent clause.

The parameter $\gamma$ captures the cost of describing contingencies relative to that of describing actions. It is intuitive that, if $\gamma$ is sufficiently low, a pure contingent approach is optimal, and if $\gamma$ is sufficiently high, a pure spot approach is optimal. As a consequence, the enrichment rule $\mathcal{E}$ can be optimal only for intermediate values of $\gamma$. Proposition 1 also states that $\mathcal{E}$ can be strictly optimal only if $p > q$. To understand this statement note that $\mathcal{E}$ is better than $\mathcal{S}$ only if $c(e)$ is smaller than the expected future costs of amending a rigid clause any time it is needed (the cost of describing the appropriate action at $t = 1$ is the same under both approaches and

hence cancels out). When $p = q$ the two states $s(e) = 0$ and $s(e) = 1$ are symmetric and hence the present value of future costs under rule $\mathcal{S}$ is independent of the current state. If this present value is larger than $c(e)$, then it must be optimal to use a contingent clause from the first period, i.e. it is optimal to use rule $\mathcal{C}$.

Notice that, absent writing costs, the model has little predictive power, because there is a vast multiplicity of optimal contracting plans. Any contracting plan that implements the first best is optimal. These include a complete long-term contingent contract, a sequence of complete spot contracts, and a whole host of intermediate solutions. However, as the analysis makes clear, an arbitrarily small writing cost is sufficient to pin down a unique optimum and deliver strong predictions on the optimal structure of formal contracts.

An interesting implication of the model is that, if the enrichment rule $\mathcal{E}$ is optimal for a nonempty subset of the tasks, the contract becomes gradually more complex over time, as noncontingent clauses get replaced by contingent clauses. As we mentioned in the introduction, this is consistent with a phenomenon that has been documented empirically by Meihuizen and Wiggins (2000) for the case of supply contracts in the natural gas industry.

The model also generates interesting comparative-statics predictions on the impact of changes in the contractual environment on the choice between spot and contingent contracting. The relative importance of spot contracting can be captured by the fraction of tasks regulated by a pure spot rule. Recall that, even though we dropped the task index $k$ from the notation, tasks may be heterogenous with respect to any of the relevant parameters, so the optimal contracting plan may regulate different tasks in different ways.

First we examine the impact of a change in the degree of uncertainty in the environment. A simple way to parametrize the degree of uncertainty in this model is to consider the case of symmetric persistence, that is $p = q$. Then an increase in $p$ decreases the degree of uncertainty (recalling the restriction $p \geq 1/2$).[21] If the $p$ parameter varies across tasks, one can increase (weakly) all the $p_k$s. We find that an increase in uncertainty implies a decrease in the fraction of tasks regulated by spot contracting. The intuition is simple: if uncertainty is higher, the

---

[21]There is also another way to increase uncertainty: one can simply decrease $p_k$ holding $q_k$ constant for each $k$. Our result holds also in this case.

15

external state is expected to fluctuate more over time, and this increases the expected cost of using a spot approach, while the cost of a contingent rule is not affected by uncertainty.

Another parameter that affects the tradeoff between contingent and spot contracting is the discount factor $\delta$. We find that the fraction of tasks regulated by spot contracting is decreasing in $\delta$. The intuitive reason is that a spot approach saves on the cost of describing external events in the first period at the price of paying the cost of describing actions in future periods. The following proposition, whose proof can be found in Appendix together with the proofs of all remaining results, summarizes these comparative statics results:

**Proposition 2.** *Under assumption (C):*
*(i) An increase in uncertainty leads to a decrease in the fraction of tasks regulated by spot contracting.*
*(ii) The fraction of tasks regulated by spot contracting is decreasing in $\delta$.*

This proposition suggests an empirical implication of the model: spot contracting should be relatively more prevalent when the external environment is more stable and when the relationship is less durable. Thus the stability of the external environment and the internal stability of the relationship have opposite effects on the tradeoff between spot and contingent contracting.

Finally it is worth highlighting a simple corollary of the above proposition: if uncertainty varies across tasks, then tasks characterized by a higher degree of uncertainty are more likely to be regulated by contingent clauses, while lower-uncertainty tasks are more likely to be regulated by a spot approach.

### 3.1. Digression 1: the role of language

Next we argue that our language-based approach to writing costs plays an important role for the predictions of the model. To make this point, we consider a more 'traditional' specification of writing costs, similar in spirit to Dye (1985a).

Let $\{s^1, ..., s^M\}$ be the set of states and $\{b^1, ..., b^M\}$ the set of behaviors (where $M = 2^N$), and assign indices so that it is efficient to do $b^j$ if and only if the state is $s^j$, for all $j$. Now

assume that, unlike our model, it is not possible to break down the description of a state or behavior into its elementary constituents. Let $c^s$ be the cost of describing a state and $c^b$ the cost of describing a behavior, and suppose that $c^s$ and $c^b$ are small, so that it is optimal to implement the first best outcome. Keep all other assumptions of our model unchanged.

In this version of the model, if $N$ is not too small, a sequence of spot contracts is optimal, and in particular it dominates a contingent contract. To see this, note that a complete contingent contract must specify the efficient behavior $b^j$ for each state $s^j$ and therefore its cost (paid once and for all) is $2^N(c^s + c^b)$; while a sequence of spot contracts costs at most $c^b$ in each period, therefore its discounted cost does not exceed $\frac{c^b}{1-\delta}$.

Thus, this alternative specification of writing costs implies that spot contracting is typically optimal. This is in stark contrast with our model, where a contingent contract may well be optimal even if $N$ is large; indeed, with our specification of writing costs the optimal Markovian policy is essentially independent of $N$ (recall that the problem is separable in the $N$ tasks). For example, suppose $N \geq \log_2(\frac{1}{1-\delta})$. In this case, under the specification à la Dye, contingent contracting is dominated for any $c^s$ and $c^b$, whereas under our specification contingent contracting is optimal for a whole region of parameters (in particular, if $c(a_k) < \pi_k - d_k$ and $\frac{c(e_k)}{r \cdot c(a_k)}$ is sufficiently small for all $k$).

This should clarify our statement that the nature of language matters for the predictions of the theory. The question of which type of language is more relevant is an open one, but we have argued elsewhere (Battigalli and Maggi, 2002) that a language of the type considered in this paper is likely to be more efficient than a Dye-type language, and that it is closer to the languages that we observe in reality.[22]

---

[22]Another remark about language is in order. We assumed that the language described at the outset is the only common-knowledge language. In principle, the parties could construct a new language, for example by attaching a new primitive sentence to each state and to each behavior, and write a contract with the new language. Note that the parties would have to attach a vocabulary that translates the new language into the original one, in order for the courts to be able to interpret the contract. If the relationship is one-shot, the new language cannot be more efficient than the original one, because the cost of writing the vocabulary in the contract is at least as large as its benefits. In a repeated relationship, however, this approach might in principle be efficient. (We thank Leonardo Felli and Luca Anderlini for bringing this point to our attention.) A more general model would allow for this kind of recoding of the language, but we conjecture that the main qualitative results would not be affected. We already allow for the possibility of "recalling" the description of an action $a$ at low cost; the effect of coding would probably be very similar.

### 3.2. Digression 2: the Maskin-Tirole argument

In a well-known 1999 paper, Maskin and Tirole have argued that unforeseen contingencies (or, by a straightforward extension of their argument, the costs of describing contingencies) do not imply inefficiencies in contracting, provided that an appropriate message-based mechanism is played after contingencies are observed and before actions are taken. Here we argue that the results of our model would not change if such mechanisms were available, and hence the Maskin-Tirole irrelevance argument does not apply to our setting.

The key observation is that a message-based mechanism would offer no advantage relative to a spot contract. To see this, first recall an important aspect of our model that distinguishes it from Maskin and Tirole's model: here there are no ex-ante investment actions. For this reason, in our model spot contracts are sufficient to induce the agent to take the efficient actions. In principle, a message-based mechanism might be useful if it allowed parties to avoid the costs of describing the *actions* to be taken. But can this be the case? The answer is no. Even if a message-based mechanism is played after contingencies are observed, the actions to be taken must still be described formally, that is, in such a way that they can be enforced in court. But if the formal description of actions is subject to the costs that we have assumed in this paper, then a message-based mechanism can never be prefereable to a spot contract. It follows immediately that these mechanisms are redundant in our setting.

## 4. Formal and informal contracting

In reality, long-term relationships are often governed by informal contracts, or by a combination of formal and informal contracts. Informal contracts have an important advantage over formal contracts, namely that they do not require a written description of actions and contingencies, so they save on writing costs.[23] The shortcoming of informal contracts, on the other hand, is the absence of an external enforcement mechanism. Since an informal contract must be self-

---

[23]Arguably, the cost of writing an enforceable contract is higher than the cost of communicating an informal agreement, because for the contract to be enforced by courts it must be written according to the commonly accepted legal standards, which may be quite cumbersome to meet. In particular, it is not sufficient that the language used in the formal contract be common knowledge to the contracting parties; it has to be common knowledge to the parties *and* the courts, and this may require effort and skills (or lawyers)

enforcing, if players are not sufficiently patient it may not be possible to implement the first best outcome with a fully informal contract. In what follows we will examine more closely this trade-off between formal and informal contracting.

