ELSEVIER European Economic Review 42 (1998) 35–59

EUROPEAN ECONOMIC REVIEW

Describability and agency problems[1]

Luca Anderlini [a], Leonardo Felli [b,*]

[a] *St. John's College, Cambridge, UK*

[b] *London School of Economics and Political Science, Department of Economics, Houghton Street, London WC2A 2AE, UK*

Received 1 November 1995; accepted 1 October 1996

Abstract

This paper suggests a reason, other than asymmetric information, why agency contracts are not explicitly contingent on the agent's performance or actions. Two ingredients are essential to this reason. The first is the *written form* that contracts are required to take to be enforceable. The second is a form of *discontinuity* in the parties' preferences and in the technology that transforms actions into a (probabilistic) outcome. We show that under these conditions the chosen contract may not be explicitly contingent on the agent's actions although, in principle, such actions are contractible and observable to all parties to the contract, court included.

JEL classification: C63; L14; J41

Keywords: Principal-agent problems; Written contracts; Hard to describe actions and outcomes

1. Introduction

1.1. Overview

The analysis of remuneration schemes which depend on the outcome, possibly noisy of the agent's action or performance, rather than on the action itself is

* Corresponding author. Tel.: + 44-171-955-7525; fax: + 44-171-831-1840; e-mail: l.felli@lse.ac.uk.

[1] This paper is a revised version of an earlier paper entitled "Endogenous Agency Problems" (Anderlini and Felli, 1994a).

PII S0014-2921(96)00062-1

central to contract theory. One reason, often quoted (Hart and Holmström, 1987), for this contractual form is the asymmetry of information that exists either between the principal and the agent or between the contracting parties, on the one hand, and the enforcing agency (the court) on the other. Indeed, the agent's action is either assumed observable only to the agent or is assumed observable to both the principal and the agent but *not* to the third party (e.g., the court) whose rôle is to enforce the contract. In this paper we argue that information asymmetries are not the only possible cause of the form of principal-agent contracts.

Indeed, in a framework in which the agent's action is contractible and observable to all the parties involved in the contract, an alternative cause of the standard principal-agent contractual form can be found in the *formal nature* of contracts coupled with *discontinuities* in the parties' preferences. We take the formal nature of contracts to be the written form that contracts are required to take by legal prescription, common practice or simply the contracting parties' will, to be enforceable in a court of law.[2] The fact that contracts must be written, clearly places some restrictions on the prescriptions which they can make. For instance, no written contract can be infinitely long.

Consider a framework in which both the agent's action as well as the outcome of such an action are assumed to be contractible and observable to all the parties involved in the contract, court included. Consider, then, a written contract consisting of a finite number of 'clauses'. Once the agent has decided on which action to take and the corresponding outcome has been realized, the parties (or the enforcing agency) will examine the available 'evidence' about the agent's action and about the realized outcome, and identify which clause(s) apply to the case at hand and hence what remuneration for the agent is prescribed by the contract (the principal's remuneration is taken to be 'residual' as it is common in this literature).

We assume that this requires some form of commitment by the parties to the set of clauses they originally choose. In other words we focus on those cases in which an ex-post conflict of interests makes it impossible to leave the

[2] For example, in the American common law a number of contracts require the *formality of writing* to be enforceable. These contracts are listed in the Statute of Frauds and the Sales Act, modified by the Uniform Commercial Code, and include: "… any special promise, to answer damages out of his own estate; … any special promise to answer for the debt, default, or miscarriage of another person; … any agreement made upon consideration of marriage; … any contract or sale of lands, tenements or hereditaments, or any interest in or concerning them; … any agreement that is not to be performed within the space of one year from the making thereof; … [any] contract for the sale of any goods, wares and merchandize, for the price of [$ 500 or more] … except the buyer shall accept part of the goods so sold, and actually receive the same, or give something in earnest to bind the bargain, or in part of payment …". The agreement for these transactions "… shall be in writing, and signed by the party to be charged therewith, or some other person thereunto by him lawfully authorized." (Calamari and Perillo, 1987, p. 775).

enforcement of the contract to 'implicit' mechanisms. Examples of economic situations which fit this description are familiar and abundant; most agency models—including models of irreversible specific investment, on which the literature on incomplete contracts has focused (Grossman and Hart, 1986; Hart and Moore, 1988, 1990)—and most models of risk-sharing come to mind.

A simple and common form of commitment is a written document signed by all parties to the contract. We restrict attention to the case in which the formality of writing is the device used by the parties to commit themselves to the original agreement. We take the view that the need for a finite written contract can be modeled as a *restriction* on the *mapping* between action–outcome pairs and remunerations which represents the contract.

In Anderlini and Felli (1994b) we develop a full theory of how these restrictions are derived from more primitive assumptions. In particular, there we explicitly model the view that a finite written contract must define a mapping between contractible variables and remunerations which is *algorithmic* in nature and refer to the formal theory of algorithmic (or effectively computable, or Turing computable) functions to characterize the resulting class of admissible mappings describing a written contract.[3]

A (Turing) computable function is a function whose values can be effectively computed using a *finite* abstract computing device (a Turing machine) in a *finite* number of steps. A written contract is clearly a finite object. It embodies a well defined procedure which yields a prescription on, for instance, a payment from one contracting party to the other, on the basis of evidence about variables upon which the contract is based. The view that written contracts can be modeled as (Turing) computable functions is therefore compelling in our opinion.

In this paper we take the restrictions on the mappings representing the written contracts as primitive and refer to Anderlini and Felli (1994b) for a full technical justification of these restrictions. This saves a considerable amount of formal details and allows us to focus on the other factor which is critical for our alternative explanation of the form of agency contracts: discontinuities in the parties' preferences and in the available technology. On an intuitive basis, it is not difficult to see how the restrictions we describe in Section 3 flow from the requirement that the mappings representing written contracts must be algorithmic in nature. We assume that a written contract must yield a mapping between action–outcome pairs and remunerations which is a step function with 'cut-off' values given by real numbers which can be described exactly using a *finite* number of digits. We expand further on the reasoning behind this assumption in Section 3.1 describable below.

[3] The notion of computability embodied in the class of abstract computing devices known as Turing machines is widely agreed in the mathematical literature to embody the widest possible notion of 'effective computability'. This claim is known as 'Church's Thesis' (Cutland, 1980, Ch. 3).

