Rigidity, Discretion, and the Costs of Writing Contracts

By Pierpaolo Battigalli and Giovanni Maggi*

In this paper we model contract incompleteness "from the ground up," as arising endogenously from the costs of describing the environment and the parties' behavior. Optimal contracts may exhibit two forms of incompleteness: discretion, meaning that the contract does not specify the parties' behavior with sufficient detail; and rigidity, meaning that the parties' obligations are not sufficiently contingent on the external state. The model sheds light on the determinants of rigidity and discretion in contracts, and yields rich predictions regarding the impact of changes in the exogenous parameters on the degree and form of contract incompleteness. (*JEL* D23, D8, L14)

It is often argued that contracts are incomplete because it is too costly to describe all the relevant contingencies and the exact behavior of the contracting parties. In his discussion of the causes of contract incompleteness, Jean Tirole (1999) classifies them in three categories: (i) unforeseen contingencies; (ii) costs of writing contracts; and (iii) costs of enforcing contracts. This paper focuses on point (ii) of this list, which according to Tirole (1999 p. 772) is a weak point in the existing literature:

> ... many have argued that contingencies are missing because of substantial costs of writing them. While there is no arguing that writing down detailed contracts is very costly, we have no good paradigm in which to apprehend such costs.

In a similar spirit, Richard Posner (1986 p. 92) emphasizes the relevance of the costs of writing contracts:

> ... some contingencies, even though foreseeable in the strong sense that both parties are fully aware that they may occur, are so unlikely to occur that the costs of careful drafting to deal with them might exceed the benefits, when those benefits are discounted by the (low) probability that the contingency will actually occur.

Just what types of costs are incurred in writing contracts is open to debate. For the purposes of this paper, we have in mind costs that are, broadly speaking, proportional to the amount of detail in the contract, such as the cost of figuring out the relevant contingencies and obligations, the cost of thinking how to describe them, the cost of time needed to write the contract, and the cost of lawyers.

We have in mind not only written contracts but also informal (oral) contracts. For example, the relationship between a baby-sitter and the child's parents is typically regulated by an informal contract, in which a set of instructions is communicated orally to the baby-sitter (and enforced by the threat of firing her). This set of instructions is typically very incomplete. We believe an important reason for this is that oral communication suffers from similar costs of complexity as written contracts. To keep the terminology lean, however, in the remainder of the paper we will simply talk about "writing costs."

* Battigalli: Università Bocconi, IEP, Via Gobbi 5, 20136 Milano, Italy (e-mail: pierpaolo.battigalli@unibocconi.it); Maggi: Department of Economics, Princeton University, Princeton, NJ 08544 (e-mail: maggi@princeton.edu). We thank Luca Anderlini, Patrick Bolton, Leonardo Felli, Drew Fudenberg, Katrina Ghironi, Gene Grossman, Faruk Gul, Joe Halpern, Andrea Ichino, Pat Kehoe, Bart Lipman, Salvatore Modica, John Moore, Fabrizio Perri, Ariel Rubinstein, Marciano Siniscalchi, Jean Tirole, Tim Van Zandt, four anonymous referees, and participants in various seminars and conferences for helpful comments and suggestions. Battigalli gratefully acknowledges financial support from the European University Institute and Università Bocconi. Maggi thanks the National Science Foundation for financial support, and the New York University and the European University Institute for their hospitality.

Even though the above quotes by Tirole and Posner emphasize the notion of "missing contingencies," it is clear at a moment's reflection that contract incompleteness can take two distinct forms: (excessive) *discretion,* meaning that the contract does not specify the parties' behavior with sufficient precision; and *rigidity,* meaning that the parties' obligations are not sufficiently contingent on the external state. For example, a construction contract is characterized by discretion if it does not specify the materials with sufficient precision (and this results in the contractor choosing low-quality materials); and is characterized by rigidity if the completion time for the project is fixed, when it would be more efficient to make it contingent on certain exogenous events.

The presence of writing costs can explain both of these forms of incompleteness. Intuitively, if it is costly to describe the external contingencies *and the parties' behavior,* then there is a potential role for both rigidity and discretion. In this paper we explore this intuition more rigorously. We will present a model that sheds light on the implications of writing costs for the optimal degrees of rigidity and discretion in contracts, and is tractable enough to generate potentially testable predictions about the impact of changes in the fundamental parameters.

We now sketch the structure of the model and our main results. A key feature of the model is that it makes explicit the language used to describe the environment and the parties' behavior. In particular, the language consists of (i) *primitive sentences* that describe elementary events and elementary actions (tasks), and (ii) logical connectives (e.g., "not," "and," "or"). This language can be used to describe state-dependent constraints on behavior, or in other words, a correspondence from states to allowable behaviors. Each primitive sentence has a cost (logical connectives have zero cost), and the total cost of writing the contract is a function (e.g., a summation) of the costs of its primitive sentences.

We consider a simple principal-agent framework with symmetric information, where parties are risk neutral, states and behavior are verifiable, and contracts are perfectly enforceable. A contract specifies obligations for the agent and a monetary transfer. After the contract is signed, the principal makes the agreed-upon transfer to the agent; next, the state is realized and observed, and finally the relevant agent takes the relevant actions.

We characterize the structure of the optimal contract and analyze how this changes with the parameters. In general, the optimal contract is characterized by both rigidity and discretion. In particular, tasks of high importance tend to be regulated by contingent clauses, tasks of intermediate importance tend to be regulated by rigid clauses, and tasks of low importance are left to the agent's discretion.

Two key parameters in our model are the writing cost and the potential surplus (gross of the writing costs). The optimal contract depends on these two parameters through their ratio, so we can describe the model's predictions regarding the impact of these parameters by focusing on one of them. When the potential surplus is relatively small, the optimal contract includes only rigid clauses, and leaves substantial discretion to the agent. When the potential surplus is higher, the contract includes contingent clauses as well as rigid clauses, and leaves less discretion to the agent. As the potential surplus increases further, discretion disappears before rigidity does.

Another interesting comparative-statics result concerns the impact of uncertainty. The model predicts that in more uncertain environments contracts should contain more contingent clauses and fewer rigid clauses, and should leave more discretion to the agent. It is interesting to note that, as uncertainty increases, the optimal contract may become simpler, in the sense of a lower total complexity cost.

At this point, the skeptical reader may still ask: how important are writing costs in reality? This is ultimately an empirical question that we will not be able to settle here, but we will offer a few remarks and casual observations. A preliminary consideration is that the cost of including one additional clause in a contract may well be small, but for most contracting situations the number of events and actions that are potentially relevant is arguably astronomical, so that the cost of writing a complete contract would be very large. The following example should strengthen this point. Consider a principal who delegates the writing of a document to an agent (this could be a lawyer delegating the writing of a letter to an assistant). Of course there is an astronomical number of possible documents that can be written. A complete contract would describe *exactly* the document that the principal wants to see, but this would involve nothing short of writing the whole document, thus defeating the whole purpose of the trade. Instead, it may be optimal to give the assistant an incomplete set of instructions, specifying some general characteristics that the document should have, the number of pages, etc.[1] In this type of situation, writing costs are relevant almost by logical necessity; more generally, the example suggests that, even if the "unit" writing cost is very small, the total cost of a complete contract is easily blown up by the dimensionality of the contracting problem.

The example we just gave concerns the costs of describing behavior. As for the costs of describing contingencies, it is not hard to find examples of contracts where relevant contingencies are missing even though they are foreseeable and verifiable, thereby suggesting the presence of writing costs. An example is provided by Hanne E. Meihuizen and Steven N. Wiggins (2000), who examine the evolution of supply contracts in the natural gas industry between 1946 and 1985 in the United States. Around 1975, most of these contracts were amended to include a new clause that provided for renegotiation of the price in case of deregulation of the industry. Our interpretation is that, before it was introduced, this was a classic "missing contingency." 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 (possibly because of the 1973–1974 oil crisis), or more generally because the expected benefit of writing this clause came to exceed its cost.[2]

This is not the first paper that explicitly models the complexity of writing contracts as a source of contractual incompleteness. The pioneering paper in this literature is by Ronald A. Dye (1985), and more recent papers include Luca Anderlini and Leonardo Felli (1994, 1998, 1999) and Stefan Krasa and Steven R. Williams (1999). Before discussing these papers in more detail, we highlight in general terms what we think is our main contribution to this literature. The above-mentioned papers model contracts as functions mapping external states into an outcome (typically a monetary transfer). As a consequence, in these models, contractual incompleteness can only take the form of rigidity. In our framework we consider other contractual obligations besides monetary transfers, and we assume that a detailed description of such contractual obligations is costly; therefore our model is capable of explaining both rigidity and discretion. A related innovation of our model is that it makes explicit the language used to write contracts; this allows a simple and intuitive formalization of the costs of describing the environment and behavior.