Consider the game of section 2. The way we allow for informal contracts is by focusing on the constrained Pareto-efficient SPE of the infinite-horizon game. In such equilibria cooperation may be sustained by formal rules and/or by informal rules: formal rules are enforced by external courts, while informal rules are enforced by the threat of reverting to an equilibrium that is worse for the deviator. The assumption that players focus on a constrained Pareto-efficient SPE is standard in the literature on self-enforcing contracts. This assumption seems reasonable for situations in which there is sufficient mutual understanding between players, so that they are able to coordinate on a "good" equilibrium.

Our objective is to understand under what conditions it is efficient to govern the relationship by formal contracting, by informal contracting, or by a combination of the two, and in the latter case, which tasks are regulated formally and which are regulated informally. Finally, we will explore how some key parameters affect the optimal mix of formal and informal contracting.

The first step in the analysis of efficient SPE is to characterize the optimal punishment strategies. In this type of game it is efficient to punish a deviator with his maxmin payoff, if this punishment is credible. We now show that, if players are sufficiently patient, maxmin punishments are indeed credible. We already know that the MPE keeps the agent at his maxmin. The following lemma identifies a critical level of $\delta$ (which is close to $\frac{1}{2}$ when writing costs are small) above which there exists also a credible punishment strategy that keeps the principal at his maxmin.

**Lemma 1.** *There exists $\delta^*$ (function of other parameters) such that for $\delta \geq \delta^*$ there is a SPE that keeps the continuation payoff of the principal at his maxmin (zero) in every subgame starting with a move by the principal. The critical level $\delta^*$ approaches $\frac{1}{2}$ as writing costs become negligible.*

To convey the basic intuition we focus on the case in which writing costs are negligible and argue that the principal can be kept at her maxmin provided $\delta$ is slightly above $\frac{1}{2}$. Consider

a punishment strategy with the following structure: after a deviation, the informal contract is abandoned and parties revert to the optimal formal contracting plan (the MPE), with the surplus going entirely to the player that has not deviated. If the principal offers a formal contract that she is not supposed to offer, the agent is supposed to reject it unless it gives him the full current net surplus. The key incentive-compatibility condition that this punishment strategy has to satisfy concerns a situation where the principal is being punished. In this situation, the principal has the option of tempting the agent with a "sweet deal" that gives the agent part of the current net surplus. According to the candidate equilibrium strategies, by accepting this offer the agent would lose all the future net surpluses. Neglecting writing costs, a sufficient condition under which the agent will not accept this offer is $\sum_{k \in \mathbf{N}}(\pi_k - d_k) < \frac{\delta}{1-\delta} \sum_{k \in \mathbf{N}}(\pi_k - d_k)$. The left hand side of this inequality is an upper bound to the benefit of accepting a "sweet deal". The right hand side is the present value of future surpluses, which is the opportunity cost of accepting (assuming that the candidate equilibrium strategies are followed from the next period). If $\delta > \frac{1}{2}$ this condition is satisfied, and the proposed punishment strategy is credible.

We note that the condition $\delta \geq \delta^*$ is essential for Lemma 1. Consider the extreme case of $\delta$ equal to zero: then there is a unique SPE (the Markovian equilibrium) in which the principal offers a formal contract and makes a positive profit in every period. However we suspect this condition is not essential for our qualitative results. If it is not satisfied, it may not be possible to keep the principal at his maxmin payoff in the punishment phase, in which case the principal's incentive constraints will be more stringent, and this is likely to result in fewer tasks being regulated informally, but our qualitative results are still likely to hold.

Having discussed the off-equilibrium-path strategies, we can now focus on the equilibrium path. We continue to assume that writing costs are small, in the sense that condition (C) holds. Under this condition, it can be shown that an efficient equilibrium must implement the first best, just as in the formal-contracting analysis. It is then natural to focus on simple equilibrium paths where (i) each task $k$ is either regulated by one of the first-best formal rules described in the previous section ($\mathcal{C}_k$, or $\mathcal{E}_k$ or $\mathcal{S}_k$), or by an *informal* contingent rule prescribing the first-best action, and (ii) the agent always accepts the proposed formal contract $(\mathbf{g}_t, m_t)$ and takes the first-best action for each task regulated by an informal rule.[24]

---

[24]Recall that, given the assumed rules of the game, the principal must pay the exact wage specified in the

Such an equilibrium path is described by a tuple $(\mathbf{I}, \mathbf{C}, \mathbf{E}, \mathbf{S}, (m_t)_{t \geq 1})$, where bold capital letters denote subsets of tasks, and $m_t$ is a random variable, that is, $m_t : H_t \to \mathbb{R}$, where is $H_t$ the set of histories of shocks $h_t = (s_1, ..., s_t)$.[25] $(\mathbf{I}, \mathbf{C}, \mathbf{E}, \mathbf{S})$ is an ordered partition of the set of tasks $\mathbf{N} = \{1, ..., N\}$ specifying that tasks in subset $\mathbf{I}$ are agreed-upon informally and tasks in $\mathbf{C}$ (respectively, $\mathbf{E}, \mathbf{S}$) are regulated formally *via* a contingent (respectively, an enrichment, or spot) rule. We will refer to a tuple $(\mathbf{I}, \mathbf{C}, \mathbf{E}, \mathbf{S})$ as a *task partition.*

We say that a task partition is *incentive compatible* if it is part of a SPE. We say that a task partition is *efficient* if it is incentive compatible and there is no incentive compatible task partition that yields a higher expected present value of the net surplus. Notice that in our model this is equivalent to the standard definition of constrained Pareto-efficiency.[26]

Next we derive a necessary condition for the incentive compatibility of a task partition. If some tasks are regulated by informal rules, the agent has the opportunity to shirk on those tasks (i.e. increase his current utility by not taking the efficient action). The most effective way to prevent the agent from shirking on the informal tasks is to give him all the net surplus from the second period onward if he does not shirk and to keep him at his maxmin if he shirks. Then the agent will not shirk only if the present value of future expected net surpluses is at least as large as the current benefit from shirking. To express this condition formally, it is convenient to introduce some more notation. Let $\pi(\mathbf{K}) = \sum_{k \in \mathbf{K}} \pi_k$, $d(\mathbf{K}) = \sum_{k \in \mathbf{K}} d_k$ and $\sigma(\mathbf{K}) = \pi(\mathbf{K}) - d(\mathbf{K})$. Thus, $\sigma(\mathbf{K})$ represent the surplus obtained by taking the efficient action for each task $k \in \mathbf{K} \subseteq \mathbf{N}$. Also, let $\widehat{c}_t(\mathbf{C}, \mathbf{E}, \mathbf{S} | h_\tau)$ denote the expected present value of writing costs from time $t$ conditional on history $h_\tau$ $(\tau \leq t)$.[27] Then it is clear that a task partition is incentive compatible only if the following set of aggregate incentive constraints hold:

$$\forall t \geq 1, \forall h_t \in H_t, \qquad \frac{\delta}{1-\delta} \sigma(\mathbf{N}) - \delta \widehat{c}_{t+1}(\mathbf{E}, \mathbf{S} | h_t) \geq d(\mathbf{I}). \qquad \text{(IC)}$$

formal contract, $m_t$. As we argue at the end of this section, in this model there is nothing to gain from paying informal bonuses to the agent.

[25]It should be clear that, even if the wage is state-dependent, it is not written as a contingent wage in the formal contract, but it is written period by period after observing the state $s_t$, so it involves no writing costs.

[26]The expected net surplus is simply the sum of the expected payoffs of principal and agent. Therefore, constrained maximization of the PDV of expected net surplus implies constrained Pareto-efficiency. It is not hard to show that also the converse holds.

[27]Note that, for $t \geq 2$, $\widehat{c}_t(\mathbf{E}, \mathbf{S} | h_\tau)$ depends only on $\mathbf{E}$ and $\mathbf{S}$ because the cost of contingent clauses is entirely paid up front.