The first question we ask in this framework is whether the restrictions we impose on the mapping representing the contract prevent the parties from choosing a contract which at least approximates, arbitrarily closely, the 'first best' contract (the contract the parties would have chosen if no restrictions were imposed on which contracts are feasible). We conclude that under certain *continuity* assumptions about the parties' preferences and the technology which transforms the agent's action into a probabilistic outcome, the contracting parties can at least approximate—in terms of their own expected utility—the first best contract. This is the *approximation* result which we present in Section 4 below.

The continuity assumptions on which the approximation result hinges are not, in our view, always appealing in a principal agent setting, however. Hence, we proceed to characterize the optimal contracts in three examples of discontinuous preferences or technology. In the first two examples we assume that the technology which transforms the agent's action into an outcome is *deterministic*, while in the third case we deal with a stochastic outcome associated with each of the possible actions of the agent.

The first example concerns a situation in which the principal's preferences are discontinuous (Section 5.1). We find that if the discontinuity occurs at a point which is 'difficult' to include in a written contract then the optimal written contract the parties choose will not be contingent on either actions or outcomes. Hence, the written nature of contracts will entail a substantial welfare loss to the parties.

We proceed (Section 5.2 below) to consider a situation in which the discontinuity occurs not in the parties' preferences but in the technology which transforms the agent's action into an outcome. In this case we find that the first best expected utilities can be approximated arbitrarily closely by the expected utilities given by a written contract. Hence, no substantial welfare loss occurs. However, the written nature of contracts endogenously determines the 'shape' of the contract selected by the parties. When contracts are restricted to take a written form the parties will select an agreement which is not directly dependent on the agent's action. The agent's remuneration will be exclusively dependent on the outcome: parties will select a purely 'profit-based', as opposed to a 'performance-based', contract.

The third and final example we consider (Section 6 below) focuses on a situation in which the discontinuity occurs in the agent's preferences *and* in the technology that transforms actions into outcomes. We take this technology to be stochastic, in the sense that it defines a probability distribution over outcomes as a function of the agent's action, and to be discontinuous in this action. We find that the optimal written contract will not be able to approximate the parties' first best level of expected utility but only a second best level will be achievable. Indeed, the optimal written contract may take exactly the form of the standard principal-agent contract derived in the literature on moral hazard (Holmström, 1979).

The plan of the paper is as follows. Section 2 sets up the general contracting problem between the principal and the agent. We then give a formal account of the restrictions which written contracts have to satisfy (Section 3). Section 4 concentrates on the approximation result mentioned above and discusses the rôle and appeal of the continuity assumptions needed in attaining this result. The main results of the paper are presented in Sections 5 and 6 in which we present the examples of agency problems outlined above. Finally, Section 7 offers some concluding remarks.

2. The general problem

We consider a very simple *principal-agent* problem. A principal hires an agent to undertake an action a which the principal does not have either the time or the ability to undertake himself. The agent's action a is assumed to take values in a closed and compact set, which we normalize to be the interval $[0,1]$, and yields an outcome y through a stochastic technology described by the conditional distribution $F(y|a)$ with a support (which may be degenerate), normalized to lie in the unit interval $[0,1]$. The principal's preferences are described by the separable utility function $V(y,I) = V_y(y) + V_I(I)$ strictly decreasing and continuous in the agent's remuneration I, while the agent's preferences are described by the separable utility function $U(I,a) = U_I(I) + U_a(a)$ strictly increasing and continuous in his remuneration I.[4] Finally, the agent has an outside employment opportunity which yields an expected utility normalized to be zero.

To make the contracting problem interesting, we assume that the principal and the agent have a *local conflict of interests* in the sense that the distribution of outcomes, hence the action, most preferred by the principal does not coincide with the one most preferred by the agent. Denote a^* the agent's action that yields the best possible distribution of outcomes for the principal, which we assume exists.

Assumption 1. *There exist two values* $0 \leq \underline{a} < a^*$ *and* $a^* < \bar{a} \leq 1$ *such that* $U(I,a)$ *is strictly decreasing in* a *for every* $a \in (\underline{a},\bar{a})$.[5]

The existence of this conflict of interests between the principal and the agent implies that it would not be enough for the principal to pay the agent a constant

[4] We assume that the preferences of both the principal and the agent are separable in their two arguments for simplicity only. This is qualitatively inessential to any of the results below but it simplifies the notation and the analysis a great deal.

[5] The value of a^* does not depend on I, because of the separability of the principal's preferences. Similarly, the interval $(\underline{a},\bar{a})$ does not depend on the value of I since the agent's preferences are separable. Further, the assumption that $U(I,a)$ be strictly decreasing in a is only essential as far as strict monotonicity is concerned. We could have assumed that $U(I,a)$ be strictly increasing in a without affecting the substance of the analysis.

amount and expect him to undertake the action a^*: an agency problem arises. In other words, the principal, using the agent's remuneration, has to create the incentives for the agent to perform the action a^*. This is realized through a contract. In our framework the parties and the enforcing agency (the court) are assumed to have *full* and *symmetric information.* Moreover, we assume that both the agent's actions and the outcomes of these actions are *contractible.* A contract is then a mapping from each action–outcome pair (a,y) into a remuneration I, $I:\mathscr{S} \to \mathbb{R}$, where $\mathscr{S} \equiv [0,1] \times [0,1]$.

We assume that contracts are chosen and signed before the agent undertakes the action. The principal makes a take it or leave it offer to the agent, which the agent accepts or rejects. We assume that the contracting problem is such that the agent accepts in equilibrium. After the binding contract is signed the agent chooses an a and subsequently the outcome y is realized and the remuneration is paid according to the chosen contract. Hence, the principal's choice of an optimal contract $I^*(\cdot,\cdot)$ is a solution of the following problem.

$$\begin{aligned} I^*(\cdot,\cdot) \in \quad & \underset{I(\cdot,\cdot)}{\operatorname{argmax}} \int_0^1 V[y, I(\hat{a},y)]\, \mathrm{d}F(y|\hat{a}) \\ \text{s.t.} \quad & \int_0^1 U[I(\hat{a},y),\hat{a}]\, \mathrm{d}F(y|\hat{a}) \geq 0 \\ & \hat{a} \in \underset{a}{\operatorname{argmax}} \int_0^1 U[I(a,y),a]\, \mathrm{d}F(y|a) \\ & I(a,y) \geq 0 \quad \forall (a,y) \in \mathscr{S} \end{aligned} \tag{1}$$

The first constraint in Problem (1) is the *participation constraint* which guarantees that the agent's expected utility from the contract does not fall below zero. The second constraint in Problem (1) is an *incentive compatibility constraint* which makes the problem consistent with the fact that it is the agent who chooses $a \in [0,1]$. The last constraint is also standard in this type of contracting problem. It stipulates that the remuneration offered to the agent cannot go below zero, whatever the values of a and y. Any other arbitrary lower bound would suffice, the crucial assumption is that the principal cannot threaten to punish the agent infinitely hard for some pairs $(a,y) \in \mathscr{S}$.