Dye (1985) explains the presence of rigidity by assuming that the cost of writing a contract is increasing in the number of its contingencies, that is, the number of cells in the partition of the state space induced by the contractual function. Our model differs from Dye's in several dimensions. First, we view the complexity of a contract in a very different way. For example, two contracts with the same number of mutually exclusive contingencies have the same cost according to Dye, but could have very different

[1] In this example, the agent does not have better information than the principal, so the incompleteness of instructions is caused solely by communication costs. Of course, if the agent had superior information there would be an additional reason for giving incomplete instructions. We are abstracting from this type of consideration here.

[2] Another example of missing contingencies can be found in the area of environmental insurance contracts. Many insurance companies have recently introduced a new contingent clause in their pollution-insurance contracts. This clause excludes injuries caused by (spores released by) certain strains of mold that grow in buildings. In the past, insurance companies had received some claims related to this type of injury, but the frequency of these claims was very low. The frequency of claims for some reason increased substantially in recent times, and as a consequence the new exclusion clause was added to the contracts. We view this anecdote as suggestive of nonnegligible writing costs. If writing additional clauses were costless, probably the exclusion clause on mold would have been introduced from the beginning.

costs in our model.[3] This is because, in our framework, the cost of a contract is not a function of the number of contingencies specified in the contract, but of how hard it is to describe those contingencies in the given language. Second, the two models yield different comparative-statics predictions, as we will discuss in Section I. Finally, as already mentioned, our model is able to explain the presence of discretion in contracts, while Dye's model is not.

Anderlini and Felli (1994) capture the difficulty of describing contingencies in a different way: in a coinsurance model with a continuum of states, they require that contracts correspond to computable functions, i.e., algorithms that for every input (state) produce an outcome in a finite number of steps. They show that the computability constraint per se does not preclude an approximate first best. But if the decision process used to select the contract is also constrained to be algorithmic, the resulting contract is incomplete.[4] Krasa and Williams (1999) consider a similar constraint on the complexity of a contract: they assume that the number of relevant contingencies (elementary dummy variables) is countably infinite, but the contractual outcome can depend only on a finite number of contingencies. They explore the conditions under which the optimal contract can be approximated (in a payoff metric) by contracts satisfying this finiteness constraint. Anderlini and Felli (1999) is closer to our work. They consider a large class of complexity measures for computable functions satisfying a few plausible axioms, and show that for any complexity measure in the given class one can find a contracting problem such that the optimal contract is incomplete. Broadly speaking, our approach differs from theirs in that we impose more structure on the problem and in return we get sharper predictions from the model.[5]

Before plunging into the analysis, we need to comment briefly on the well-known irrelevance result by Maskin and Tirole (1999). They show that the possibility of unforeseen contingencies and the costs of describing contingencies need not imply inefficiencies in contracting, provided a message-based mechanism can be played after the state is observed and before actions are taken. We think our approach is useful in spite of the Maskin-Tirole result. First, in many situations it is not feasible to play games after the state is realized and before actions are taken.[6] Second, even if it is feasible to play a mechanism à la Maskin-Tirole, it is still necessary to describe *behavior,* which can be quite complicated. Third, a Maskin-Tirole mechanism can itself be quite complex, and the costs of describing and implementing the mechanism might not be lower than those of describing the relevant contingencies. Fourth, as shown in Battigalli and Maggi (2000a), if parties interact repeatedly and can contract at any point in time, writing costs can lead to inefficiencies even if mechanisms à la Maskin-Tirole are available.[7]

[3] Consider the following two contracts: contract A specifies behavior b^0 if the exogenous event E occurs and behavior b^1 otherwise; contract B specifies behavior b^0 if the exogenous events E and F occur and behavior b^1 otherwise. These contracts have the same complexity cost according to Dye's assumption, whereas contract B is more costly according to our model.

[4] Anderlini and Felli (1998) show that the approximation result of Anderlini and Felli (1994) fails when the parties' utilities are discontinuous. Nabil Al Najjar et al. (2001) present a model with a countable state space, finitely additive probabilities, and continuous utilities, where the approximation result also fails.

[5] The literature has pointed out a number of other potential causes of contract incompleteness beside the costs of writing contracts. Franklin Allen and Douglas Gale (1992), Kathryn E. Spier (1992) and Mathias Dewatripont and Eric Maskin (1995) argue that the presence of asymmetric information can be a source of contract incompleteness. Arnoud W. A. Boot et al. (1993) argue that an optimal contract may exhibit discretion when some contingencies are not verifiable. The reason is that, if contingencies are not verifiable, a contract that completely specifies behavior would force noncontingent actions, while a contract leaving some discretion may induce the agent to respond efficiently to contingencies. A similar argument is made in B. Douglas Bernheim and Michael D. Whinston (1998). In their model, the parties take actions in sequence. If the first-mover's actions are not verifiable, the optimal contract may leave some discretion in the second-mover's choice of actions. Sujoy Mukerji (1998) argues that, if parties are ambiguity averse rather than expected-utility maximizers, the equilibrium contract may be excessively rigid.

[6] Consider the baby-sitting example: the baby-sitter must react quickly to contingencies, and playing a mechanism with the baby's parents before taking action is out of the question.

[7] Further qualifications to the Maskin-Tirole result have been pointed out by Ilya Segal (1999) and Oliver Hart and John Moore (1999). Segal (1999) considers a hold-up problem in which contingencies cannot be described *ex ante,* parties cannot commit not to renegotiate, and only a finite

The paper is structured as follows. In Section I we present the basic model and derive the main results. In Section II we discuss the implications of more general payoffs, of richer languages, and of unforeseen events. Section III offers concluding remarks.

I. The Basic Model

A. *Language*

Our starting point is a simple formalization of the language used to write a contract.

$\Pi^e = \{e_1, e_2, e_3, \cdots\}$ is a finite collection of primitive sentences, each of which describes an *elementary event.* These are the exogenous aspects of the world that are relevant to the contracting problem, for example, e_1: "the baby cries," e_2: "the baby smells," e_3: "it rains." With a slight abuse of terminology, we will use the notation e_k to indicate both an elementary event and the primitive sentence that describes it.

$\Pi^a = \{a_1, a_2, a_3, \cdots\}$ is a finite collection of primitive sentences describing *elementary actions* (or *tasks*), for example, a_1: "feed the baby," a_2: "change the baby's diapers," a_3: "take the baby for a walk," a_4: "talk to the baby."

Using the primitive sentences, the logical connectives $\neg$ ("not"), $\wedge$ ("and"), $\vee$ ("or"), $\rightarrow$ ("if ... then"), the parentheses and the logical constant $\top$ ("tautology") we can derive well-formed formulae about the exogenous environment and/or about behavior.[8] A formula about the environment describes a *contingency,* for example $(e_1 \vee e_2) \wedge (\neg e_3)$ ("it does not rain and the baby cries or smells"). A formula about behavior describes an *instruction,* for example $(a_1 \vee a_3)$ ("feed the baby or take him for a walk"). We will use interchangeably the expressions "formula about the environment" and "contingency," and likewise for expressions "formula about behavior" and "instruction." The logical constant $\top$ in our setting will be used only in two ways: as a formula about the environment it will mean "whatever happens," and as a formula about behavior it will mean "anything." The set of well-formed formulae about the environment is denoted Λ^e, and the set of well-formed formulae about behavior is denoted Λ^a.

An important assumption is that the language just described 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. In Section II, subsection C we will discuss how results are likely to change if parties can use richer languages to write contracts.

B. *Contracts*

We consider formal contracts between a principal and an agent. A contract stipulates a number of clauses of the form "if contingency η_k occurs then the agent must follow the instruction β_k," and a monetary transfer from the principal to the agent. We will represent a nonmonetary clause as a formula $\eta_k \rightarrow \beta_k$, and we will call "job description" a conjunction of such clauses. We could allow transfers to be contingent on the external environment and/or behavior, but there would be no gains from doing so, due to the assumptions (to be introduced shortly) of verifiable states and behavior, risk neutrality, and conflict of interests.[9]

Definition 1: A *contract* is a job description, $g = \bigwedge_{k=1}^{K} (\eta_k \rightarrow \beta_k)$ (where $\eta_k \in \Lambda^e$, $\beta_k \in \Lambda^a$), and a transfer $t \in \mathbf{R}$.

Examples of contract clauses are: (1) $\neg e_3 \rightarrow a_3$, "if it does not rain, take baby for a walk;" (2) $(e_1 \vee e_2) \rightarrow (a_1 \wedge a_2)$, "if baby cries or smells, feed him and change his diapers;" (3) $\top \rightarrow a_4$, "always talk to the baby."

Note that the different contingencies η_k, $k =$

number of actions can be described *ex post.* He shows that, even if message-contingent mechanisms à la Maskin-Tirole are available, the first-best outcome cannot be achieved. Moreover, as the size of the action space grows, the benefit from any message-contingent mechanism shrinks. Hart and Moore (1999) consider a hold-up problem similar to Segal's, and show that the first-best outcome may be unattainable *even if* states can be costlessly described *ex ante.*

[8] A formula is "well formed" if it is constructed according to the rules of the language, which are quite similar to those used in algebra. See, e.g., Alan G. Hamilton (1988) for details.

[9] We provide a formal proof of this claim in our working paper version (Battigalli and Maggi, 2000b).