The left hand side of (IC) is the expected present value of future net surpluses conditional on history $h_t$. The right hand side of (IC) is the agent's disutility from performing the informal tasks, which is also the benefit from shirking.[28]

If $\delta \geq \delta^*$, so that maxmin punishments are credible, it can be shown that condition (IC) is also sufficient for incentive compatibility. We then obtain the following:

**Lemma 2.** *Suppose condition (C) holds, and $\delta \geq \delta^*$ (where $\delta^*$ is defined as in Lemma 1). Then a task partition is efficient if and only if it solves the following problem*

$$\min_{\mathbf{I},\mathbf{C},\mathbf{E},\mathbf{S}} \widehat{c}(\mathbf{C}, \mathbf{E}, \mathbf{S}) \qquad (P)$$
$$s.t. \ (IC)$$

*where $\widehat{c}(\mathbf{C}, \mathbf{E}, \mathbf{S})$ denote the unconditional PDV of writing costs.*

The solution to problem (P) in some cases may not be unique, for the following reason. Since the set of task partitions is finite, a solution will typically not satisfy the (IC) constraint as an equality. It may therefore happen that switching the rules of some tasks from formal to informal and *viceversa* does not violate (IC) and has no impact on the objective function. This is the case, for example, if tasks $j$ and $k$ have the same writing costs, and at a solution of (P), $j$ is regulated by informal rule and $k$ by contingent rule. This example, however, holds for nongeneric values of the cost parameters (the contingent clauses $C_j$ and $C_k$ must have the same writing cost). Indeed, it can be shown that the solution to (P) is *generically* unique.[29] From now on we take for granted that the assumptions of Lemma 2 hold and that the solution to (P) is unique.

---

[28]Another way to interpret (IC) is the following. Suppose the principal pays a stochastic "efficiency wage" $m_t(h_t)$ which in expectation is just enough to keep the agent from shirking. Then the expected wage must be $\overline{m} = d(\mathbf{N}) + \frac{1-\delta}{\delta}d(\mathbf{I})$. The principal would not renege on the promised wage after the first period if

$$\sigma(\mathbf{N}) - m_{t+1}(h_{t+1}) + \delta\frac{\sigma(\mathbf{N}) - \overline{m}}{1-\delta} - \widehat{c}_{t+1}(\mathbf{E}, \mathbf{S}|h_{t+1}) \geq 0.$$

Taking the expectation conditional on $h_t$ and substituting the value of $\overline{m}$ we obtain (IC), which is therefore a necessary condition for the inequality above. Setting $m_{t+1}(h_{t+1}) = \overline{m} - [\widehat{c}_{t+1}(\mathbf{E}, \mathbf{S}|h_{t+1}) - \widehat{c}_{t+1}(\mathbf{E}, \mathbf{S}|h_t)]$ minimizes the strain on the binding constraint and makes (IC) fully equivalent to the inequality above.

[29]More precisely, it can be shown that the set of parameter vectors for which there are multiple solutions is *nowhere dense* in the parameter space (i.e., its closure has empty interior).

We first highlight some properties of the efficient task partition with respect to the choice between formal and informal contracting. Then we discuss how informal contracting affects the choice between spot and contingent contracting.

The next proposition states conditions under which the optimum is fully informal, or fully formal, or a mix of formal and informal contracting. Let $d^{\min} \equiv \min_k d_k$ .

**Proposition 3.** *At the solution of problem (P):*
*(i) If $d(\mathbf{N}) \leq \frac{\delta}{1-\delta}\sigma(\mathbf{N})$, all tasks are regulated informally.*
*(ii) If $d^{\min} < \frac{\delta}{1-\delta}\sigma(\mathbf{N}) < d(\mathbf{N})$, and writing costs are sufficiently small, some tasks are regulated formally and some are regulated informally.*
*(iii) If $d^{\min} > \frac{\delta}{1-\delta}\sigma(\mathbf{N})$, all tasks are regulated formally.*

This result states that formal and informal contracting coexist if the total disutility incurred by the agent over all tasks is relatively large ($d(\mathbf{N}) > \frac{\delta}{1-\delta}\sigma(\mathbf{N})$) and there is at least one task characterized by a relatively low disutility ($d^{\min} < \frac{\delta}{1-\delta}\sigma(\mathbf{N})$). If the first condition is not satisfied, then a fully informal contract is incentive compatible, and hence efficient; if the second condition is not satisfied, then introducing any of the tasks in the informal contract violates (IC), therefore the optimum is fully formal.[30]

An interesting aspect of this result is that, as long as $d^{\min}$ is relatively low, the optimal contracting plan is partly or fully informal, and this is true even if writing costs are very small. This is in sharp contrast with a result obtained by Baker *et al.* (1994) who show that if the imperfection in formal contracting is small, it is impossible to sustain any informal contracting. This is not due to the differences between their contracting game and ours, but rather to a restriction that they make on the punishment strategy: instead of focusing on maximal punishments, they assume that deviations are punished by reversion to the static Nash equilibrium of the contracting game. In our model, the analogous restriction would be to assume that deviations are punished by reversion to the Markov perfect equilibrium. If writing

---

[30]The reason we need sufficiently small costs at point (ii) is that if noncontingent formal rules are used, then the present value of future writing costs enters the incentive constraint. See the proof of Proposition 3 in the appendix.

costs are sufficiently small, this punishment strategy cannot deter deviations by the principal, hence it cannot support any informal contracting.[31]

An interesting question is: which tasks are regulated formally and which are regulated informally? Intuitively, informal contracting should be harder to sustain for tasks characterized (other things equal) by higher disutility, because the agent has a stronger incentive to shirk on these tasks. The next proposition confirms this intuition, under a *ceteris paribus* assumption on other task characteristics:

**Proposition 4.** *Suppose that writing costs and transition probabilities are the same for all* $k \in \mathbf{N}$. *Then, at the solution to (P), there exists a critical level* $\bar{d}$ *such that tasks with* $d_k < \bar{d}$ *are regulated informally and tasks with* $d_k > \bar{d}$ *are regulated formally.*

The basic argument behind this result is the following. Consider two tasks characterized by different disutilities, and suppose that one task must be regulated formally and the other informally. Which one will be handled informally? The surplus is the same independently of which task is chosen, but the low-$d_k$ task implies a lower incentive to cheat, hence it is better to regulate this task informally.[32]

A natural question is how other task characteristics, and in particular the degree of uncertainty and the benefit $\pi_k$, affect the choice between formal and informal contracting. Consider first the role of uncertainty. Even if tasks differ only in the degree of uncertainty, there is no clear-cut result: comparing two tasks it is possible that the higher-uncertainty task is regulated informally and the lower-uncertainty task is regulated formally, but also the viceversa is possible. This is clear if one considers the case in which all formal clauses are contingent:

---

[31]To see this, note first that if some tasks are regulated informally, the principal must pay an efficiency wage in order to motivate the agent to work (see footnote 28). This creates a temptation for the principal to renege on the wage promised to the agent, and hence the principal has a strictly positive one-time gain from cheating. Second, note that the MPE gives all the surplus to the principal, and if writing costs are small, this is close to the maximum potential surplus. Thus the future loss from punishment for the principal, if any, is negligible, and cannot outweigh his one-time gain from cheating.

[32]Recall that we are assuming that the solution to (P) is unique. While this holds generically in the unrestricted parameter space, it need not hold generically in the subspace where writing costs and transition probabilities are the same. However we can show that, if there is a continuum of tasks, the efficient task partition conforms to the proposition above.

then uncertainty is irrelevant for the choice between formal and informal rules. Also the task-specific benefit $\pi_k$ has no role in the optimal allocation of tasks between formal and informal contracting: even if tasks differ only with respect to $\pi_k$, there is no way of saying which ones are regulated formally and which ones are regulated informally. This is because the $\pi_k$ parameters enter problem (P) only through their sum, $\pi(\mathbf{N})$.[33]

Thus far we have examined the efficient partition of tasks with respect to the choice between formal and informal contracting. The model also offers some insights as to how changes in the parameters affect this choice. The key observation is that any parameter change that relaxes the incentive constraints makes informal contracting easier to sustain. Parameter changes that have this effect are (a) an increase in the total benefit $\pi(\mathbf{N})$; (b) a decrease in uncertainty; (c) a decrease in writing costs; and (d) an increase in the discount factor $\delta$. The impact of $\pi(\mathbf{N})$ is apparent from inspection of (IC): other things equal, an increase in $\pi(\mathbf{N})$ implies an increase in $\sigma(\mathbf{N})$, which relaxes the constraint. A decrease in uncertainty or in writing costs relaxes (IC) because it reduces the present value of future writing costs, $\widehat{c}_{t+1}$. Note that this effect is present only to the extent that some tasks are governed by spot or enrichment rule. An increase in $\delta$ increases the PDV of future gross supluses more than it increases the the PDV of future writing costs.

Any of these parameter changes increases the slack in (IC) *for a given task partition*. However, this does not necessarily mean that such changes increase the optimal *number* of informal tasks, because the heterogeneity of tasks, the discrete nature of problem (P) and the interaction between spot and contingent contracting may cause offsetting effects.[34] We can only state broadly that these parameter changes *tend* to increase the relative importance of informal contracting.

We can say something more precise when tasks are homogeneous with respect to writing costs and transition probabilities (as assumed in Proposition 4) and formal rules are all of the

---

[33]As the next section makes clear, this is no longer true if writing costs are large.