On the other hand, to make the problem interesting we need the principal to be able to threaten the agent with some punishment below his outside employment opportunity. To make this possible we assume $U(0,a) < 0$ for all $a \in [0,1]$. In other words, if the agent is paid 0 by the principal, his utility falls below his participation level, whatever the action he chooses to undertake.

A solution to Problem (1) is called *first best.* We assume that the parties preferences $V(y,I)$, $U(I,a)$ and the distribution of outcomes $F(y|a)$ are such that there exists at least one first best contract. Throughout the rest of the paper we will denote the set of solutions to Problem (1) by $\mathscr{I}^*$, with typical element

$I^*(\cdot,\cdot)$. In general, the set $\mathscr{I}^*$ will be quite a large set. This is especially true if the distribution $F(y|a)$ is degenerate (that is when the outcome associated with any given action is deterministically given by a function like $y(a)$).

When $\mathscr{I}^*$ contains many contracts, we focus on a subset of it which we call *simple* contracts. These are the contracts which maximize the principal's expected utility as well as 'minimize' the partitioning of the set $\mathscr{S}$ into more and more subsets giving different values of $I(a,y)$. In other words we concentrate on those contracts which guarantee 'minimal variability' of the remuneration mapping $I(a,y)$ either on pairs (a,y) which will certainly not happen 'in equilibrium' or on subsets of $\mathscr{S}$ of measure zero. Our choice of simple contracts can be interpreted as coming from the existence of a *lexicographic* cost (to be minimized after the principal's expected utility has been maximized) of partitioning more and more the domain $\mathscr{S}$ of the contract.[6]

Given any contract I, let $P(I)$ denote the partition of the space $\mathscr{S}$ induced by I. The notation $P(I') \succ P(I)$ is used to mean that the partition $P(I')$ is strictly coarser than $P(I)$, or equivalently $P(I)$ is strictly finer than $P(I')$.

Definition 1. An optimal agency contract $I^* \in \mathscr{I}^*$ is simple if there does not exist a contract $I^{*\prime} \in \mathscr{I}^*$ such that $P(I^{*\prime}) \succ P(I^*)$.

3. Written contracts

3.1. A model of describability

Any contract the principal chooses for the agent will be effective only if it can be enforced. We focus on situations in which this is possible only if the contract is a *written* agreement between the parties. In many economic situations this is indeed the case and we shall take it simply as given.

Further, we take the view that this has consequences as far as the nature of the *enforcement mechanism* of a contract is concerned. Our written contracts, in fact, contain all the enforcement prescriptions which the court must apply. In other words, our written contracts embody both the actual prescriptions, contingent on actions and outcomes, *and* how they must be enforced. The court is a *passive* subject in this framework; it simply ensures that the prescriptions of the contracts are complied with by the contracting parties.

We view a written contract as a *finite* set of clauses which, given an action and its outcome, yield a remuneration in a finite number of 'steps'. In our context, a finite number of 'steps' can be interpreted as the fact that, examining the contract for a given action and outcome it yields a remuneration in finite time.

[6] For a different equilibrium characterization of simple contracts see Matthews (1995).

It is easy to imagine a written contract that, for instance, 'loops' in the sense that clause α calls on clause β and vice-versa. We exclude such contracts by assumption.

In Anderlini and Felli (1994b), we relate the intuitive notion of a finite set of clauses which must give a result in finite time to the formal notion of *algorithmic function* as embodied in the notion of Turing computability. We will not do this xplicitly here, but refer directly to the class of functions which is generated in this way. Strictly speaking, we will use a wider class of functions than Turing-computable ones, but this only strengthens our results since it relaxes some of the constraints we have to deal with when choosing an optimal contract.[7]

The intuitive justification for the definition of a mapping representing a written contract which we give below is as follows. Imagine that a written contract has been signed and subsequently the agent has chosen an action a and an outcome y has been realized. The values of (a,y) now constitute the 'input' of the written contract which the parties, or the enforcing agency, must 'examine' to decide what remuneration is prescribed in this case by the contract. The 'evidence' consists of the 'digits' of a and similarly for y.

We then imagine that the contract first prescribes the observation of a finite set of digits to look at, and then, as a result of which digits have been observed, of a further set of digits and so on. The crucial requirement is that this 'information gathering procedure' must end in finite time. The finite set of digits looked at during the information gathering stage is called the 'examined evidence' in the instance (a,y).

The next step is to take the examined evidence and using a well defined finite sequence of 'steps', derive the actual remuneration for the agent for the instance (a,y). What constitutes a well defined finite sequence of steps can be formalized using the notion of Turing computability, but this is irrelevant for our purposes here. We will in fact allow any remuneration function which satisfies the restriction that the remuneration for the agent is the same whenever the examined evidence is the same.

Throughout the paper we imagine that the evidence available to the parties, or the enforcing agency, once a and y are realized, is the *binary form* (or binary expansion) of these two real numbers. This way of proceeding has an appealing intuitive interpretation. Both actions and outcomes are 'lists' of a possibly infinite array of different *characteristics*. The binary expansion of, say, a then tells us which ones of these possible characteristics each value of a has. The natural convention is to say that a possesses characteristic i if the ith digit of its binary expansion is 1, and it does not possess the characteristic if the ith digit is 0.

[7] In a previous version of this paper (Anderlini and Felli, 1993) we proved all the results contained in this paper using the exact class of functions given by the restriction of Turing computability.

Although a formal treatment involves a certain amount of notation and lengthy manipulations, it is intuitively clear that since we only allow a written contract to prescribe the observation of a *finite* number of digits of a and y, the resulting remuneration mapping must partition the possible values of a into collections of *non degenerate intervals*, and similarly for the values of y. Indeed, it turns out that the intervals in question must be half-open (open above), in other words they must be of the form $[c,d)$.[8] Moreover, the end-points of the intervals in which actions and outcomes are partitioned must be real numbers which can be written in binary using only a finite number of 1's: all their digits but a finite number of them must be zero.