TABLE 1—A SIMPLE EXAMPLE SHOWING THE RELATIONSHIP BETWEEN ELEMENTARY EVENTS AND STATES

	Event	
Event	e_1	$\neg e_1$
e_2	baby cries and smells	baby smells and does not cry
$\neg e_2$	baby cries and does not smell	baby does not cry and does not smell

1, ... , K, will in general *not* be mutually exclusive, as a contract with mutually exclusive contingencies may be more complex than an equivalent contract with nonexclusive contingencies. Similarly, a contingency η_k will in general not be a complete description of the environment and an instruction β_k will in general not be a complete specification of behavior.

Since the transfer will be determined so as to make the agent indifferent between accepting and rejecting the contract, it will play no interesting role in the analysis. For this reason, with a slight abuse of our terminology, we will refer to the job description g simply as the "contract."

C. *Costs of Writing Contracts*

We assume that a writing cost c is incurred for each primitive sentence included in the contract. We also assume that writing the logical connectives, the logical constant, and the transfer has no cost. Thus, if n^g is the number of *distinct* primitive sentences occurring in contract g, the cost of writing g is

$$C(g) = cn^g.$$

For example, if g consists of clauses (1)–(3) above, the cost of writing contract g is $7c$.[10]

[10] We have implicitly assumed that it is costless to "recall" a primitive sentence within the contract. The qualitative results would not change if we introduced a cost r of recalling a primitive sentence, or more formally, if we assumed $C(g) = cn^g + r(\bar{n}^g - n^g)$, where n^g is the number of distinct primitive sentences occurring in contract g, $\bar{n}^g$ the total number of primitive sentences contained in contract g, and $0 \leq r \leq c$. We note in passing that our remark on mutually exclusive contingencies in the previous section is valid for any $r > 0$ but not for $r = 0$. For $r = 0$, it turns out that contracts with mutually exclusive contingencies are not more costly to write than other contracts specifying the same obligations.

D. *States and Behavior*

A *state of the environment* (or simply 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. In other words, $s(e_k)$ is a dummy variable that takes value 1 if elementary event e_k occurs and 0 otherwise; and a state is a realization of the vector of dummy variables $(s(e_1), s(e_2), \cdots)$. We let $S = \{0, 1\}^{\Pi^e}$ denote the set of possible states.

A simple example can illustrate the relationship between elementary events and states. Suppose the only relevant elementary events are e_1 ("the baby cries") and e_2 ("the baby smells"). Then there are four relevant states as shown in Table 1.

Function s is extended on Λ^e in the standard inductive way:

$$s(\top) = 1$$
$$s(\neg\eta) = 1 \text{ if and only if } s(\eta) = 0$$
$$s(\eta \vee \varepsilon) = \max(s(\eta), s(\varepsilon))$$
$$s(\eta \wedge \varepsilon) = \min(s(\eta), s(\varepsilon)).$$

Similarly, a *behavior* is a complete description of all elementary actions, represented by a valuation function b: $\Pi^a \rightarrow \{0, 1\}$ (that is, $b \in \{0, 1\}^{\Pi^a}$). Here $b(a_k) = 1$ means that elementary activity a_k is executed, and $b(a_k) = 0$ that a_k is not executed. The function b is extended on Λ^a analogously as function s. We let $B = \{0, 1\}^{\Pi^a}$ denote the set of possible behaviors.

To illustrate the relationship between elementary actions and behaviors, suppose the only relevant elementary actions are a_1 ("feed the baby") and a_2 ("change the baby's diapers"). Then there are four relevant behaviors as shown in Table 2.

TABLE 2—A SIMPLE EXAMPLE SHOWING THE RELATIONSHIP BETWEEN ELEMENTARY ACTIONS AND BEHAVIORS

	Action	
Action	a_1	$\neg a_1$
a_2	feed and change diapers	change diapers and do not feed
$\neg a_2$	feed and do not change diapers	do not feed and do not change diapers

E. *The Behavioral Correspondence*

A contract g, if enforced, imposes state-dependent constraints on the behavior of the agent. These constraints constitute what we call the *behavioral correspondence* induced by the contract. This behavioral correspondence can be derived logically from the contract g in the following way.

For every contingency η_k, define the *truth set* of η_k, denoted by $\|\eta_k\|$, as the set of possible states where contingency η_k is true; that is,

$$\|\eta_k\| = \{s \in S : s(\eta_k) = 1\}.$$

and define analogously the truth set of β_k, denoted by $\|\beta_k\|$, as the set of behaviors b such that the instructions β_k are satisfied. The behavioral correspondence induced by g is

$$B^g(s) = \bigcap_{\{k : s \in \|\eta_k\|\}} \|\beta_k\|.$$

In words, $B^g(s)$ is the set of behaviors allowed by the contract at state s, namely those behaviors that satisfy the instructions specified in all the clauses that apply to state s. Once the contract is signed, the agent has to choose his behavior in set $B^g(s)$. We assume that a contract is enforceable only if it specifies feasible (hence, noncontradictory) obligations for all states.[11]

We say that contract g is *feasible* if $B^g(s) \neq \emptyset$ for all $s \in S$. We denote the set of feasible contracts by F.

[11] This is by no means the only possible assumption. An alternative assumption would be that courts enforce the contract in all states for which the contract specifies feasible obligations, and enforce no obligations in all other states. Under this alternative assumption, it is possible that the optimal contract will stipulate unfeasible obligations (i.e., will contain contradictory prescriptions) in some states, as this may potentially economize on writing costs.

F. *Efficiently Written Contracts*

An important element of our logical construct is the definition of an efficiently written contract. For this, however, we first need appropriate notions of equivalence between contracts.

Definition 2: Two contracts g and h are *behaviorally equivalent* if they specify the same constraints on behavior at each state, that is, $B^g(s) = B^h(s)$ for all $s \in S$. Two contracts g and h are *equivalent* if they are behaviorally equivalent and they have the same cost $(C(g) = C(h))$.

The following is our notion of efficiency in writing a contract:

Definition 3: A contract g is *efficiently written* if $C(g) \leq C(h)$ for every behaviorally equivalent contract h.

G. *Environment and Payoffs*

We are now ready to describe the game between the principal and the agent. The principal proposes a feasible contract (job description and transfer) to the agent, who can either accept it or reject it. If the offer is accepted, the transfer is made; after observing the realized state, the agent chooses her behavior subject to the state-dependent constraints implied by the contract. If the offer is rejected, the agent gets her reservation utility. The principal pays the writing costs regardless of the agent's choice.

Note the assumption that only the principal—who has all the bargaining power—pays the writing costs.[12] Implicit in this timing is also the

[12] This assumption is not entirely innocuous. Anderlini and Felli (1997) show that, if both parties must incur a

assumption that it is not feasible for parties to negotiate or communicate after the state is realized and before the agent takes action.[13]

In this basic model we consider a very simple payoff structure. We will later discuss how the results extend to more general payoffs. We assume that there is a one-to-one correspondence between elementary events and elementary activities. The principal wants elementary activity a_n to be carried out if and only if elementary event e_n occurs. For example, in our baby-sitting situation, if it is sunny the baby-sitter should take the baby for a walk (and if it is not sunny she should keep him at home), if the baby cries she should feed him (and if he does not cry she should not feed him), and so on.

The principal gets incremental benefit $\pi_n > 0$ from "matching" e_n with a_n (or $\neg e_n$ with $\neg a_n$), while he gets zero incremental benefit if there is a "mismatch." π_n may depend on n but is independent of other contingencies and tasks. We adopt the convention that $\pi_n > \pi_{n+1} > 0$ for all $n \geq 1$, so we can interpret task 1 as the most important task. Formally, the principal's payoff is $\pi(s, b) - t - C(g)$, where

$$\pi(s, b) = \sum_{n=1}^{N} \pi_n[s(e_n)b(a_n) + (1 - s(e_n))(1 - b(a_n))].$$

The agent's interests are in full conflict with the principal's, in the sense that her preferred actions are always the opposite of the principal's preferred actions. This assumption greatly simplifies the derivation of the optimal contract, as will become clear in the next subsection. Formally, the agent's utility function is $U(s, b, t) = t - d(s, b)$, where

$$d(s, b) = \delta\pi(s, b) \qquad (0 < \delta < 1).$$

We assume for simplicity that the agent's reservation utility is $\bar{U} = 0$.

The parameter δ captures the strength of the conflict of interests between principal and agent. Note that the potential gross surplus is proportional to $1 - \delta \equiv A$. The parameter A thus captures the potential gains from contracting. For future reference, we let

$$BR^g(s) = \arg\min_{b \in B^g(s)} d(s, b)$$

denote the best response of the agent at state s given contract g (we can assume without essential loss of generality that the best response is unique because the agent is indifferent if and only if the principal is also indifferent). Also, we will often use the phrase "first best" to refer to the behavior that maximizes the surplus gross of writing costs, $\pi - d$.

The state and the agent's behavior are verifiable in court, but preferences and realized payoff levels are not. If preferences were verifiable, the first-best behavior could trivially be implemented by a contract of the form "The agent's behavior b must maximize the sum of the parties' utilities."[14] 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. We also assume that the principal cannot "sell the activity" to the agent (i.e., the agent cannot be made the recipient of the gross payoff π); this would be essentially equivalent to specifying a contingent transfer as in the previous point.