[34]Consider the following example. There are three tasks, tasks 1 and 2 characterized by small disutility and writing costs, task 3 by relatively high disutility and writing costs. Suppose that in the initial situation it is possible to sustain informally task 1 and 2 but not task 3, because of this task's high disutility. If a parameter change relaxes (IC), it is possible that in the new situation it is incentive compatible to regulate informally task 3 alone, but not in conjunction with any other task. Then it is optimal to switch from regulating tasks 1 and 2 informally to regulating only task 3 informally.

same type. Then parameter changes that increase the slack in the incentive constraints increase (weakly) the number of informal tasks, and hence decrease (weakly) the number of formal tasks.

**Remark 1.** *Suppose that writing costs and transition probabilities are the same for all tasks, and consider any of the following parameter changes: (a) an increase in the total benefit $\pi(\mathbf{N})$; (b) a decrease in uncertainty; (c) a decrease in writing costs; and (d) an increase in the discount factor $\delta$. If at the solution of (P) all formal tasks are governed by rules of the same type (before and after the change), then the optimal number of formal tasks (weakly) decreases.*

The effect of an increase in $\pi(\mathbf{N})$ is very intuitive: this increases the gross potential surplus, and thus increases the opportunity cost of deviations. Broadly speaking, this result suggests that informal contracting should be used more intensively in long-term relationships where the gains from contracting are larger.

As for the second result, an increase in uncertainty decreases the expected number of future contract amendments, thus increases the net worth of the relationship in future periods, and hence relaxes the incentive constraint. This in turn makes it informal contracting easier to sustain. This same effect explains the surprising result that a decrease in writing costs may shift the balance in favor of *in*formal contracting. The analysis of a change in the discount factor is subtler: an increase in $\delta$ increases both the PDV of future gross surpluses and the PDV of future writing costs, but it can be shown that the net effect is to relax the incentive constraint.

We now come back to the choice between contingent and spot contracting. In the previous section we analyzed this choice when only formal contracting is available (i.e. at the MPE of the game). Here we ask how this choice changes when informal contracting is available (i.e. at an efficient SPE of the game). Recall that the optimal formal contracting plan minimizes the PDV of writing costs, whereas when informal contracting is available, the cost-minimization problem allows for informal rules and is subject to the (IC) constraint. Intuitively, with a spot approach writing costs must be incurred repeatedly over time, and this imposes a strain on the incentive constraint; whereas with a contingent approach writing costs are incurred only at $t = 1$, and this does not create incentive problems. Thus contingent contracting has better incentive properties

26

than spot contracting. As a consequence, the presence of informal contracting should tend to favor contingent contracting. The following remark confirms this intuition.

**Remark 2.** *In the efficient task partition it is possible that a task $k$ is governed by contingent rule even though a spot rule is less costly $[\hat{c}(\mathcal{S}_k) < \hat{c}(\mathcal{C}_k)]$. The reverse case is not possible.*

This result suggests that the relative importance of contingent versus spot contracting should be higher in relationships where formal and informal contracting coexist, compared with relationships that are governed exclusively by formal contracting.

A related question is how the presence of informal contracting affects the results of the previous section on the choice between spot and contingent contracting. The qualitative insights about the partition of formal tasks for given parameter values are unchanged,[35] but the comparative statics results can change in interesting ways. The new effect is that, when costly contingent rules are used instead of cheaper spot rules to put less strain on the incentive constraint, a parameter change that relaxes the incentive constraint allows a more intensive use of spot rules. This new effect can generate surprising possibilities. In particular we note the following:

**Remark 3.** *An increase in $\pi(\mathbf{N})$ or $\delta$ may increase the fraction of formal tasks governed by spot rule.*

Note that, when only formal contracting was available, increasing $\pi(\mathbf{N})$ had no effect on the tradeoff between spot and contingent contracting, and increasing $\delta$ shifted the balance in favor of contingent contracts. The reason we only state a possibility result ("may") is again that the heterogeneity of tasks and the discreteness of the problem may generate counterintuitive results, but the tendency is clear: parameter changes that relax (IC) alleviate the incentive problems associated with spot contracts, hence facilitate their use, whereas a relaxation of (IC) does nothing to favor contingent contracting.

---

[35]For example, consider the role of uncertainty in the choice between spot and contingent contracting. It can be shown that, if tasks differ only by the degree of uncertainty, a set of low-uncertainty tasks will be regulated by formal spot contracting, and a set of high-uncertainty tasks will be regulated by informal or contingent formal contracting.

Before concluding the section, we want to discuss our assumptions about timing, the possibility to pay "bonuses" (i.e. payments in excess of what is specified by the agreed-upon formal contract) and the possibility to write multi-period contracts.

According to the rules of the game, the principal pays the agent before he acts and the payment is exactly the one specified by the offered contract (if accepted). It is clear that each of these assumptions is without loss of generality given the other. If bonuses are not allowed, it does not matter when the payment $m$ occurs, because the principal has to pay $m$ independently of whether the agent has shirked on the informally regulated tasks or not. On the other hand, if any payment has to be made before the agent acts, bonuses are redundant, because the best incentive to keep the agent from shirking is still to hold him down to his maxmin from period $t+1$ if he shirks in period $t$. We now argue that even the *joint* assumption about timing and bonunes is without loss of generality. Suppose that the principal is allowed to pay an informal bonus immediately after the agent acts. In this case, the agent has a stronger incentive not to shirk, because shirking will prevent him from enjoying an immediate reward. On the other hand, with the modified assumption, the principal has an incentive to renege on the informally promised bonus, whereas with our current assumption he can only omit to offer the formal contract specified by the equilibrium. It can be shown that these two effects cancel out.

Similarly, it can be shown that there is no gain from writing formal multi-period contracts. To gain intuition, suppose there is only one task. Even though committing to future wages in the current contract would remove the principal's incentive to renege on payments, this would ruin the agent's incentives: if future wages are assured, the agent will surely cheat if the task is regulated informally. And if the task is regulated formally, then there is no gain relative to one-period formal contracts.

## 5. Large writing costs

If writing costs are large, the main change in results is that it may not be optimal to implement the first best for some (or all) of the tasks. In particular, two additional possibilities emerge: it might be optimal to regulate a task by *rigid rule*, that is by writing a noncontingent $R_k$ clause

once and for all, or to leave a task to the agent's *discretion* with no informal agreement to take the efficient action. Given our assumptions on payoffs, a rigid rule yields a gross incremental surplus that is positive but lower than a first-best rule, and leaving a task discretionary yields zero gross incremental surplus.

We interpret rigidity and discretion as two forms of contractual incompleteness, while we think of a contract implementing the first-best outcome as a "complete" contract. If one adopts this terminology, then the main implication of large writing costs is that they lead to contractual incompleteness. Next we show how the analysis can be extended to this more general case.

We drop assumption (C). The only condition we assume on writing costs is that they are non-prohibitive, in the sense that for every $\delta \in [0, 1)$ the optimal formal contracting policy yields a strictly positive present value of the net surplus.[36] A contracting plan can now be represented by a (possibly) non exhaustive task partition $(\mathbf{I}, \mathbf{C}, \mathbf{E}, \mathbf{S}, \mathbf{R})$ and a transfer process $(m_t)_{t \geq 1}$, where $\mathbf{R}$ is the set of tasks regulated by rigid rules. The tasks that are not covered by the partition are discretionary.

Let $\hat{\sigma}_t(\mathbf{R}|h_\tau)$ denote the expected present value of future surpluses associated with the set of rigid rules conditional on $h_\tau$ $(\tau \leq t)$, and let $\mathbf{FB} \equiv \mathbf{I} \cup \mathbf{C} \cup \mathbf{E} \cup \mathbf{S}$ denote the set of tasks regulated in a first best way. In this more general case the expected present value of future net surpluses is $\frac{\delta}{1-\delta}\sigma(\mathbf{FB}) + \delta\hat{\sigma}_{t+1}(\mathbf{R}|h_t) - \delta\widehat{c}_{t+1}(\mathbf{E}, \mathbf{S}|h_t)$. By the argument of the previous section, a task partition is incentive compatible only if

$$\forall t \geq 1, \forall h_t \in H_t,$$

$$\frac{\delta}{1-\delta}\sigma(\mathbf{FB}) + \delta\hat{\sigma}_{t+1}(\mathbf{R}|h_t) - \delta\widehat{c}_{t+1}(\mathbf{E}, \mathbf{S}|h_t) \geq d(\mathbf{I}). \tag{IC'}$$

This is a generalization of the incentive constraints in the previous section. The only difference is that now the surplus terms in the left hand side must be calculated taking into account that the tasks regulated by rigid rules yield a stochastic stream of surpluses, and that discretionary tasks yield no surplus at all. With relatively large writing costs, however, it is possible