We are now ready to state formally the restrictions on a remuneration mapping which characterize written contracts. We start with a formal definition of what can be *described* in a contract. In other words, which values can constitute the end points of the action and outcome intervals. We call the real numbers which have a binary expansion with only finitely many 1's *easy to describe* while the other reals will be called *hard to describe.*

Definition 2. A real number is called easy to describe if and only if it has a binary expansion with only finitely many digits equal to 1. Otherwise a real number is called hard to describe.

There is a sense in which easy to describe numbers are rare, but another in which they can be used to approximate any real number, whether hard to describe or not. The following is a consequence of the fact that all easy to describe real numbers are rational and of the fact that easy to describe numbers are 'dense' in the real line.

Remark 1. There is a countable infinity of easy to describe real numbers in the real line. Therefore easy to describe numbers have (Lebesgue) measure zero, while hard to describe numbers have full measure in the real line. Given any real number and any arbitrarily small $\varepsilon > 0$ there exists an easy to describe real number which is within ε of the given real. In other words easy to describe real numbers can be used to approximate any real number.

We can now be precise about what constitutes a written contract. We say that a contract (or remuneration mapping) $I(a,y)$ can be *written*, or simply is a *written contract* if and only if it is a step function which partitions the sets of possible a's

[8] The half-openness of the intervals follows from a choice of which binary expansion to use as evidence when either a or y admit more than one such expansion. The number 1/2 for instance, can be written in binary as either 0.10000 ... or 0.01111 ... (with the last digit being repeated forever in both cases). In such cases we imagine that, arbitrarily, the expansion with the *least* number of 1's is chosen. Choosing the one with the *most* number of 1's would also have yielded half-open intervals, but open below rather than above.

and y's into a collection of disjoint half-open (above) intervals with easy to describe end-points. In other words, using the interpretation of easy to describe values introduced above a written contract has to contain an exact description (a complete list of the defining characteristics) of the relevant values of the performance, outcome and remuneration accruing to the agent. Formally, this implies the following definition.

Definition 3. A remuneration mapping $I: \mathscr{S} \mapsto \mathbb{R}$ constitutes a written contract if and only if there exist two (possibly infinite) sets of easy to describe real numbers $(a_0, \ldots, a_i, \ldots)$ and $(y_0, \ldots, y_j, \ldots)$, and an array of real numbers I_{ij} such that

$$I(a,y) = I_{ij} \quad \text{whenever } a_{i-1} \leq a < a_i \text{ and } y_{j-1} \leq y < y_j \tag{2}$$

Throughout the rest of the paper, the set of all possible written contracts is denoted by $\mathscr{W}$.

3.2. Describability vs. observability

Before we proceed any further, it is legitimate to ask whether (lack of) describability in the sense we have defined and (lack of) observability in the sense of standard principal-agent theory are in fact the same thing. Notice that, by observable in the sense of standard principal-agent theory we mean observable by all parties, court included. In the jargon of contract theory such variables are often called *verifiable*.

We contend that, although related, these two notions are distinct. Intuitively, the distinction between observability and describability (in a written form) is a simple one. The economic environment determines which variables are observable and which ones are not. The 'input' to a written contract are the observable variables.[9] The restrictions placed on the contract by the fact that it must be written are what embodies our notion of describability. In particular, a written contract can only specify outcomes which 'change' on the basis of describable values of the observable variables.

To dwell further on this point, let us focus attention on describability vs. observability of the agent's action. Analytically, the fact that the agent's action cannot be observed, or can only be imperfectly observed, corresponds to a *fixed partition* of the agent's action set. Only the element of the given partition corresponding to the action taken is observable and can be included in the contract.

The constraint imposed by the requirement that the contract be a written one is plainly far from measurability with respect to a given, *fixed* partition of the agent's action set. Indeed all that the writing restriction imposes is that the

[9] In principle, of course, the parties could decide to include unobservable variables in a written contract. This, however, would have no effect on the outcome of the contract since these variables could not be used by the enforcing agency (the court) as evidence to decide what the contract prescribes.

contract must partition the agent's action set in *some* way so that the relevant end-points are actions which are easy to describe in the sense of Definition 2.

In particular the 'fineness' of the partition is *endogenous* in the writing restriction case. If there is any further useful information which can be incorporated into the contract by refining the partition of the agent's action set, and still have easy to describe end-points, then the contracting parties are free to choose a contract which takes it into account.

The fact that the writing constraint and observability constraint can give rise to very different contracts in some environments at least is apparent from the approximation result which we present in the next section contrasted to the 'endogenous principal-agent' result of Section 6. The approximation result tells us that if the contracting problem is 'sufficiently continuous' then the writing restriction has an arbitrarily small effect on the parties' expected utilities, relative to the case in which the agent's action is perfectly observable.

The *same* writing restriction imposed on a contracting problem with noise and suitable discontinuities, turns out to yield the same contract that the parties would write if the agent's action were not observable at all to the principal.

Clearly the two constraints are not equivalent. According to the type of environment, the writing restriction can be seen to have the same effect as the two opposite extreme assumptions about the observability of the agent's action.

4. Does the writing restriction matter?

4.1. The approximation result

The first question we ask in the framework described by the previous sections is whether the contracting parties are at all constrained by the restriction that they must use a written contract. The first answer we obtain is a negative one. Whenever the principal's and the agent's preferences and choice of action to undertake are continuous (in a sense to be made precise below) the parties will write agency contracts which approximate, in terms of the parties' expected utilities, the first best contract.

Assume, for simplicity, that Problem (1) has a unique solution I^* (the first best contract) which is 'well behaved' ($I^*(a,y)$ can be integrated over y and a). Assume now that the principal's and the agent's preferences are continuous in both their arguments. Assume further that the technology described by the family of conditional distributions $F(y|a)$ is continuous in the sense that the distribution of y changes by a 'small amount' for small variations in a.[10]

Assume further that the first best contract I^* has the following property. Given a contract I, which is 'close' to I^*, the action chosen by the agent when

[10] Technically, continuity in the 'weak topology' is what we have in mind.

offered the contract I, is close to the action chosen given the first best contract.[11] In other words, assume that the agent's action is continuous in the reward function, near the first best contract. Notice that while the first set of continuity assumptions we have listed are all 'primitive' in the sense that they concern the 'data' of the contracting problem—technology and preferences—we are now assuming a property which concerns an *endogenously* determined function, namely the solution to Problem (1). The status of these two sets of assumptions is clearly different and we will return to this question shortly.

Before any further discussion we state what we have called above the approximation result.