The principal and agent's prior beliefs about the exogenous state are given by a common

transaction cost before the negotiation takes place, it is possible that in equilibrium no contract will be signed even though it would be efficient to do so.

[13] Note that this assumption rules out message-based mechanisms à la Maskin-Tirole (1999). This kind of setting is realistic in a variety of situations, for example when the agent must react quickly to contingencies (think of the baby-sitting case). Battigalli and Maggi (2000a) extend the analysis to a setting where parties can contract at any point in time and interact repeatedly. Among other things, they show that, in a certain parameter region, the qualitative results are similar to those obtained in the present static setting.

[14] Sometimes we do observe general "first-best" clauses of this kind, for example when a contract requires an employee to "act in the company's best interest." This kind of clause makes sense if the company's payoff function can, at least imperfectly and at a cost, be verified in court. A more general model would allow for imperfect and costly verification of payoff functions. In this case, it is conceivable that an optimal contract might include both a general "first-best" clause as well as specific behavioral clauses.

probability measure $\mu \in \Delta(S)$.[15] In particular, we assume that the elementary events e_n, $n = 1, \dots, N$, are independently and indentically distributed (i.i.d.) with marginal probability p. By convention we assume $p > 1/2$ (we do not consider the knife-edge case $p = 1/2$ to avoid ties that would make the analysis more tedious). Formally we are assuming:

$$\mu\left(\left(\bigwedge_{m \in K} e_m\right) \wedge \left(\bigwedge_{n \in N \setminus K} \neg e_n\right)\right) = p^{\#K}(1 - p)^{N - \#K} \quad \forall K \subseteq N$$

where N denotes both the number and the set of tasks, and $\#K$ denotes the cardinality of set K. We can think of p as capturing the degree of uncertainty in the environment: the higher p, the lower the uncertainty [notice that the variance of the random variable $s(e_n)$ is decreasing in p].

Finally, we assume that the parameters satisfy a genericity condition, to avoid ties: $c/pA \neq \pi_n \neq [c/(1 - p)A]$ for all $n \in N$.

H. *The Optimal Contract*

Our main objective is to characterize the *optimal* contract, i.e., the contract that maximizes the expected joint surplus net of writing costs:

$$\max_{g \in F} \left\{ \sum_{s \in S} \mu(s)[\pi(s, BR^g(s)) - d(s, BR^g(s))] - C(g) \right\}.$$

Note that at any subgame-perfect equilibrium of the game the offered contract must be optimal in the sense just specified. If the writing cost c is not very high, g is nonempty, the transfer satisfies the participation constraint as an equality and the agent accepts. If c is very high, g is empty.

We start by defining some useful benchmark notions. We say that g^* is a *first-best* contract if it implements the first-best outcome in every state, or equivalently if it is optimal in the absence of writing costs. Formally, a first-best contract solves

$$\max_{g \in F} \left\{ \sum_{s \in S} \mu(s)[\pi(s, BR^g(s)) - d(s, BR^g(s))] \right\}.$$

Of course, there are many first-best contracts, because there are many ways to write a contract that implements the first-best outcome. A more interesting benchmark for our purposes is the *efficiently written first-best* contract (see the definition of efficiently written contracts given earlier). It is easy to show that, if c is strictly positive and sufficiently close to zero, an optimal contract is an efficiently written first-best contract.

We can distinguish between two basic forms of incompleteness: (1) We say that a contract g is *rigid* if there are two distinct states $s, s' \in S$ such that $B^g(s) = B^g(s')$. (2) We say that a contract g exhibits *discretion at s* if $\#B^g(s) > 1$.

In words, a contract exhibits discretion if the behavioral correspondence B^g is not single-valued, and is rigid if B^g is not one-to-one. Rigidity is a lack of sensitivity of the contractual obligations to the external state. Our notion of discretion includes as a special case a notion of incompleteness that is fairly common in the literature, namely that the contract is "silent" (it specifies no obligations) at a given state.

Rigidity and discretion are two ways of saving on writing costs. Omitting from the contract an elementary sentence e_n saves on the cost of describing contingencies, but makes the contract rigid. Omitting from the contract an elementary sentences a_n saves on the cost of describing behavior, but gives discretion to the agent. A key objective of the analysis will be to examine under what conditions the optimal contract displays one or the other form of incompleteness.

Note that, in our basic model, any rigidity or discretion implies a divergence from the first-best outcome. In a more general model this need not be true, as a contract exhibiting rigid-

[15] With a slight abuse of notation we will sometimes write sentences and formulae as an argument of μ, as in $\mu(\eta) = \mu(\{s : s(\eta) = 1\})$.

ity or discretion might implement the first-best outcome.

We can now start with the analysis of the model. Following is a remark on the form of the efficiently written first-best contract:

Remark 1: Every efficiently written first-best contract is equivalent to

$$\bigwedge_{k=1}^{N} (e_k \rightarrow a_k) \wedge (\neg e_k \rightarrow \neg a_k)$$

and therefore does not exhibit any rigidity or discretion.

PROOF:
Straightforward.

We are now ready to characterize the optimal contract. In the next proposition, we let

$$K_1 = \max\left\{k \in N : \pi_k > \frac{c}{(1-p)A}\right\}$$

and

$$K_1 + K_2 = \max\left\{k \in N : \pi_k > \frac{c}{pA}\right\}.$$

PROPOSITION 1: *Every optimal contract is equivalent to the following:*

$$\left(\bigwedge_{k=1}^{K_1} (e_k \rightarrow a_k) \wedge (\neg e_k \rightarrow \neg a_k)\right)$$

$$\wedge\left(\bigwedge_{k=K_1+1}^{K_1+K_2} (\top \rightarrow a_k)\right)$$

(by convention, if $K_1 = 0$ the first subset of clauses is empty, if $K_2 = 0$ the second subset of clauses is empty).

PROOF:
See Appendix.

This proposition states that, in general, the optimal contract is characterized both by rigidity and by discretion. In particular, the set of N tasks is partitioned in three groups: a group of more important tasks is regulated by contingent clauses; a group of less important tasks is regulated by rigid clauses; and the least important tasks are left to the agent's discretion (i.e., they are not regulated at all). Any of these three subsets of tasks may be empty, depending on parameters.

We now sketch the basic intuition for the result. Using the assumption that expected payoffs are separable with respect to the dimensions $n = 1, \ldots, N$, and that there is full conflict of interests, we show that there is no loss of generality in focusing on contracts with a "separable" structure, in the sense that each dimension n is handled by a clause that depends only on e_n and/or a_n.[16]

There is only a small number of candidate clauses for each n. Refer to Table 3. We have attached labels to clauses to simplify the notation. Label C stands for "contingent," R stands for "rigid," and D stands for "discretion." Any other clause about task n is clearly suboptimal, because it prescribes the wrong action. Next we need to select the clause with the largest incremental net surplus for each task n. It is immediate to verify that the simple contingent clauses $\tilde{C}_n$ and $\tilde{C}_n^{\neg}$ and the rigid clause $R_n^{\neg}$ cannot be optimal. The choice is thus narrowed down to clauses C_n, R_n, and D, which cost respectively $2c$, c, and zero.

TABLE 3—CANDIDATE CLAUSES FOR TASK n

Clause	Label	Incremental net surplus
$e_n \rightarrow a_n$	$\tilde{C}_n$	$pA\pi_n - 2c$
$\neg e_n \rightarrow \neg a_n$	$\tilde{C}_n^{\neg}$	$(1-p)A\pi_n - 2c$
$(e_n \rightarrow a_n) \wedge (\neg e_n \rightarrow \neg a_n)$	C_n	$A\pi_n - 2c$
$\top \rightarrow a_n$	R_n	$pA\pi_n - c$
$\top \rightarrow \neg a_n$	$R_n^{\neg}$	$(1-p)A\pi_n - c$
$\top \rightarrow \top$	D	0

[16] It must be emphasized that this is not true if the agent's interests are partially aligned with those of the principal. In this case, the optimal contract may not be separable in the N tasks; in particular, it might be optimal to include clauses of the form $(a_i \vee a_j)$. This type of clause may be sufficient to induce the agent to take the right action in the right contingency, thus saving the costs of describing contingencies. See our working paper version (Battigalli and Maggi, 2000b) for an example where the parties' interests are partially aligned.