---

[36]This is equivalent to the following simple condition: there is at least one task $j$ such that $c(a_j) < \pi_j - d_j$. The condition is obviously sufficient, because it allows to obtain a postitive net surplus for task $j$ by using clause $R_j$ or $\overline{R}_j$, as appropriate, in every period. To see that the condition is necessary, suppose that $c(a_k) \geq \pi_k - d_k$ for all $k$ and let $\delta = 0$. Then introducing any (nonempty) clause in the contract would yield a (weakly) negative net surplus in the first period that could not be compensated by positive surpluses in future periods.

that (IC') holds and yet the expected present value of net surpluses conditional on some history of shocks $h_t$ is negative. In this case the task partition cannot be part of a subgame perfect equilibrium. Therefore we have a separate necessary condition:

$$\forall t \geq 1, \forall h_t \in H_t, \ \frac{\sigma(\mathbf{FB})}{1-\delta} + \hat{\sigma}_t(\mathbf{R}|h_t) - \hat{c}_t(\mathbf{E}, \mathbf{S}|h_t) \geq 0. \tag{PC'}$$

Relying on Lemma 1, we can show that (IC') and (PC') are also sufficient for incentive compatibility if the discount factor is high enough. Therefore we obtain the following characterization result:

**Lemma 3.** *Suppose $\delta \geq \delta^*$ (where $\delta^*$ is as in Lemma 1). Then a task partition is efficient if and only if it solves the following problem*

$$\max_{\mathbf{I}, \mathbf{C}, \mathbf{E}, \mathbf{S}, \mathbf{R}} \left[ \frac{1}{1-\delta}\sigma(\mathbf{FB}) + \hat{\sigma}_1(\mathbf{R}) - \hat{c}(\mathbf{C}, \mathbf{E}, \mathbf{S}, \mathbf{R}) \right] \tag{P'}$$
$$s.t. \ (IC') \ and \ (PC').$$

Note that the presence of large writing costs modifies not only the incentive constraints but also the objective function of the problem: since it may not be optimal to implement the first best outcome, we need to consider explicitly the surplus implications of different formal rules, therefore we no longer have a cost-minimization problem.

In general this is a fairly complex problem, but we can say something about the case in which tasks differ only by their disutility:

**Proposition 5.** *Suppose tasks differ only with respect to $d_k$. Then, at the solution to (P'), there exist three critical levels, $d' < d'' < d'''$, such that tasks with $d_k < d'$ are regulated by informal rule, tasks with $d' < d_k < d''$ are regulated by formal first-best rule, tasks with $d'' < d_k < d'''$ are regulated by formal rigid rule, and tasks with $d_k > d'''$ are left to the agent's discretion.*

This proposition highlights that the main consequence of large writing costs is that the optimal contracting plan may include rigidity and/or discretion. The logic that drives this

result is simple. Putting aside informal contracting for a moment, let us ask: which tasks should be regulated by first-best rule, which by rigid rule, and which should be discretionary? Note that, since the benefit $\pi_k$ is assumed constant across tasks, lower $d_k$ implies higher gross surplus ($\sigma_k = \pi_k - d_k$). Intuitively, high-surplus tasks are regulated by a first best rule, intermediate-surplus tasks are handled with a rigid rule, and low-surplus tasks are left to the agent's discretion. This part of the result is analogous to Proposition 1 in Battigalli and Maggi (2002), and is a consequence of the fact that the writing cost is higher for a first-best rule than for a rigid rule, and is zero for a discretionary task. Now consider also the possibility of informal contracting. As we saw in the previous section, with small writing costs, lower disutility favors informal contracting over formal contracting. This continues to be true with large writing costs, hence the ranking described in Proposition 5 follows.

Another interesting point that is brought out in the presence of large writing costs is the following. Formal and informal contracting not only tend to be used jointly, but are *complementary* in a stricter sense: when some tasks are discretionary, increasing the number of tasks regulated formally (for example as a consequence of a reduction in writing costs) increases the surplus from the relationship, thus relaxing the incentive constraint and making it easier to regulate more tasks informally.[37]

We conclude with a note on the nature of writing costs in our model. We considered only "variable" writing costs, that is costs that increase with the number of modifications to the contract. One may also consider the impact of "quasi-fixed" writing costs, i.e. costs that are incurred every time the contract is modified. Quasi-fixed writing costs would affect the mode of formal contracting in two ways: first, they would tilt the balance in favor of contingent contracting as opposed to spot contracting, since the latter involves contract modifications over time and the former does not; second, if spot contracting is nevertheless efficient, there will be a tendency to postpone some contract modifications so as to "cluster" several modifications in a single period. This would complicate the analysis without adding much insight.

---

[37]A similar complementarity between formal and informal contracting is highlighted by Baker *et al.* (1994).

# 6. Conclusion

In this paper we have implicitly assumed away an alternative mode of governance that could avoid the costs of writing detailed contracts, namely giving *authority* to the principal. If the principal could instruct the agent on what actions to take in each period, there would be no need to specify contingencies or actions in a contract. In this concluding section we discuss how results would change if we allowed for authority as a governance mode.

As a premise, it is useful to distinguish between *formal* and *informal* authority. We speak of formal authority when the principal's authority is specified in the formal contract, and the agent can be punished by courts for disobeying the principal. We speak of informal authority when the principal's orders are not enforced by courts, but by credible punishment mechanisms.

Let us focus on formal authority first. It is critical to note that formal authority is enforceable only if two conditions are met: (i) the principal can send verifiable messages to the agent; this requires that messages be written, or at least recorded; and (ii) messages must be expressed in a language understood by the courts. In other words, messages must be *formal*. For this reason, even if a system of formal messages is feasible, it is not clear that its costs would be significantly lower than a system of formal contracts.[38]

Informal authority is a more common mode of governance in real organizations. In our model, however, there is no role for such an arrangement, due to the assumption of symmetric information. Given that the principal and the agent have the same information, informal authority cannot improve on an informal contract as we defined it, because in the latter arrangement the agent knows what actions to take under any contingency, hence there is no need for further instructions from the principal. A role for informal authority would probably arise if the principal had private information). An extension of the model in this direction is left for future research.

---

[38]In their analysis of construction contracts, McLeod and Chakravorty (2004) consider an environment where formal authority may be optimal. An essential feature that distinguishes their model from ours is that external shocks are observed after the contract is signed.

# 7. Appendix

**Proof of Proposition 2**

(i) In the symmetric-persistence case the relevant thresholds for the parameter $\gamma = \frac{c(e)}{c(a)}$ (defined in the proof of Proposition **1**) collapse to the same value $\gamma^*_{\mathcal{C}/\mathcal{S}} = \gamma^*_{\mathcal{E}/\mathcal{S}} = \gamma^*_{\mathcal{C}/\mathcal{E}} = \frac{r\delta(1-p)}{(1-\delta)}$, which is decreasing in $p$. By Proposition 1, the number of tasks regulated by a spot rule way is therefore decreasing as uncertainty increases ($p$ decreases).

[Using the general expressions for $\gamma^*_{\mathcal{C}/\mathcal{S}}$ and $\gamma^*_{\mathcal{E}/\mathcal{S}}$ it is also easy to check that these thresholds are decreasing in $p$ keeping $q$ fixed. This proves the claim made in footnote that the number of tasks regulated by a spot rule is increasing if, for all tasks $k$, $p_k$ decreases with $q_k$ fixed, which is another way to increase uncertainty.]

(ii) We verify that the threshold $\gamma^*_{\mathcal{E}/\mathcal{S}}$ is increasing in $\delta$:

$$\frac{\partial \gamma^*_{\mathcal{E}/\mathcal{S}}}{\partial \delta} = r(1-q)\frac{[1 - \delta(2p - 1)]\left\{(1-\delta) + \delta(1-q)[1 - \delta(p + q - 1)]\right\}}{(1-\delta)^2[1 - \delta(p + q - 1)]^2} > 0.$$

By Proposition 1, it follows that the number of tasks regulated by a spot rule is decreasing in $\delta$. ∎

**Proof of Lemma 1**

Let $\underline{v}(\delta)$ be the minimum present value of net surpluses in the Markov perfect equilibrium with respect to all possible states of the system, i.e.

$$\underline{v}(\delta) = \min_{\mathbf{g},\mathbf{M},s} v(\mathbf{g},\mathbf{M},s;\delta),$$

where $v(\mathbf{g},\mathbf{M},s;\delta)$ is the present value of net surpluses when the set of clauses from the previuous period is $\mathbf{g}$, the memory is $\mathbf{M}$ and the current state of nature is $s$, given discount factor is $\delta$. Function $\underline{v}(\delta)$ is nondecreasing in $\delta$, and $(1-\delta)\underline{v}(\delta)$ it is bounded below by $\sum_{k:c(a_k)<\pi_k-d_k}[\pi_k - d_k - c(a_k)]$, which is strictly positive because we assume that $c(a_k) < \pi_k - d_k$ for at least some task $k$. This implies that $\delta\underline{v}(\delta)$ converges monotonically to $\infty$ as $\delta \to 1$. Therefore there is exactly one value of $\delta$ in $(0,1)$ solving the following equation:

$$\sigma(\mathbf{N}) = \delta\underline{v}(\delta). \tag{7.1}$$

Let $\delta^*$ be the solution to (7.1). Then $\sigma(\mathbf{N}) \leq \delta \underline{v}(\delta)$ for each $\delta \geq \delta^*$.