Proposition 1. *Under the continuity assumptions stated above, it is always possible to approximate arbitrarily closely the expected utilities which the parties achieve through the first best contract using a written contract.*

Instead of proving Proposition 1 formally, we describe intuitively how a proof can be constructed and refer to Anderlini and Felli (1994b) for the full formal details. Given the first best contract I^*, it is always possible to approximate $I^*(a,y)$ with a written contract giving a step function $I(a,y)$ in the sense that the distance between $I^*(a,y)$ and $I(a,y)$ will not exceed some small number, for any pair $(a,y) \in \mathcal{S}$.[12] Moreover, since any real number can be approximated by an easy to describe real number, we can ensure that the end-points of the steps of I are easy to describe reals. Finally, it is not hard to see that the intervals in which I partitions the sets of possible a's and y's can be taken to be half open above.

In other words, given a well behaved first best contract, we can always find a written contract which is arbitrarily close to it in the sense defined above. Given that I is close to I^*, we know by assumption that the agent's action under contract I will be close to his action under contract I^*. Since the technology and the preferences are continuous, it is now clear that the expected utility achieved by both parties under the contract I is arbitrarily close to what they, respectively, achieve under the first best contract I^*.

4.2. The rôle of continuity

An intuitive way to restate the continuity assumptions which give the approximation result is the following. Let $\hat{a}^*$, be the action which the agent chooses so as to maximize his expected utility, given that he is offered the first best contract I^*. For the approximation result to hold we then need first of all that the

[11] Technically we can think of the distance between two contracts as the 'supremum' over a and y of the distance between the values of the agent's rewards in the two contracts.

[12] Recall that we are assuming a unique first best which is integrable as a function of a and y.

principal's expected utility, given by $\int_0^1 V[y,I(a,y)]\,dF(y|a)$, be continuous as a function of a over a small interval around $\hat{a}^*$. For this, continuity of the principal's preferences and of the technology will suffice.

Secondly, we need the agent's choice of utility maximizing action to be continuous in the remuneration function, again over a small neighbourhood of the first best remuneration function I^*. This is essentially a property of the endogenously determined first best remuneration function, and it will ensure that the principal's utility does not vary much for small variations in the remuneration function.

Lastly, we need the agent's expected utility not to vary too much when the remuneration function is changed by a small amount, and the agent's own choice of action changes (by a small amount) as a result. Continuity of the agent's preferences and of the technology will suffice for the last continuity requirement.

Let us start with the requirement that the first best remuneration schedule be such that the agent's action is continuous in the remuneration schedule itself for small variations around the first best. From the existing literature (Grossman and Hart, 1983), it is clear that a discontinuity in the agent's choice of action around the first best remuneration function is far from a pathological feature of the standard principal-agent model.

In Fig. 1 we have depicted a situation in which given I^* the agent's expected utility as a function of his own action has two distinct peaks yielding the same expected utility. The value $\hat{a}^*$ is the agent's *unique* action in the first best,

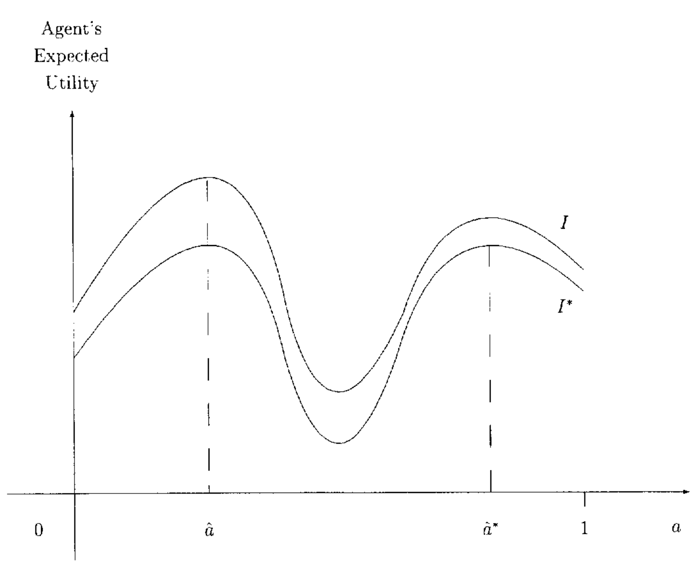

Fig. 1. Example of endogenous discontinuities.

however. It is now clear that as we deform continuously the function giving the agent's expected utility as a function of a (as a result of a small change in the remuneration function, away from the first best I^*), the agent's choice of action may 'jump' to a value like $\hat{a}$ in Fig. 1, which is *not close* to the first best value $\hat{a}^*$. In this sense, the assumption of a continuous reaction of the agent to small changes in the remuneration schedule is clearly not satisfied in general.

Let us now look at the rest of the continuity assumptions necessary for the approximation result. We do not consider the assumption of continuity of both parties' preferences in the remuneration schedule a particularly bad one. The standard interpretation of the reward function I is that of a *transfer* (real or monetary) from the principal to the agent. It is therefore difficult to take issue with the assumption that the parties' utility should be continuous in the amount transferred.

Continuity of the principal's utility in the outcome y, and analogously continuity of the agent's utility in the action a, are much less compelling assumptions in our view.

For concreteness, let us focus on preferences over the outcome y. Recall that we interpret the binary expansion of y as a list of 'characteristics'. It is then plausible that such characteristics could have a lexicographic ordering (different from the 'natural' one), thus making the principal's preferences discontinuous. For instance, the principal could be commissioning a bridge to the agent, and its structural characteristics could matter to him discontinuously more than its aesthetic qualities.

The last continuity assumption we need for the approximation result to hold concerns the technology which transforms the action of the agent into (a probability distribution over) outcomes. Again, we find this assumption less than compelling, at least for some contracting environments one can imagine as relevant.

Consider for example the case in which the agent is hired to write a computer program containing a large number of 'lines'. A change in a single line of the program may change its output by a large amount (as it is well known to anyone who has ever attempted to write a computer program of any significant complexity!) Hence, a small variation in the agent's action may trigger a large change in the distribution of outcomes.

A further example is an agent hired to fill out the principal's income tax return. A completely correct tax return may produce a very different outcome from a tax return with a 'small' mistake in it which may be nevertheless pursued vigorously by the tax authorities.

In the following sections we explore the consequences of discontinuities in the principal's preferences in outcomes and in the technology when the output is deterministic. Lastly, we turn to the case in which the discontinuities occur in the agent's preferences in the action he is supposed to undertake and in the outcome of the technology which is taken to be a non-degenerate probability distribution. It will turn out to be crucial whether or not the discontinuities occur at a point which is hard to describe according to Definition 2.