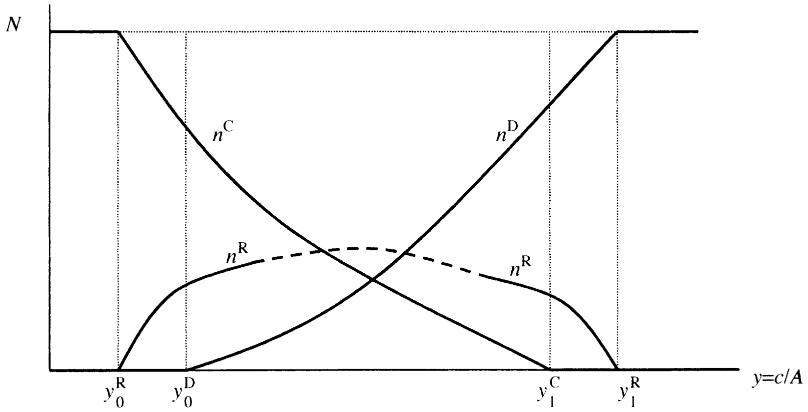

FIGURE 1. IMPACT OF CHANGES IN y ON THE OPTIMAL CONTRACT

Having narrowed down the choice in this way, the result that the most important tasks are regulated by contingent clauses and the least important tasks are left to the agent's discretion is intuitive. The reason a task of intermediate importance is regulated by a rigid clause, on the other hand, is that a rigid clause "gets it right" with probability $p > 1/2$. The task is important enough that the expected benefit of a rigid clause outweighs its cost c, but is not so important that the additional benefit of a contingent clause (which "gets it right" with probability one) over a rigid clause exceeds the cost of including a contingency.

We next examine how changes in the key parameters affect the optimal amounts of rigidity and discretion. Let $y \equiv c/A$. This parameter captures the importance of writing costs relative to the potential gross surplus. The degree of rigidity is captured by the number of rigid clauses (n^R), while the amount of discretion is captured by the number of missing clauses (n^D); the number of contingent clauses is denoted by n^C.

Let y_0^R (respectively, y_0^D) denote the minimum level of y for which there are rigid (missing) clauses in the optimal contract and y_1^C (y_1^R) the maximum level of y for which there are contingent (rigid) clauses in the optimal contract. (In the next proposition, when we use the expression "increasing" or "decreasing" without further specification, we mean it in the weak sense.)

PROPOSITION 2: *(i) n^C is decreasing in y; n^D is increasing in y; n^R is increasing for low values of y and decreasing for high values of y. Furthermore, $y_0^R < y_0^D$ and $y_1^C < y_1^R$. (ii) n^R is increasing in p; n^C and n^D are decreasing in p.*

PROOF:
Straightforward.

Figure 1 illustrates Proposition 2(i), which highlights how changes in y affect the optimal contract. For simplicity, the integer constraint is ignored in the figure. (Part of the n^R curve is dotted to indicate that it need not be single-peaked, and it need not pass above the intersection of the other two curves.)

This comparative-statics result can be described by considering an increase in A, which captures the overall value of the contract. When A is small, no contract is signed. For moderate levels of A, the optimal contract includes only rigid clauses, and leaves substantial discretion to the agent. When A is higher, the contract includes contingent clauses as well as rigid clauses, and leaves less discretion to the agent. As A increases further, discretion disappears before rigidity does, and if A is very high the contract becomes complete.

We believe this prediction is potentially testable. At a broad level, we can think of two ways of going about this. First, one could take a cross-section approach, by looking at the variation of contracts within an industry. For example, our model predicts that when the value of trade (A) is relatively small, the contract should be short and contain only a few rigid clauses, leaving substantial discretion to the parties; when the value of trade is higher, on the other hand, the contract should be longer and contain both rigid and contingent clauses, leaving less discretion to the parties.[17]

Another possibility would be to observe how contracts change over time in situations where the value of contracting (captured by A) increases. For example, this might be the case in a growing industry, to the extent that the size of individual firms (not only the number of firms) tends to grow. One could then check whether the evolution of contracts is consistent with the model's predictions about increases in A.

Proposition 2(ii) focuses on the impact of uncertainty on the optimal contract. In more uncertain environments (i.e., when p is closer to ½), the optimal contract contains more contingent clauses, fewer rigid clauses, and leaves more discretion to the agent. This is intuitive: when uncertainty is higher the efficiency cost of ignoring low-probability events and writing rigid clauses is higher, hence the number of rigid clauses is lower. Moreover, when uncertainty is higher, both contingent clauses and missing clauses increase in number. In Figure 2, the two dots indicate the critical levels of π_n that separate, respectively, missing clauses from rigid clauses, and rigid clauses from contingent clauses. As p decreases, both contingent and missing clauses become more attractive than rigid clauses, hence the result in the picture.

[17] We took a casual look at the area of construction contracts, and what we saw seems consistent with our model. Harold D. Hauf (1968) and James A. Douglas and Jeanne O'Neill (1994), for example, report the most frequent types of construction contracts used in the United States. Projects of smaller value are generally handled by contracts that are fairly simple and short. These short contracts typically contain only a limited set of noncontingent instructions, including a plan of the facility and a specification of the materials to be used, and leave much discretion to the constructor. On the other hand, bigger projects tend to be handled by longer contracts. These longer contracts give much less discretion to the constructor: they contain a fair number of noncontingent instructions, as well as several contingent clauses, describing for example what the contractor is supposed to do if the site conditions change, or describing the contingencies under which the owner can request a change in the construction specifications.

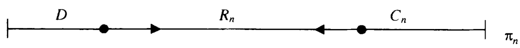

FIGURE 2. EFFECT OF AN INCREASE IN UNCERTAINTY ON THE OPTIMAL CONTRACT

Note that, as uncertainty increases, the optimal contract contains fewer clauses, and it may even be simpler, in the sense of having a lower total complexity cost $C(g)$ (it is easy to show examples where this occurs). This should be contrasted with the prediction of more traditional transaction-cost models, such as Dye (1985). In these models, an increase in uncertainty typically leads to more complex contracts, because contract incompleteness can only take the form of rigidity. What makes a difference in our model is the interplay of rigidity and discretion, which is absent from models à la Dye.

II. Discussion and Extensions

In the previous section we made a number of assumptions on the payoff structure and on the language used to write contracts, and we ignored the possibility of unforeseen events. In this section we discuss how results are likely to change when these assumptions are relaxed.

A. *More General Payoffs*

Here we remove all the symmetry assumptions and the one-to-one correspondence between elementary actions and events. We retain only a minimum of assumptions to ensure that the problem is separable in the N tasks, in the sense that we can optimize the contract task by task. This requires that expected payoffs are separable in the N tasks *and* that there is full conflict of interests between the parties.

We can drop the assumption that the number of elementary actions equals that of elementary events, and we can replace the gross benefit function of the basic model with the following:

$$\pi(b, s) = \sum_{n=1}^{N} \pi_n g_n(b_n; \mathbf{s}_n)$$

where b_n stands for $b(a_n)$, $\mathbf{s}_n \equiv (s_{n_1}, s_{n_2}, \ldots,) \equiv (s(e_{n_1}), s(e_{n_2}), \ldots,)$ is the set of elementary random variables that are relevant for task a_n, and $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_N$ do not overlap (i.e., each elementary event is relevant for at most one task). We can also replace the assumption of i.i.d. elementary events with the weaker condition that the vectors $(\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_N)$ are mutually independent. For example, we could have $\pi(b, s) = \pi_1 g_1(b_1; s_2, s_3) + \pi_2 g_2(b_2; s_1, s_4, s_6) + \pi_2 g_2(b_2; s_5)$, where the vectors (s_2, s_3), (s_1, s_4, s_6) and (s_5) are mutually independent. The task-specific scaling parameter π_n captures the "importance" of task n: an increase in π_n (holding all else equal) blows up the gain from contracting on task n. To avoid tedious ties, we assume that arg $\max_b g_n(b_n, \mathbf{s}_n)$ is unique for all n.

The agent's payoff is still given by $U = t - \delta\pi(b, s)$, and the reservation utility is still zero for both players.

We refer to this setting as the *generalized match-the-state* model. The reason we did not conduct the whole analysis within this more general model is twofold. First, as we will see shortly, the characterization of the optimal contract and the comparative-statics results are not as neat as in the simple match-the-state model. Second, in the simpler version of the model we could capture the degree of uncertainty with a single parameter (p), whereas in this more general setting there is no simple way to examine the comparative-statics effects of changes in the degree of uncertainty.

Let us focus on task n. Performing this task $(b_n = 1)$ is efficient for a certain subset of states, say $E_n^* \subseteq S$. An efficiently written first-best contract will then take the form $\bigwedge_{n=1}^{N} (\eta_n^* \rightarrow a_n) \wedge (\neg\eta_n^* \rightarrow \neg a_n)$, where η_n^* is an efficiently written formula with truth set E_n^*. Note that the efficiently written first-best contract may be partially rigid, in the sense that it may prescribe the same behavior at distinct states.

Next we examine the optimal contract.[18] Without loss of generality, we label elementary actions in such a way that the rigid clause $(\top \rightarrow a_n)$ is preferred to the rigid clause $(\top \rightarrow \neg a_n)$ for each n. We also assume that this preference is strict for each n, to avoid knife-edge cases. In the optimal contract, each task is regulated by a clause of one of three types: (1) A contingent clause of the form $(\eta_n \rightarrow a_n) \wedge (\neg\eta_n \rightarrow \neg a_n)$. If c is sufficiently low, η_n will coincide with η_n^*, and the clause will implement the first best for task n; if c is higher, η_n may be a simpler formula than η_n^*, and the clause may not implement the first best. In any case, a contingent clause costs *at least* $2c$. (2) A fully rigid clause $(\top \rightarrow a_n)$, which costs c. In what follows we refer to this clause simply as *the* rigid clause. (3) A discretionary (empty) clause, which is costless.