Since

$$\sigma(\mathbf{N}) = \delta^* \underline{v}(\delta^*) \geq \frac{\delta^*}{1 - \delta^*} \left[ \sigma(\mathbf{N}) - \sum_{k \in \mathbf{N}} c(a_k) \right],$$

it follows that $\delta^* \to \frac{1}{2}$ as $c(a_k) \to 0$ for all $k$.

From now on we keep $\delta$ fixed and therefore we suppress in our notation the dependence of values on $\delta$. We exhibit a SPE keeping the principal at his maxmin under the parameter restriction $\delta \geq \delta^*$. Consider the following strategies: for all Markov states $(\mathbf{g}, \mathbf{M}, s)$ the principal makes amendments as in the MPE. Wages are determined according to a punishment phase.[39] There are two punishment phases $\mathcal{P}_P$ and $\mathcal{P}_A$. The system starts in phase $\mathcal{P}_P$. When the system is in phase $\mathcal{P}_P$ the (offered) wage is the net profit generated by the offered contract. When the system is in phase $\mathcal{P}_A$ the (offered) wage is the disutility generated by the offered contract. As soon as player $i$ deviates from his strategy the system switches immediately to phase $\mathcal{P}_i$. If the system is in phase $\mathcal{P}_A$ the agent accepts the offered contract (and chooses the one-shot best response). Thus, in phase $\mathcal{P}_A$ the MPE is played. If the system is in phase $\mathcal{P}_P$, the new set of default clauses and memory is $(\mathbf{g}', \mathbf{M}')$, the offered contract is $(\mathbf{g}', m)$ and the state of nature is $s$, then the agent accepts if and only if

$$m - d(\mathbf{g}', s) > \delta v(\mathbf{g}', \mathbf{M}'|s), \tag{7.2}$$

where $\delta v(\mathbf{g}', \mathbf{M}'|s)$ is the expected present value of the of the net surpluses generated by the MPE starting from next period at $(\mathbf{g}', \mathbf{M}')$ given the current shocks $s$, and $d(\mathbf{g}', s)$ is the disutility induced by the set of clauses $\mathbf{g}'$ given $s$.

By construction, the agent has no incentive to deviate. In particular, suppose that the state of nature is $s$, the principal moves the set of clauses and memory to $(\mathbf{g}', \mathbf{M}')$ and offers $(\mathbf{g}', m)$ so that, as a consequence, the system enters (or stays in) phase $\mathcal{P}_P$.

If $m - d(\mathbf{g}', s) \leq \delta v(\mathbf{g}', \mathbf{M}'|s)$, the agent is supposed to reject. The expected payoff if the agent conforms is $\delta v(\mathbf{g}', \mathbf{M}'|s)$. The expected payoff of a one-shot deviation is $m - d(\mathbf{g}', s)$, because after the deviation the system enters phase $\mathcal{P}_A$ where the agent gets his maxmin (zero). Therefore rejection is indeed a best response.

---

[39]Implicitly, we describe strategies as finite automata.

If $m - d(\mathbf{g}', s) > \delta v(\mathbf{g}', \mathbf{M}'|s)$, the agent is supposed to accept and this is obviously a best response.

We now check that the principal has no incentive to deviate in phase $\mathcal{P}_P$. The only way for the principal to make a profitable one-shot deviation is to offer a contract $(\mathbf{g}', m)$ (by way of appropriate amendments and exceptions) satisfying the acceptance condition (7.2). Observe that by definition $v(\mathbf{g}', \mathbf{M}'|s) = \sum_{s'} \Pr[s'|s] v(\mathbf{g}', \mathbf{M}', s') \geq \underline{v}$. Therefore the net payoff the principal can get by "tempting" the agent is bounded above by

$$\max_{\mathbf{g}', \mathbf{M}'} [\sigma(\mathbf{g}', s) - \delta v(\mathbf{g}', \mathbf{M}'|s)] \leq \sigma(\mathbf{N}) - \delta \underline{v},$$

where $\sigma(\mathbf{g}', s)$ is the surplus induced by $\mathbf{g}'$ at state $s$. Recall that $\sigma(\mathbf{N}) - \delta \underline{v} \leq 0$ because we assumed $\delta \geq \delta^*$. Therefore the principal has no profitable one-shot deviation. $\blacksquare$

**Proof of Lemma 2**

It can be shown that under assumption (C) (small writing costs) at an efficient task partition all tasks are regulated by first best (formal or informal) rules. Furthermore, (C) implies (PC'). Therefore the result follows as a corollary of Lemma 3, which is proved just below.$\blacksquare$

**Proof of Lemma 3**

We first prove that (IC') and (PC') are necessary for incentive compatibility. Suppose that the tuple $(\mathbf{I}, \mathbf{C}, \mathbf{E}, \mathbf{S}, \mathbf{R}, (m_t)_{t \geq 1})$ is part of a SPE. Then it must be the case that the present value of the principal's expected profits is always (weakly) positive and that the agent's expected utility from following $(\mathbf{I}, \mathbf{C}, \mathbf{E}, \mathbf{S}, \mathbf{R}, (m_t)_{t \geq 1})$ is (weakly) larger than what the agent can get by accepting the formal contract offered by the principal, "shirking" on the tasks in $\mathbf{I}$ and rejecting all future offers. To write these incentive constraints in a relatively simple form, define

$$\widehat{m}_{t+1}(h_t) := \sum_{k=1}^{\infty} \delta^{k-1} \mathrm{E}(m_{t+k}|h_t)$$

and let $\mathbf{FB} = \mathbf{I} \cup \mathbf{C} \cup \mathbf{E} \cup \mathbf{S}$, $\mathbf{FFB} = \mathbf{C} \cup \mathbf{E} \cup \mathbf{S}$ denote, respectively, the set of tasks governed by first best rules and the set of tasks governed by *formal* first best rules. Then for all $t \geq 0$, all $h_{t+1} = (h_t, s_{t+1})$

$$\frac{\pi(\mathbf{FB})}{1 - \delta} + \widehat{\pi}_{t+1}(\mathbf{R}|h_t, s_{t+1}) - \widehat{c}_{t+1}(\mathbf{C}, \mathbf{E}, \mathbf{S}|h_t, s_{t+1}) \geq m_{t+1}(h_t, s_{t+1}) + \delta \widehat{m}_{t+2}(h_t, s_{t+1}), \quad (\mathrm{IC}_P^{t+1})$$

and for all $t \geq 1$, all $h_t = (h_{t-1}, s_t)$

$$m_t(h_t) + \delta \widehat{m}_{t+1}(h_t) - \frac{d(\mathbf{FB})}{1-\delta} - \widehat{d}_t(\mathbf{R}|h_t) \geq m_t(h_t) - d(\mathbf{FFB}) - d(\mathbf{R}, s_t), \qquad (\text{IC}_A^t)$$

where $\widehat{\pi}_{t+1}(\mathbf{R}|h_{t+1})$ is the expected present value of gross benefits for tasks in $\mathbf{R}$ at date $t+1$ conditional on $h_{t+1}$, $\widehat{d}_t(\mathbf{R}|h_t)$ is the expected present value of disutilities from tasks in $\mathbf{R}$ at date $t$ conditional on $h_t$, and $d(\mathbf{R}, s_t)$ is the disutility of the tasks regulated by rigid rules at $s_t$. Incentive constraint $(\text{IC}_A^t)$ can be written as

$$\widehat{m}_{t+1}(h_t) \geq \frac{1}{1-\delta}\left[ d(\mathbf{FFB}) + \frac{d(\mathbf{I})}{\delta} \right] + \widehat{d}_{t+1}(\mathbf{R}|h_t)$$

[note that $\widehat{d}_t(\mathbf{R}|h_t) - d(\mathbf{R}, s_t) = \delta \widehat{d}_{t+1}(\mathbf{R}|h_t)$]. Taking the expected value of both sides of $(\text{IC}_P^{t+1})$ w.r.t. $s_{t+1}$ (conditional on $h_t$) and combining with the above inequality we obtain

$$\frac{\pi(\mathbf{FB})}{1-\delta} + \widehat{\pi}_{t+1}(\mathbf{R}|h_t) - \widehat{c}_{t+1}(\mathbf{C}, \mathbf{E}, \mathbf{S}|h_t) \geq \widehat{m}_{t+1}(h_t) \geq \frac{1}{1-\delta}\left[ d(\mathbf{FFB}) + \frac{d(\mathbf{I})}{\delta} \right] + \widehat{d}_{t+1}(\mathbf{R}|h_t),$$

which yields (IC'):

$$\frac{\delta}{1-\delta}\sigma(\mathbf{FB}) + \delta\widehat{\sigma}_{t+1}(\mathbf{R}|h_t) - \delta\widehat{c}_{t+1}(\mathbf{C}, \mathbf{E}, \mathbf{S}|h_t) \geq d(\mathbf{I}).$$

The agent has no incentive to reject the proposed contract if for all $t \geq 1$ and all $h_t$

$$m_t(h_t) + \delta\widehat{m}_{t+1}(h_t) \geq \frac{d(\mathbf{FB})}{1-\delta} + \widehat{d}_t(\mathbf{R}|h_t). \qquad (\text{PC}_A^t)$$

$(\text{PC}_A^t)$ and $(\text{IC}_P^t)$ yield

$$\frac{\pi(\mathbf{FB}) - d(\mathbf{FB})}{1-\delta} + \widehat{\pi}_t(\mathbf{R}|h_t) - \widehat{d}_t(\mathbf{R}|h_t) - \widehat{c}_t(\mathbf{C}, \mathbf{E}, \mathbf{S}|h_t) \geq 0,$$

which is the aggregate participation constraint (PC').