5. Discontinuities with deterministic output

In this section we construct two examples of discontinuities in the principal's preferences and in the technology, concentrating on a simple version of the framework introduced in Section 2. We assume that the action the agent undertakes is transformed in a contractible outcome through a deterministic technology in the sense that the conditional distribution $F(y|a)$ is degenerate. Hence, the agent controls a simple production function $y = y(a)$ which associates an outcome y to each action a.

5.1. Discontinuous preferences

We start with the case in which the discontinuities arise in the utility the principal derives from the outcome of the agent's action. In other words, we assume that the function $V(y,I)$ may be discontinuous in y. To make the analysis as simple as possible we further assume that the agent's action coincides with the contractible outcome which is relevant for the principal. The technology controlled by the agent is then described by the following equality:[13]

$$y = a. \tag{3}$$

Among the actions which the agent may undertake, there exists one a^* which the principal most prefers. Any other action yields a discontinuously lower utility to the principal. These discontinuous preferences can be described by the following utility function.

$$V(y,I) = \begin{cases} H + V_I(I), & \text{if } y = a^* \\ V_I(I), & \text{otherwise} \end{cases} \tag{4}$$

where $H > 0$.

Since $a = y$, the remuneration mapping can just be taken to be a map from $a \in [0,1]$ into the non-negative reals. Therefore, the principal's problem which in general we have stated as in Problem (1), may now be re-written as follows

$$\begin{aligned} \max_{I(\cdot)} \quad & V[\hat{a},I(\hat{a})] \\ s.t. \quad & U[I(\hat{a}),\hat{a}] \geq 0 \\ & \hat{a} \in \underset{a}{\operatorname{argmax}} \; U[I(a),a] \\ & I(a) \geq 0 \;\; \forall a \in [0,1] \end{aligned} \tag{5}$$

[13] Because of this assumption, throughout the rest of the example we use the words action and outcome as synonyms.

One of the simple first best contracts that solve problem (5) is then

$$I^*(a) = \begin{cases} T^* > 0, & \text{if } a = a^* \\ 0, & \text{if } a \neq a^* \end{cases} \tag{6}$$

where $T^* > 0$ is implicitly defined by $U(T^*,a^*) = 0$. Intuitively, the principal gives the agent an acceptable remuneration only if the agent performs the action the principal most prefers, in any other case the agent receives the lowest possible amount which the principal is allowed to pay.

We are now ready to impose the additional constraint that the contract which binds the parties must be a written contract. Recall that $\mathscr{W}$ is the set of all possible written contracts. The optimization problem giving the optimal written contract can be written as

$$\begin{aligned} \max_{I(\cdot)} \quad & V[\hat{a}, I(\hat{a})] \\ \text{s.t.} \quad & U[I(\hat{a}),\hat{a}] \geq 0 \\ & \hat{a} \in \operatorname*{argmax}_a \; U[I(a),a] \\ & I(a) \geq 0 \;\; \forall a \in [0,1] \\ & I(\cdot) \in \mathscr{W} \end{aligned} \tag{7}$$

The effect on the solution to Problem (5) of the constraint that $I(\cdot)$ must be in $\mathscr{W}$ is quite dramatic, as the following shows.

Proposition 2. *Suppose that the outcome y^* most preferred by the principal is hard to describe, and that the principal's preferences are discontinuous as described in Eq.* (4). *Then there exists a unique optimal written contract* (*a solution to Problem* (7)) *which is simple according to Definition* 1. *This contract is given by*

$$I(a) = \bar{T} > 0 \qquad \forall a \in [0,1] \tag{8}$$

where denoting by a^{**} the action most preferred by the agent,[14] *$\bar{T}$ satisfies*

$$U(\bar{T},a^{**}) = 0. \tag{9}$$

Proof. Since the parties must use a written contract, the remuneration schedule must be a step function with easy to describe end-points. Since $a^* = y^*$ is hard to describe by assumption, it cannot be the end point of any of the intervals in which any written contract partitions the set of possible actions. Since the remuneration of the agent does not change within each of the intervals of any step function, the agent will choose an action which maximizes $U_a(a)$ within one of the intervals of the step function given by a written contract. Consider the

[14] Recall that we are assuming that the agent has separable preferences.

interval containing a^*. Since a^* cannot be the end point of any interval for any written contract, using the fact that the parties have a local conflict of interest around a^* (Assumption 1), it is clear that for any written contract, the agent's choice of action will be *different* from a^*. Therefore using Eq. (4) for any written contract, the principal's utility will be given by $V_I(I(\hat{a}))$, where $\hat{a}$ is the action chosen by the agent given the written contract in force. Recall that $V_I(\cdot)$ is monotonically decreasing in the amount paid to the agent. Therefore, the optimal written contract must minimize $I(\cdot)$ for the chosen action $\hat{a}$. It follows that any written contract solving Problem (7) must satisfy $U(I(\hat{a}),\hat{a}) = 0$. But the principal is indifferent between any of the outcomes that can be sustained by any written contract. Therefore, he may as well offer a contract which sustains a^{**} as the agent's action since this generates the lowest payment to the agent, $U(I(a^{**}),a^{**}) = 0$. Indeed, since action a^{**} is the one the agent most prefers it is induced by a remuneration function which does not vary with a. It follows that the unique *simple* contract solving Problem (7) is as in Eq. (8). □

Before moving on to our next example, we notice that Proposition 2 could have been formulated in a more general way simply at the expense of additional notation. In particular, there is nothing in the assumption that a *single* action gives the principal discontinuously higher utility than the rest. A whole set (possibly with full measure) of most preferred actions could have been accommodated, provided that they were all hard to describe.

It is best to delay further the discussion of Proposition 2 until we are able to compare it to our next example below.