Consider first the robustness of Proposition 1. In this more general setting, tasks may not only differ in "importance" (π_n), but also in a number of other ways, since we allow the function g_n to vary by task. For this reason, we cannot hope the result of Proposition 1 to hold exactly as stated. But the result still holds in a ceteris paribus sense:

Remark 2: Consider the generalized match-the-state model. As π_n increases, holding everything else constant, the optimal clause for task n switches from discretionary, to rigid, to contingent.

PROOF:
See Appendix.

Broadly speaking, then, we still have the result that tasks of high importance tend to be regulated by contingent clauses, tasks of intermediate importance tend to be regulated by rigid clauses, and the least important tasks tend to be discretionary.

As for the impact of changes in c and A on the efficient contract, the result of Proposition 2(i) continues to hold:

Remark 3: In the generalized match-the-state model, Proposition 2(i) holds as stated.

PROOF:
See Appendix.

Intuitively, the reason our comparative-statics results are robust is that the rankings

[18] To keep the exposition simple, we continue to refer to the optimal contract, i.e., the contract that maximizes the net surplus. In this extended setting, the optimal contract need not be part of a subgame-perfect equilibrium, because it may yield a negative net surplus (in this case, in equilibrium no contract is signed). At any rate, the comparative-statics results we state for the optimal contract are valid also for any contract signed in a subgame-perfect equilibrium, in the relevant parameter region.

between a contingent clause, a rigid clause, and a discretionary clause in terms of writing costs and in terms of expected benefits have not changed: a contingent clause costs at least $2c$, a rigid clause costs c, and a discretionary clause costs nothing; on the other hand, a contingent clause yields a higher expected benefit than a rigid clause, which in turn yields a higher expected benefit than a discretionary clause.

Proposition 2(ii) cannot easily be extended to the generalized match-the-state model, because the impact of uncertainty can no longer be gauged by a single parameter. However, we believe that the main insight—that reducing uncertainty tends to increase rigidity—should still hold.

An important question is to what extent our results hold when the parties' interests are partially aligned, or when payoffs are not separable across tasks. This more general setting is hard to analyze because the optimal contract may not be separable in the N tasks, and we do not have techniques to solve the general optimization problem. But we conjecture that our qualitative results would hold, to the extent that discretion causes a greater loss of surplus than rigidity. The intuition behind this conjecture is that rigidity saves on the cost of describing contingencies, while discretion saves on the cost of describing contingencies *and* on the cost of describing actions. This insight is quite general, and is the main driving force of our comparative-statics results.

B. *Unforeseen Events*

In this subsection we discuss how the model can be extended to allow for unforeseen events. As a preliminary consideration, there are two types of unforeseen events: unforeseen aspects of the *environment* and unforeseen aspects of *behavior*. Even though the latter notion is rarely emphasized in the literature, we think it is quite relevant in contexts where the complexity of behavior is an important issue. When parties face a nonstandard contracting situation, they have to think hard about all the possible ways that each party can take advantage of the other, so that these actions can be prohibited by the contract. In what follows we use the expression "unforeseen events" to encompass both aspects of the environment and of behavior.

Suppose that, in addition to the elementary exogenous events (binary random variables) that the parties have in mind, $s_1, \ldots, s_N$, there is an additional set of "latent" elementary events, $s'_1, \ldots, s'_{N'}$, that the parties do not have in mind because they are normally turned "off." An example of latent elementary event might be the appearance of the Internet (for someone living a few decades ago). *Ex post,* the parties may become aware of these elementary events if they are turned "on." By convention, let us identify the "off" state of a latent elementary event s'_j with the value $s'_j = 0$. Similarly, we can think of a set of latent elementary actions (binary choice variables), $b'_1, \ldots, b'_{M'}$, that the principal does not have in mind when drafting the contract, because they are normally "off."[19] A latent elementary action might be "the agent gets an autotransfusion."[20] We identify the "off" state of a latent elementary action b'_j with the value $b'_j = 0$. If the "true" benefit function is $\tilde{\pi}(s_1, \ldots, s_N, s'_1, \ldots, s'_{N'}; b_1, \ldots, b_M, b'_1, \ldots, b'_{M'})$, we can think of the principal as having a "perceived" benefit function at the time of writing the contract, given by $\pi(s_1, \ldots, s_N; b_1, \ldots, b_M) = \tilde{\pi}(s_1, \ldots, s_N, 0, \ldots, 0; b_1, \ldots, b_N, 0, \ldots, 0)$. From the point of view of *ex ante* perceived payoffs, the optimal contract will be the same as the one we characterized. From the point of view of the "true" payoffs, however, the presence of unforeseen events implies an additional incompleteness of the contract. If, *ex post,* an unforeseen event does occur, this additional incompleteness will be revealed.

The point we want to stress here concerns the *form* of the incompleteness that is caused by unforeseen events: from the previous arguments it follows directly that unforeseen aspects of the *environment* increase the degree of *rigidity* of

[19] Since the b'_js are actions of the agent, there are two relevant possibilities: one is that neither party is aware of these possible actions, and the other is that only the principal is not aware of them. Since we assumed that the contract is drafted by the principal, what matters most is the principal's (un)awareness.

[20] This example is motivated by a well-known case in the world of cycling. At some point in the history of this sport, some cyclists started to get autotransfusions (transfusions of blood to themselves) to enhance their performance. Soon afterwards, the regulations were changed to prohibit this trick.

the contract, while unforeseen aspects of *behavior* increase the degree of *discretion.* This is a simple but novel point that the previous literature could not make because it did not take into account the complexity of behavior as a source of contract incompleteness.

C. *Richer Languages*

A key aspect of our approach is that we model explicitly the language used to write contracts. We imposed some restrictions on the language at two levels. First, we assumed a propositional language, as opposed to more mathematical languages such as the language of predicates. Second, within the propositional language, we imposed some restrictions on the set of elementary sentences. We will discuss these restrictions in turn. In what follows, we refer to the set of elementary sentences as the "vocabulary" of the language.

In our setting, payoffs depend on a set of binary random variables (s_1, s_2, $\cdots$) and on a set of binary choice variables (b_1, b_2, $\cdots$). Given that these variables are qualitative, the propositional language is the natural one to assume. In a more general setting with quantitative or ordinal variables, it would be necessary to work with a richer language. This would complicate the analysis considerably, because we would have to define a suitable measure of contract complexity in the richer language, and we can think of no simple way to do this.

As far as the vocabulary is concerned, we assumed that each of the relevant economic variables (s_1, s_2, ... ; b_1, b_2, $\cdots$) is associated with an elementary sentence. We think this is a natural vocabulary to assume as a first step of the analysis, however it would probably be more realistic to allow for a richer set of elementary sentences. Suppose that a society is initially endowed with the simple vocabulary we postulated. If a particular formula φ (which might describe for example an external contingency, or a contract clause) turns out to be used very frequently in this society, it may be socially efficient to denote φ with a new elementary sentence: this will save on writing costs (or more generally on communication costs), and these savings may outweigh the social cost of increasing the size of the vocabulary. For example, if e_1: "humid weather" and e_2: "hot weather," and the formula $e_1 \wedge e_2$ occurs frequently, this may be denoted by a new elementary sentence, $\tilde{e}$: "tropical weather." Or, if a_1: "watch television with the baby" and a_2: "sing to the baby," the formula $a_1 \vee a_2$ can be denoted by a new elementary sentence, $\tilde{a}$: "entertain the baby." Similarly, if a particular contract clause, say $\eta_j \rightarrow \beta_k$, is frequently used, this can be denoted by a new label. Even a whole contract may be frequent enough to warrant the addition of a new elementary sentence in the vocabulary.

How robust are the predictions of the model to such enrichments of the vocabulary? Our comparative-statics results are unlikely to change if one relabels formulas about the environment (as in the example of $\tilde{e}$: "tropical weather" mentioned above) or about behavior (as in the example of $\tilde{a}$: "entertain the baby"). The intuition is the usual one: even with this relabeling, it is still true that rigidity saves on the cost of describing contingencies, while discretion saves both on the cost of describing contingencies and on the cost of describing actions.[21] Results may change, however, if the contracting problem features a substantial number of standard *clauses,* i.e., formulas of the form $\eta \rightarrow \beta$ that can be replaced by simple labels, because in this case contingent clauses need not be more costly than rigid clauses. Broadly speaking, if the parties can use standard clauses for some aspects of the contract, then our analysis can be applied only to the non-standard part of the contract.[22]

[21] We can prove this claim rigorously in a particular case. Consider a contracting problem as in Section II, subsection A. Suppose that, in addition to the natural language, there is an additional set of primitive sentences $\tilde{e}_k$, $k = 1, \ldots, n^e$, that can replace more complex formulas φ_k about the environment. Furthermore, suppose that every replaced formula φ_k involves elementary events that are relevant for only one task. Then the contracting problem can still be analyzed task by task and the results of Section II, subsection A hold as stated, because it is still true that a contingent clause costs at least $2c$ while a rigid clause costs only c. In a more general setting, the introduction of standard formulas about the environment or behavior may break the separability of the problem. This is the reason for our rather cautious claim in the text.