Now suppose that $(\mathbf{I}, \mathbf{C}, \mathbf{E}, \mathbf{S}, \mathbf{R})$ satisfies (IC') and (PC'). Consider a strategy profile based on four phases, $\mathcal{N}_1, \mathcal{N}_2, \mathcal{P}_A, \mathcal{P}_P$. In the normal phases $\mathcal{N}_j$ each task covered by task partition $(\mathbf{I}, \mathbf{C}, \mathbf{E}, \mathbf{S}, \mathbf{R})$ (which need not be exhaustive) is regulated informally or by the corresponding formal rule and the other tasks are left to the agent's complete discretion. The system starts in phase $\mathcal{N}_1$ and then from the following period moves to phase $\mathcal{N}_2$ if no deviation occurs. In phase $\mathcal{N}_1$ (period 1) the transfer $m_1$ is such that the participation constraints of principal and

agent are satisfied (such $m_1$ must exist because (PC') holds). In phase $\mathcal{N}_2$ and period $t$ ($t \geq 2$) the transfer is such that all the net surplus goes to the agent:

$$m_t(h_t) = \sigma(\mathbf{FB}) + \sigma(\mathbf{R}, s_t) - c_t(\mathbf{E}, \mathbf{S}, h_t).$$

In normal phases the agent accepts the proposed contract and chooses the efficient action for all the tasks $k \in \mathbf{I}$. As soon as player $i$ deviates the system switches immediately from the current phase to the punishment phase $\mathcal{P}_i$ which is defined as in the proof of Lemma 1. Hence, there are no incentives to deviate in the punishment phases.

The proof that the principal has no incentive to deviate in a normal phase is essentially the same as in the proof of Lemma 1. By construction the agent has not incentive to reject the contract in the first period. (PC') implies that also in the following periods the agent has no incentive to reject, because he is the recipient of the net surplus, which has positive present value. To see that the agent has no incentive to cheat in a normal phase simply note that the LHS of (IC') is the expected present value of future net benefits to the agent (which he forgoes if he cheats), whereas the RHS is the temptation to cheat, i.e. the disutility the agent avoids by cheating on informal tasks. ∎

**Proof of Proposition 3**

Parts (i) and (iii) are obvious. To see that (ii) holds, first note that $d(\mathbf{N}) > \frac{\delta}{1-\delta}\sigma(\mathbf{N})$ implies that a fully informal contract violates (IC). Next suppose without loss of generality that $d^{\min} = d_1$, and consider the task partition corresponding to the MPE modified by replacing the formal rule for task 1 with the informal rule. The expected present value of writing costs conditional on any history is bounded above by $\frac{1}{1-\delta}\sum_{k>1} c(a_k)$. If $d_1 \leq \frac{\delta}{1-\delta}[\sigma(\mathbf{N}) - \sum_{k>1} c(a_k)]$, then the modified partition satisfies (IC), which means that a fully formal partition cannot be a solution to problem (P). The latter inequality is satisfied if $d_1 < \frac{\delta}{1-\delta}\sigma(\mathbf{N})$ (that is $\sigma(\mathbf{N}) - \frac{1-\delta}{\delta}d_1 > 0$) and writing costs are sufficiently small, i.e. $\sum_{k>1} c(a_k) \leq \sigma(\mathbf{N}) - \frac{1-\delta}{\delta}d_1$. ∎

**Proof of Proposition 4**

Suppose that a (complete) task partition $(\mathbf{I}, \mathbf{C}, \mathbf{E}, \mathbf{S})$ satisfies (IC) and there are tasks $j$ and $k$ such that $j \in \mathbf{I}$, $k \in \mathbf{C} \cup \mathbf{E} \cup \mathbf{S}$ and $d_j > d_k$. Since by assumption, $j$ and $k$ have the

same writing costs and transition probabilities, swapping the respective rules does not affect the present value of expected costs and relaxes the constraints (IC). Then $(\mathbf{I}, \mathbf{C}, \mathbf{E}, \mathbf{S})$ cannot be the unique solution to (P). This shows that the (unique) solution to (P) must satisfy the property that for some $\overline{d}$ all tasks $k$ with $d_k < \overline{d}$ are informal and all tasks $k$ with $d_k \geq \overline{d}$ are formal. ∎

### Proof of Remark 1

Under the stated assumptions, Proposition 4 implies that for some threshold $\overline{d}$ all tasks $k$ with $d_k < \overline{d}$ are governed by informal rule, and all tasks $k$ with $d_k > \overline{d}$ are governed by spot rule. Assume without loss of generality that $d_1 \leq d_2 \leq ...d_{N-1} \leq d_N$. If it is optimal to use the same type of rule for all formal tasks, then the solution to (P) can be found by solving the following simpler problem:

$$\max I$$
$$\text{s.t.} \ \sum_{k=1}^{I} d_k \ \leq \ \frac{\delta}{1-\delta}[\pi(\mathbf{N}) - d(\mathbf{N})] - (N-I)\delta \hat{c}_{+1}^0,$$

where $I = |\mathbf{I}|$ is the number of informal tasks and $\delta \hat{c}_{+1}^0$ is the present value future writing costs for any given task $k$ under the optimal formal rule in the most unfavourable situation, i.e. in the Markov state where the PDV of future writing costs is the highest. If the formal rule is $\mathcal{C}$, then $\hat{c}_{+1}^0 = 0$; if it is $\mathcal{E}$ then $\hat{c}_{+1}^0 = [r \cdot c(a) + c(e)] \frac{\delta(1-p)}{1-\delta p}$; if it is $\mathcal{S}$, then $\hat{c}_{+1}^0 = \frac{r(1-q)[1-\delta(2p-1)]}{(1-\delta)[1-\delta(p+q-1)]} c(a)$ (under a contingent rule all writing costs are paid upfront; under the enrichment rule future writing costs are positive only if the initial rigid clause has not yet been amended to a contingent clause; under the spot rule the PDV of future writing costs is higher when event $\neg e_k$ has just occurred, because in this case the probability that future shocks will match the current rigid clause is lower). It can be checked that the PDV of future net surpluses generated by a formal task, $\frac{\delta}{1-\delta}(\pi_k - d_k) - \delta \hat{c}_{+1}^0$, is increasing in $\delta$ and (weakly) decreasing in $p$, $q$ (parametrizing uncertainty) and in writing costs. By inspection of the constraint, it is then clear that any of the parameter changes under consideration (weakly) increases the optimal number of informal tasks, and correspondingly (weakly) reduces the number of formal tasks. ∎

### Proof of Remark 2

Consider a model with two symmetric tasks and symmetric i.i.d. random processes: $p_k = q_k = \frac{1}{2}$, $c(e_k) = c^e$, $c(a_k) = c^a = r \cdot c(a_k)$, $\pi_k = 1$, $d_k = d < 1$, $k = 1, 2$. The cost of a spot rule is $\hat{c}(\mathcal{S}) = c^a + \frac{\delta}{2(1-\delta)} c^a$. We exhibit parameters values $c^e$, $c^a$, $d$, $\delta$ that satisfy the following inequalities:

$$d > \delta, \tag{7.3}$$

$$\frac{\delta}{1-\delta}(2 - 2d - \frac{1}{2}c^a) < d \le \frac{\delta}{1-\delta}(2 - 2d) \tag{7.4}$$

$$c^a + \frac{\delta}{2(1-\delta)}c^a < c^a + c^e < 2\left[c^a + \frac{\delta}{2(1-\delta)}c^a\right], \tag{7.5}$$

$$2 - 2d \le \frac{\delta}{1-\delta}(2 - 2d) - 2\delta(c^e + c^a), \tag{7.6}$$

$$c^e + c^a < 1 - d \tag{7.7}$$

Eq. (7.3) implies that fully informal contracting is not incentive compatible (see Proposition 4); (7.4) says that having one informal task and one task governed by spot rule violates (IC), but having one informal task and one governed by contingent rule satisfies (IC); (7.5) says that contingent contracting is more expensive than spot contracting, so that both tasks are governed by a spot rule in the MPE, but one contingent rule is less costly than two spot rules; (7.6) guarantees that it is possible to keep the principal at his maxmin in a subgame perfect equilibrium; finally (7.7) is the small cost assumption. The last two conditions imply that a solution to (P) is an efficient task partition (see Lemma 2). If all these conditions are satisfied, in the efficient task partition one task is governed by contingent rule and the other is informal, which proves the claim.