5.2. Discontinuous production function

We now move to a different situation in which the discontinuities occur not in the principal's preferences but in the technology that transforms contractible actions into outcomes, $y(a)$. We assume that the principal's preferences $V(y,I)$ are continuous and monotonic in the outcome of the agent's action y. Moreover, we assume that the production function is the following discontinuous function.

$$y(a) = \begin{cases} y^* > 0, & \text{if } a = a^* \\ 0, & \text{if } a \neq a^* \end{cases} \tag{10}$$

Under this new set of assumptions the principal's Problem (1) can be re-written as

$$\begin{aligned} \max_{I(\cdot,\cdot)} \quad & V[y(\hat{a}),I(\hat{a},y(\hat{a}))] \\ s.t. \quad & U[I(\hat{a},y(\hat{a})),\hat{a}] \geq 0 \\ & \hat{a} \in \operatorname*{argmax}_a \; U[I(a,y(a)),a] \\ & I(a,y) \geq 0 \;\; \forall (a,y) \in \mathcal{S} \end{aligned} \tag{11}$$

where the remuneration schedule $I(\cdot,\cdot)$ is now again a function of both the agent's action a and its outcome y. The following remuneration schedule is one of the many first best contracts that solve Problem (11).

$$I^*(a,y) = \begin{cases} T^*, & \text{if } a = a^* \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

where T^* is defined by $U(T^*,a^*) = 0$, as in our previous example. The remuneration schedule above depends only on the agent's action, and not on its outcome. The agent is induced to take action a^*, since this is the only way he can achieve an acceptable remuneration. Any other action will lead him to be punished by the principal to the maximum possible extent. Clearly, since both actions and outcomes are contractible variables, it is also possible to find first best remuneration schedules which have the opposite feature of being dependent on outcomes only, and not on the agent's actions. An example is the following,

$$I^*(a,y) = \begin{cases} T^*, & \text{if } y \geq \bar{y} \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

where $0 < \bar{y} \leq y^*$. This is a contract which induces the agent to pick action a^* by means of a threshold value on outcomes. If the value of y exceeds the threshold $\bar{y}$, then the agent gets an acceptable remuneration, otherwise he is 'punished' to the maximum possible extent. Clearly, given the discontinuous technology described in (10) any threshold value greater than 0 and less than or equal to y^*, will have the same effect on the agent's action, forcing him to pick a^*.

It is easy to see that many other first best contracts exist, which are qualitatively a mixture of (12) and (13). For instance, a first best contract is one which gives the agent an acceptable remuneration only if both conditions $a = a^*$ and $y = y^*$ are satisfied, while punishing the agent to the maximum extent for any other action–outcome combination possible.

The effect on the solution to Problem (11) of the additional constraint that the chosen contract must be a written one is rather less dramatic than in our previous example. The optimal written contract which the principal will choose will be one of the first best contracts solving the unrestricted Problem (11). However, if the action a^* producing the good outcome y^* is hard to describe, the optimal (simple) written contract will *not* be contingent on actions. Thus, the constraint that the parties must choose a written contract determines the 'shape' of the contract chosen, among many possible first best contracts.

Problem (11), with the additional constraint that a written contract must be used by the parties is written simply as

$$\begin{aligned} \max_{I(\cdot,\cdot)} \quad & V[y(\hat{a}),I(\hat{a},y(\hat{a}))] \\ \text{s.t.} \quad & U[I(\hat{a},y(\hat{a})),\hat{a}] \geq 0 \\ & \hat{a} \in \operatorname*{argmax}_a \; U[I(a,y(a)),a] \\ & I(a,y) \geq 0 \;\; \forall (a,y) \in \mathscr{S} \\ & I(\cdot,\cdot) \in \mathscr{W} \end{aligned} \tag{14}$$

The set of written contracts solving Problem (14) can be characterized as follows.

Proposition 3. *The solution to Problem* (14) *is one of the first best contracts solving Problem* (11). *Moreover, if the action a^* producing the best outcome is hard to describe, any simple contract solving Problem* (14) *is not contingent on actions and is of the same form as Eq.* (13).

Proof. Clearly, any contract of the same form as Eq. (13) is a written contract, provided that $\bar{y}$ is easy to describe. Any contract of the same form as (13) achieves first best utility for the parties, provided that $0 < \bar{y} \leq y^*$. By Remark 1 we can always choose a threshold $\bar{y}$ which is easy to describe. This proves that any contract solving Problem (14) also solves Problem (11). If a^* is hard to describe, contracts contingent on actions can be ruled out in exactly the same way as in the proof of Proposition 2. Using the fact that the parties have a local conflict of interest around a^* (Assumption 1), it is clear that any contract which partitions the set of possible y's in a strictly coarser way than contracts of the same form as Eq. (13), cannot possibly achieve the first best level of utility, since it will be contingent on neither actions nor outcomes. Therefore all simple written contracts solving Problem (14) must be of the same form as Eq. (13) whenever a^* is hard to describe. □

5.3. *Comparison*

We conclude this section with a comparison of the two examples of discontinuities we have analyzed so far. The difference between the effects of imposing written contracts in the two cases is quite dramatic. In the case of discontinuous preferences if the action producing the good outcome is hard to describe, we get an optimal written contract which is very far from the first best. The principal chooses not to affect at all the choice of action of the agent and chooses a flat remuneration function. In the case of discontinuous technology on the other hand, the parties achieve their first best level of expected utility using a written contract, regardless of whether the action producing the good outcome is hard or easy to describe.

The main difference between the two cases can be thought of as follows. In both cases, a discontinuity occurs at an action which is hard to describe, and therefore impossible to enforce directly by means of a written contract. In the case of discontinuous technology, however, there is an alternative contractible variable (the outcome y) which 'jumps' at the discontinuity in a. Therefore a written contract can prescribe a punishment if the 'jump does not occur' and an acceptable remuneration if it does, thus enforcing the choice of a^* on the part of the agent.

In the case of discontinuous preferences, the only thing that 'jumps' at a^* is the principal's utility which is *not* a contractible variable by assumption. In our extreme example, there is no hope of using the outcome as a 'signal' of the action since the good outcome is just as hard to describe as a^* for our degenerate technology $a = y$.

In all cases, except the one in which an easy to describe outcome is associated with the good action a^*, attempting to induce the choice of a^* using outcome-based written contracts is worthless to the principal. The main culprit for the dramatic effect of the constraint that contracts must be written in our first example is the *non-contractibility* of the principal's utility level, coupled of course with the fact that the good action and outcome are hard to describe.

6. Noisy output: a principal-agent problem

In this section we consider the general case in which the technology through which actions are transformed into outcomes is stochastic, and it is described by the non-degenerate conditional distribution $F(y|a)$. We further assume that there exists a discontinuity in the agent's preferences as well as in this stochastic technology and we explore the type of optimal written contracts the parties choose in such a situation.

We show that in a noisy environment the discontinuity in the agent's preferences and in the technology are enough to make the best written contract the parties can write sub-optimal relative to the first best. Moreover, in our example, the optimal simple written contract takes the same form as the principal-agent contract obtained by the moral hazard literature (Holmström, 1979). We view this principal-agent problem as stemming entirely from describability problems since we do not assume any asymmetry of information among the contracting parties and/or the court on the action the agent undertakes.