[22] There is a subtle but important distinction to make. We are talking about situations where parties can take advantage of *existing* standard contract clauses, not about the *creation* of a standard contract, in the sense for example

To summarize, if the vocabulary available to contracting parties is richer than the simple one we assumed, the predictions of the model are still valid, but subject to an important qualification: our results apply more tightly the less standard is the contracting problem.

One could legitimately ask: why not take the idea of vocabulary enrichment to the extreme consequences, and consider for example a vocabulary that associates an elementary sentence to each possible contract (e.g., contract A, contract B, etc.)? With this vocabulary, the parties could always write a first-best contract at the cost of c, and we would have no contract incompleteness (if c is not too large). However, we believe this type of "complete" vocabulary is highly unrealistic. The common-knowledge vocabulary of a society must serve a large population of heterogeneous contracting parties. Thus, a complete vocabulary would have to include an elementary sentence for each *conceivable* first-best contract. The number of conceivable first-best contracts in reality is astronomical. If there is a social cost of having a richer vocabulary (because a richer vocabulary is more costly to learn, to teach, to remember), then a complete vocabulary will be excessively costly.

We conclude this section with a remark on "private" vocabularies. Consider two contracting parties that can only use the common-knowledge language understood by courts to write an enforceable contract. Can they save on writing costs by creating new elementary sentences? Given our assumption that the cost of writing a contract is proportional to the number of distinct elementary sentences that appear in the contract, the answer is no. The reason is that, in order for the courts to understand the contract, any new elementary sentence needs to be defined *within the contract* in terms of the common-knowledge language, and doing so is at least as costly as writing the contract in the common-knowledge language. If we had a positive cost r of "recalling" elementary sentences within the contract (see footnote 10), then the creation of new elementary sentences could save on writing costs, but the comparative-statics results would be very similar to the ones we presented.

of a company drafting a contract to be offered to multiple customers. Our analysis is broadly applicable to the latter type of situation, as long as the contract writer faces a fresh contracting problem.

III. Concluding Remarks

We developed a multitask, principal-agent model of contract incompleteness where rigidity and discretion arise endogenously from the costs of describing the external environment and the agent's behavior. In this concluding section we briefly discuss another potential application of our way of modeling complexity costs.

Although we chose to focus on a setting characterized by conflict of interests between principal and agent, our approach is potentially useful also for a different type of setting, where the key problem is not one of incentives, but rather one of efficient communication of information. This could be the case in situations where a scientific authority issues directives for practitioners (e.g., the U.S. Center for Disease Control issuing protocols for doctors and nurses on how to diagnose or treat a certain disease), or when the head of a large organization issues protocols for lower-level employees (e.g., the U.S. Postal Service issuing instructions for local postal offices on how to process and handle mail under various contingencies), or in employment relationships where the main reason to instruct the agent is that the principal has better information (this could apply to the baby-sitting case). In situations of this kind, the presence of complexity costs may lead to rigidity and/or discretion in the set of instructions communicated by the principal.

To exemplify how this type of setting can be captured with our framework, consider a simple variant of our model of Section I: suppose that the interests of the principal and the agent are aligned, but the principal is better informed than the agent on the relevant parameters of the payoff functions. Then, if the principal leaves discretion to the agent, there will be a positive probability (from the point of view of the principal) that the agent will take "wrong" actions. If this probability is relatively high, then discretion implies a larger expected loss of surplus than rigidity, hence the qualitative results are likely to be the same as in our basic model. We are able to prove this rigorously in the extreme case where the agent chooses at random within the set of behaviors that do not violate the principal's instructions. Intuitively, discretion (for a given task) in this case implies that the agent will take the wrong action with 50-percent probability; therefore leaving discretion implies a larger expected loss of surplus than giving a rigid instruction. Extending the analysis to a more general setting with asymmetric information is an ambitious task, and will have to await future research.

Appendix

PROOF OF PROPOSITION 1:

Here it is convenient to represent a whole contract (set of clauses) as a logical formula involving both primitive sentences about the environment and primitive sentences about behavior. Consider an arbitrary contract g^0. We will construct a contract that has the features described in the proposition, yields weakly higher expected gross surplus than g^0, and has a weakly lower writing cost than g^0. We construct this contract in three steps:

1. From g^0 we construct a contract g' which induces a constraint set that is a Cartesian product for each s: $B^{g'}(s) = \Pi_{n=1}^{N} B_n^{g'}(s)$, where $B_n^{g'}(s)$ is the nth projection of $B^{g'}(s)$. Define the following index sets: $E(g^0) = \{n \in N$: e_n occurs in $g^0\}$, $A(g^0) = \{n \in N$: a_n occurs in $g^0\}$, $E_1(s, g^0) = \{k \in E(g^0)$: $s(e_k) = 1\}$, $E_0(s, g^0) = \{\ell \in E(g^0)$: $s(e_\ell) = 0\}$, $A_1(b, g^0) = \{m \in A(g^0)$: $b(a_m) = 1\}$, $A_0(b, g^0) = \{n \in A(g^0)$: $b(a_n) = 0\}$. The following is a logically equivalent formulation of g^0:

$$\bigwedge_{s \in S} \left(\left(\bigwedge_{k \in E_1(s,g^0)} e_k \right) \wedge \left(\bigwedge_{\ell \in E_0(s,g^0)} \neg e_\ell \right) \rightarrow \bigvee_{b \in B^{g^0}(s)} \left(\left(\bigwedge_{m \in A_1(b,g^0)} a_m \right) \wedge \left(\bigwedge_{n \in A_0(b,g^0)} \neg a_n \right) \right) \right)$$

(by convention, conjunctions ranging over empty sets should be replaced by $\top$). Now consider the following contract:

$$g': \bigwedge_{s \in S} \left(\left(\bigwedge_{k \in E_1(s,g^0)} e_k \right) \wedge \left(\bigwedge_{\ell \in E_0(s,g^0)} \neg e_\ell \right) \rightarrow \left(\bigwedge_{m \in A_1^*(s,g^0)} a_m \right) \wedge \left(\bigwedge_{n \in A_0^*(s,g^0)} \neg a_n \right) \right)$$

where $A_0^*(s, g^0) = \{n : e_n$ and a_n occur in g^0 and $s(e_n) = 0\}$, and $A_1^*(s, g^0) = A(g^0) \backslash A_0^*(s, g^0)$.

We argue that contract g' yields a (weakly) higher expected surplus than g^0. First note that $C(g^0) = C(g')$ because g^0 and g' contain the same set of elementary sentences. Next observe that by additive separability of payoffs we need only compare the expected incremental gross surplus for each aspect n of the contractual problem. Since the agent minimizes the gross surplus, under both contracts and for each a_n not contemplated in g^0, he chooses to "mismatch," i.e., he chooses $b(a_n) = 1 - s(e_n)$, which yields zero incremental gross surplus. Therefore we only have to compare the agent's behavior under g^0 and g' for elementary actions a_n contemplated in g^0, that is, actions with index $n \in A(g^0)$. If e_n and a_n occur in g^0, then g' forces the agent to take the right action (a_n or $\neg a_n$) in all states, so the incremental gross surplus for aspect n is maximum. If a_n occurs in g^0 but e_n does not, then g' forces the agent to take action a_n in all states [note that in this case $n \in A_1^*(s, g^0)$]. This yields expected incremental gross profit $p\pi_n$. By conflict of interests and independence, this is an upper bound to what can be achieved without including e_n in the contract. Therefore we have $V(g') \geq V(g^0)$, where $V(g)$ denotes the net surplus induced by contract g.