It can be verified that the following set of parameter values satisfies all the conditions: $c^e = \frac{1}{9}$, $c^a = \frac{1}{9}$, $d = \frac{3}{4}$, $\delta = \frac{5}{8}$. ∎

### Proof of Remark 3

Consider again a model with two tasks where $p_k = q_k = \frac{1}{2}$, $c(e_k) = c^e$, $c(a_k) = c^a$, $r = 1$, $\pi_k = 1$, $k = 1, 2$, $d_1 \le d_2 < 1$. The cost of a spot rule is $\hat{c}(\mathcal{S}) = c^a + \frac{\delta}{2(1-\delta)}c^a$. We exhibit parameter values $d_1$, $d_2$, $c^e$, $c^a$, $\delta$ that satisfy the following inequalities:

$$d_1 + d_2 \;>\; 2\delta, \tag{7.8}$$

$$\frac{\delta}{1-\delta}(2 - d_1 - d_2) \;>\; d_1 = \frac{\delta}{1-\delta}[2 - d_1 - d_2 - \frac{1}{2}c^a], \tag{7.9}$$

$$2\left[c^a + \frac{\delta}{2(1-\delta)}c^a\right] \;>\; c^a + c^e > c^a + \frac{\delta}{2(1-\delta)}c^a, \tag{7.10}$$

$$2 - d_1 - d_2 \;<\; \frac{\delta}{1-\delta}(2 - d_1 - d_2) - 2\delta(c^e + c^a) \tag{7.11}$$

$$c^a + c^e \;<\; 1 - d_2 \tag{7.12}$$

Condition (7.8) says that a fully informal task partition is not incentive compatible; (7.9) says that governing task 1 by informal rule and task 2 by contingent rule satisfies (IC) as a strict inequality, and that if the contingent rule is replaced by the spot rule (IC) is satisfied as an equality; (7.10 says that two spot rules cost more than one contingent rule, which costs more than one spot rule; (7.11) implies that P can be punished at his maxmin; (**??**) is the small cost assumption (S). By Lemma 2, if these conditions hold the efficient partition is $(\mathbf{I}, \mathbf{C}, \mathbf{E}, \mathbf{S}) = (\{1\}, \emptyset, \emptyset, \{2\})$ (task 1 is informal, task 2 is governed by spot rule). After a sufficiently small decrease of $\pi(\mathbf{N})$ below 2, or $\delta$, the strict inequalities still hold, but the equality in (7.9) does not hold. Therefore the efficient partition changes from $(\mathbf{I}, \mathbf{C}, \mathbf{E}, \mathbf{S}) = (\{1\}, \emptyset, \emptyset, \{2\})$ (which now is not incentive compatible) to $(\mathbf{I}, \mathbf{C}, \mathbf{E}, \mathbf{S}) = (\{1\}, \{2\}, \emptyset, \emptyset)$ (task 2 goverend by contingent rule). This proves the claim in reverse.

The following parameter values satisfy all the conditions: $d_1 = \frac{3}{4}$, $d_2 = \frac{31}{40}$, $c^e = \frac{1}{10}$, $c^a = \frac{1}{20}$, $\delta = \frac{5}{8}$. ∎

### Proof of Proposition 5

Suppose that tasks differ only with respect to $d_k$ and a task partition $(\mathbf{I}, \mathbf{C}, \mathbf{E}, \mathbf{S}, \mathbf{R})$ satisfying (IC') and (PC') violates the property stated in the proposition. Note that $\pi_k - d_k < \pi_j - d_j$ if and only if $d_j < d_k$, because $\pi_k = \pi_j$. If $d_j < d_k$, but $j \notin \mathbf{FB}$ and $k \in \mathbf{FB}$, then it is possible to increase the net surplus and relax constraints (IC') and (PC') by swapping the rules for $j$ and $k$ (the present value of expected costs is not affected because writing costs and transition probabilities are the same for each task). Therefore the original partition cannot be the solution to (P'). Similarly, if $d_j < d_k$, but $j \notin \mathbf{FB} \cup \mathbf{R}$ and and $k \in \mathbf{FB} \cup \mathbf{R}$, it is possible

to increase the net surplus and relax constraints (IC') and (PC') by swapping the rules for $j$ and $k$. Again, this means that the original partition cannot be the solution to (P'). This shows that high surplus (low disutility) tasks are governed by first best rules, intermediate surplus (intermediate disutility) tasks are governed by rigid rules, and low surplus (high disutility) tasks are left to the agent's discretion. Relying on the uniqueness of the solution to (P') one can show as in the proof of Proposition 4 that, among the first best tasks, those with lower disutility are governed informally. ∎

# References

[1] AL NAJJAR, N., L. ANDERLINI and L. FELLI (2002): "Unforeseen Contingencies," Theoretical Economics Discussion Paper TE/02/431, STICERD, London School of Economics.

[2] ANDERLINI, L. and L. FELLI (1994): "Incomplete Written Contracts: Undescribable States of Nature," *Quarterly Journal of Economics,* **109**, 1085-1124.

[3] ANDERLINI, L. and L. FELLI (1999): "Incomplete Contracts and Complexity Costs," *Theory and Decision,* **46** (1), 23-50.

[4] BAKER, G., R. GIBBONS and K. MURPHY (1994): "Subjective Performance Measures in Optimal Incentive Contracts," *Quarterly Journal of Economics,* **109**, 1125-1156.

[5] BATTIGALLI, P. and G. MAGGI (2002): "Rigidity, Discretion, and the Cost of Writing Contracts," *American Economic Review,* **92**, 798-817.

[6] BULL, C. (1987): "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics*, **102**, 147-159.

[7] DYE, R. A. (1985a): "Costly Contract Contingencies," *International Economic Review*, **26**, 233-50.

[8] DYE, R. A. (1985b): "Optimal Length of Labour Contracts," *International Economic Review*, **26**, 251-270.

[9] ECCLES, R. (1981): "The quasifirm in the construction industry," *Journal of Economic Behavior and Organization,* **2**, 335-58.

[10] FUDENBERG, D. and J. TIROLE (1991): *Game Theory.* Cambridge MA: MIT Press.

[11] GRAY, J. (1976): "Wage Indexation: A Macroeconomic Approach," *Journal of Monetary Economics,* **2**, 221-236.

[12] GRAY, J. (1978): "On Indexation and Contract Length," *Journal of Political Economy*, 86, 1-18.

[13] HART, O. and B. HOLMSTROM (1987): "The Theory of Contracts," in T. Bewley (Ed.) *Advances in Economic Theory, Fifth World Congress.* Cambridge UK, C.U.P.

[14] KRASA, S. and S.R. WILLIAMS (2001): "Incompleteness as a Constraint in Contract Design," mimeo, University of Illinois.

[15] LEVIN, J. (2003): "Relational Incentive Contracts," *American Economic Review*, **93**, 835-857.

[16] MACAULAY, S. (1963): "Non-contractual relations in business: a preliminary study," *American Sociological Review*, **28**, 55-67.

[17] MacLEOD, W. B. (2000): "Complexity and Contract," *Revue d'Economie Industrielle*, **92**, 149-178.

[18] MacLEOD, W. B. and J.M. MALCOMSON (1989): "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica,* **57**, 447-480.

[19] MacLEOD, W. B. and S. CHAKRAVORTY (2004): "On the Efficiency of Standard Form Contracts: The Case of Constructions," mimeo.

[20] MASKIN, E. and J. TIROLE (1999): "Unforeseen Contingencies and Incomplete Contracts," *Review of Economic Studies*, **66**, 83-114.

[21] MEIHUIZEN, H.E. and S.N.WIGGINS (2000), "Information Cascades and Contractual Incompleteness in Natural Gas Contracting," mimeo.

[22] PEARCE, D. and E. STACCHETTI (1998): "The Interaction of Implicit and Explicit Contracts in Repeated Agency," *Games and Economic Behavior,* **23**, 75-96.

[23] WILLIAMSON, O. (1985): *The Economic Institutions of Capitalism.* New York: Free Press.