2. From g' we will construct a contract g^* that is separable in the N dimensions. But first we introduce some convenient notation. For each n, s, and b, let $s_n = s(e_n)$, $b_n = b(a_n)$. Let $s_{-n} = (s_1, \ldots, s_{n-1}, s_{n+1}, \ldots, s_N)$ and $(s'_n, s_{-n}) = (s_1, \ldots, s_{n-1}, s'_n, s_{n+1}, \ldots, s_N)$. The marginal probability of s_n is denoted $\mu_n(s_n)$. Similarly, the nth coordinate of the best-response function $BR^{g'}(s)$ this function is denoted by $BR_n^{g'}(s)$, that is, $BR_n^{g'}(s) = 1$ if under contract g' the agent

chooses a_n at state s and $BR_n^{g'}(s) = 0$ if the agent chooses $\neg a_n$ at state s. Also, let

$$\sigma_n(s_n, b_n) = A\pi_n[s_n b_n + (1 - s_n)(1 - b_n)]$$

denote the nth term of the gross surplus. For each $n = 1, \ldots, N$, pick

$$s_{-n}^* \in \arg\max_{s_{-n}} \left[\sum_{s_n \in \{0,1\}} \mu_n(s_n) \times \sigma_n(s_n, BR_n^{g'}(s_n, s_{-n})) \right].$$

We construct g^* in the following way: $g^* = \bigwedge_{n=1}^{N} \gamma_n^*$, where

$$\gamma_n^* = \begin{cases} C_n & \text{if } B_n^{g'}(0, s_{-n}^*) \neq B_n^{g'}(1, s_{-n}^*) \\ D & \text{if } B_n^{g'}(0, s_{-n}^*) = B_n^{g'}(1, s_{-n}^*) = \{0, 1\} \\ R_n & \text{otherwise.} \end{cases}$$

[Recall that $B_n^{g'}(s)$ denotes the nth projection of the constraint set $B^{g'}(s)$.] First note that every elementary sentence contained in g^* is also contained in g' [for example, e_n occurs in g^* only if $\gamma_n^* = C_n$, which implies that $B_n^{g'}(s)$ depends on s_n; thus e_n must also occur in g']. Therefore $C(g^*) \leq C(g')$. Next we argue that g^* yields a weakly higher expected gross surplus than g'. Note that, by definition of g^*, $BR_n^{g^*}(s_n, s_{-n})$ is independent of s_{-n}; thus it makes sense to write $BR_n^{g^*}(s_n)$. By definition of g^*, additive separability, independence, and conflict of interests, we have

$$V(g') + C(g')$$

$$= \sum_{s} \prod_{j=1}^{N} \mu_j(s_j) \left[\sum_{n=1}^{N} \sigma_n(s_n, BR_n^{g'}(s)) \right]$$

$$= \sum_{n=1}^{N} \left\{ \sum_{s_{-n}} \prod_{j \neq n} \mu_j(s_j) \times \sum_{s_n \in \{0,1\}} \mu_n(s_n) \sigma_n(s_n, BR_n^{g'}(s_n, s_{-n})) \right\}$$

$$\leq \sum_{n=1}^{N} \max_{s_{-n}} \left[\sum_{s_n \in \{0,1\}} \mu_n(s_n) \sigma_n \times (s_n, BR_n^{g'}(s_n, s_{-n})) \right]$$

$$\leq \sum_{n=1}^{N} \sum_{s_n \in \{0,1\}} \mu_n(s_n) \sigma_n(s_n, BR_n^{g^*}(s_n))$$

$$= V(g^*) + C(g^*).$$

Since $C(g^*) \leq C(g')$, we obtain $V(g^*) \geq V(g')$. Therefore $V(g^*) \geq V(g^0)$.

3. We have thus far shown that there is no loss of generality in restricting attention to separable contracts where each clause n is one of the three candidates: C_n, R_n, or D. The last step is to determine which of these is optimal to include in the contract for each n. This depends on the parameters p, δ, c, and π_n. Since $p > 1/2$, the threshold values for π_n are ordered as follows

$$\frac{c}{pA} < \frac{c}{(1-p)A}$$

where D is optimal for $\pi_n < (c/pA)$, R_n is optimal for $(c/pA) < \pi_n < [c/(1-p)A]$, and C_n is optimal for $\pi_n > [c/(1-p)A]$. Taking into account that π_n is decreasing in n we obtain that the contract stated in the proposition is optimal. Our genericity assumption implies that any optimal contract must be equivalent to this one.

PROOF OF REMARK 2:

Using techniques similar to parts 1 and 2 of the proof of Proposition 1, one can show that it is possible to maximize task by task. Then the claim is implied by the following observations:

1. The cost of a contingent clause is at least $2c$, the cost of a rigid clause is c, and the cost of a discretionary clause is zero.

2. The benefit of a contingent clause is at most equal to

$$\pi_n\left(\sum_{s\in E_n^*} g_n(1, s)\mu(s) + \sum_{s\in \bar{E}_n^*} g_n(0, s)\mu(s)\right) \equiv \pi_n G_n^{FB}$$

where E_n^* is the set of states where it is efficient to execute a_n and $\bar{E}_n^* = S\backslash E_n^*$ is its complement.

3. The benefit of a discretionary clause is

$$\pi_n\left(\sum_{s\in E_n^*} g_n(0, s)\mu(s) + \sum_{s\in \bar{E}_n^*} g_n(1, s)\mu(s)\right) \equiv \pi_n G_n^{D}.$$

4. The benefit of the best rigid clause is

$$\pi_n \max\left\{\sum_{s\in E_n^*} g_n(1, s)\mu(s) + \sum_{s\in \bar{E}_n^*} g_n(1, s)\mu(s), \sum_{s\in E_n^*} g_n(0, s)\mu(s) + \sum_{s\in \bar{E}_n^*} g_n(0, s)\mu(s)\right\} \equiv \pi_n G_n^{R}.$$

5. $G_n^D < G_n^R < G_n^{FB}$ and $G_n^R > ½(G_n^{FB} + G_n^D)$ (the latter inequality is strict because we assumed that the two rigid clauses are strictly ordered in terms of expected benefit).
6. The critical value of π_n for which a rigid clause is equivalent to an empty clause is $\pi_n^{R/D} = [c/(G_n^R - G_n^D)]$.
7. For a contingent clause to be preferred to a rigid clause it must be $\pi_n > \pi_n^{R/C} = [c/(G_n^{FB} - G_n^R)]$.
8. Finally note that $\pi_n^{R/C} > \pi_n^{R/D}$, which implies the claim.

PROOF OF REMARK 3:

Let us look at a single task a_n. Using observations 1–5 in the previous proof, it is easy to conclude that, as y increases, the optimal clause for task n switches from contingent, to rigid, to discretionary. Aggregating over the N tasks, the claim follows immediately.

REFERENCES

Allen, Franklin and Gale, Douglas. "Measurement Distortion and Missing Contingencies in Optimal Contracts." *Economic Theory*, January 1992, *2*(1), pp. 1–26.

Al Najjar, Nabil; Anderlini, Luca and Felli, Leonardo. "Unforeseen Contingencies." Mimeo, London School of Economics, 2001.

Anderlini, Luca and Felli, Leonardo. "Incomplete Written Contracts: Undescribable States of Nature." *Quarterly Journal of Economics*, November 1994, *109*(4), pp. 1085–124.

______. "Costly Coasian Contracts." Mimeo, London School of Economics, 1997.

______. "Describability and Agency Problems." *European Economic Review*, January 1998, *42*(1), pp. 35–59.

______. "Incomplete Contracts and Complexity Costs." *Theory and Decision*, February 1999, *46*(1), pp. 23–50.

Battigalli, Pierpaolo and Maggi, Giovanni. "Costly Contracting in a Long-Term Relationship." Mimeo, Princeton University, 2000a.

______. "Imperfect Contracting." Mimeo, Princeton University, 2000b.

Bernheim, B. Douglas and Whinston, Michael D. "Incomplete Contracts and Strategic Ambiguity." *American Economic Review*, September 1998, *88*(4), pp. 902–32.

Boot, Arnoud W. A.; Greenbaum, Stuart I. and Thakor, Anjan V. "Reputation and Discretion in Financial Contracting." *American Economic Review*, December 1993, *83*(5), pp. 1165–83.

Dewatripont, Mathias and Maskin, Eric. "Contractual Contingencies and Renegotiation." *RAND Journal of Economics*, Winter 1995, *26*(4), pp. 704–19.

Douglas, James A. and O'Neill, Jeanne. *Construction and development forms: Planning, financing, building*. Boston, MA: Warren, Gorham and Lamont, 1994.

Dye, Ronald A. "Costly Contract Contingencies." *International Economic Review*, February 1985, *26*(1), pp. 233–50.

Hamilton, Alan G. *Logic for mathematicians.* Cambridge: Cambridge University Press, 1988.

Hart, Oliver and Moore, John. "Foundations of Incomplete Contracts." *Review of Economic Studies*, January 1999, *66*(1), pp. 115–38.

Hauf, Harold D. *Building contracts for design and construction.* John Wiley & Sons, Inc., 1968.

Krasa, Stefan and Williams, Steven R. "Incompleteness as a Constraint in Contract Design." Mimeo, University of Illinois, 1999.

Maskin, Eric and Tirole, Jean. "Unforeseen Contingencies and Incomplete Contracts." *Review of Economic Studies*, January 1999, *66*(1), pp. 83–114.

Meihuizen, Hanne E. and Wiggins, Steven N. "Information Cascades and Contractual Incompleteness in Natural Gas Contracting." Mimeo, Texas A&M University, 2000.

Mukerji, Sujoy. "Ambiguity Aversion and Incompleteness of Contractual Form." *American Economic Review*, December 1998, *88*(5), pp. 1207–31.

Posner, Richard. *Economic analysis of law*, 3rd Ed. Boston, MA: Little Brown, 1986.

Segal, Ilya. "Complexity and Renegotiation: A Foundation for Incomplete Contracts." *Review of Economic Studies*, January 1999, *66*(1), pp. 57–82.

Spier, Kathryn E. "Incomplete Contracts and Signalling." *RAND Journal of Economics*, Autumn 1992, *23*(3), pp. 432–33.

Tirole, Jean. "Incomplete Contracts: Where Do We Stand?" *Econometrica*, July 1999, *67*(4), pp. 741–81.