For simplicity, we take the distribution over outcomes to have a discrete finite support $\{0, \dots, y_n, \dots, 1\}$, which is the same for any $a \in [0,1]$. So, the choice of a affects only the probabilities of each of the discrete outcomes.[15] We assume that the agent's action a^* yields a discontinuously 'better' distribution of

[15] Assuming finite support greatly simplifies the analysis, but is inessential to the qualitative nature of our results. We return to this point below.

outcomes than any other feasible action. In formal terms we take 'better' to mean that the agent's choice of action a^* yields a conditional distribution which dominates, in a first order stochastic sense, the conditional distribution generated by any other choice of action $a \neq a^*$. The technology we are assuming can therefore be summarized as follows

$$F(y|a) = \begin{cases} F^*(y), & \text{if } a = a^* \\ F(y), & \text{if } a \neq a^* \end{cases} \tag{15}$$

$$F(y) > F^*(y) \qquad \forall y \in (0,1) \tag{16}$$

We further depart from Assumption 1 in making the conflict of interests between the principal and the agent extreme. Indeed, we assume that the agent's utility is discontinuously lower at the action a^* and independent of a whenever $a \neq a^*$. These discontinuous preferences can be described by the following utility function.

$$U(I,a) = \begin{cases} K + U_I(I), & \text{if } a = a^* \\ U_I(I), & \text{otherwise} \end{cases} \tag{17}$$

where $K < 0$.

One canonical example of a principal-agent problem is the one in which the principal is risk neutral and the agent is risk-averse. For simplicity we concentrate on this case. Thus, the principal's utility function is linear in both the outcome y and the agent's remuneration I: $V(y,I) = y - I$, and the agent's utility function is strictly concave in the remuneration I: $U_I''(I) < 0$.

Recall that both outcomes and actions are contractible. The principal's problem can therefore be written as

$$\begin{aligned} &\max_{I(\cdot,\cdot)} \quad \int_0^1 [y - I(\hat{a},y)] \, \mathrm{d}F(y|\hat{a}) \\ &s.t. \quad \int_0^1 U[I(\hat{a},y),\hat{a}] \, \mathrm{d}F(y|\hat{a}) \geq 0 \\ &\qquad \hat{a} \in \operatorname*{argmax}_a \int_0^1 U[I(a,y),a] \, \mathrm{d}F(y|a) \\ &\qquad I(a,y) \geq 0 \ \ \forall (a,y) \in \mathscr{S} \end{aligned} \tag{18}$$

It is easy to check that Problem (18) yields a unique *simple* first best contract. This is

$$I(a,y) = \begin{cases} T^* > 0, & \text{if } a = a^* \\ 0, & \text{otherwise} \end{cases} \tag{19}$$

where T^* is defined by $U(T^*,a^*) = 0$ as in Section 5.

In the canonical principal-agent problem it is assumed that the agent's remuneration mapping does not depend on a because of the *non-observability* of the agent's action: I maps from $y \in [0,1]$ into the non-negative reals. The form of the principal's problem in this case is

$$\begin{aligned} \max_{I(\cdot)} \quad & \int_0^1 [y - I(y)]\, \mathrm{d}F(y|\hat{a}) \\ s.t. \quad & \int_0^1 U[I(y),\hat{a}]\, \mathrm{d}F(y|\hat{a}) \geq 0 \\ & \hat{a} \in \operatorname*{argmax}_a \int_0^1 U[I(y),a]\, \mathrm{d}F(y|a) \\ & I(y) \geq 0 \quad \forall y \in [0,1] \end{aligned} \tag{20}$$

According to the different assumptions made on the technology, preferences and outside opportunities of the parties, a solution to a problem of the form of Problem (20) may or may not exist (see for instance Grossman and Hart (1983)), and the solution may or may not give the principal an expected utility level strictly below the first best. We have made enough assumptions above to guarantee that we are in the standard well behaved case.

Remark 2. A solution to Problem (20) exists and it is sub-optimal relative to the solution to Problem (18).

This sub-optimality is intuitively due to the inability of the principal to simultaneously 'insure' the agent against the variation in y and induce him to take action a^* using a remuneration mapping which does not depend on a.

We now impose the restriction that the contract which regulates the parties' relationship must be a written one. As usual, we maintain the assumption that both a and y are variables which are observable, and hence contractible in principle. Because of the describability constraints which are embodied in the definition of a written contract, the principal's problem becomes

$$\begin{aligned} \max_{I(\cdot,\cdot)} \quad & \int_0^1 [y - I(\hat{a},y)]\, \mathrm{d}F(y|\hat{a}) \\ s.t. \quad & \int_0^1 U[I(\hat{a},y),\hat{a}]\, \mathrm{d}F(y|\hat{a}) \geq 0 \\ & \hat{a} \in \operatorname*{argmax}_a \int_0^1 U[I(a,y),a]\, \mathrm{d}F(y|a) \\ & I(a,y) \geq 0 \quad \forall (a,y) \in \mathscr{S} \\ & I(\cdot,\cdot) \in \mathscr{W} \end{aligned} \tag{21}$$

It is possible to characterize the form of the best written simple contract and to show that it is essentially the same as the solution to the standard principal-agent Problem (20).

Proposition 4. *Problem* (21) *has a solution, regardless of whether* a^* *is easy or hard to describe. If* a^* *is easy to describe the solution to Problem* (21) *yields the parties their first best level of expected utility. If the good action* a^* *is hard to describe, then all simple written contracts solving Problem* (21) *do not depend directly on the agent's action a. Moreover, all simple written contracts prescribe the same reward for the agent as the contract solving Problem* (20) *for all y in the support of the distribution of outcomes* $\{0,\dots,y_n,\dots,1\}$.

Proof. The case in which a^* is easy to describe is straightforward. The claim is obvious from the observation that a contract like Eq. (19) is a written contract if a^* is easy to describe. Consider now the case of a^* hard to describe, and assume temporarily that a simple best written contract I exists. Two cases are possible. Either contract I induces $\hat{a} \neq a^*$ or it induces $\hat{a} = a^*$. Consider the former case first. Since $\hat{a} \neq a^*$ the principal is only interested in minimizing the cost of meeting the agent's participation constraint. It then follows that a 'flat' contract which depends neither on a nor on y, and which guarantees an expected utility of 0 to the agent when he takes his most preferred action, is the best simple written contract.[16] Suppose now that I induces $\hat{a} = a^*$. Notice that since a^* is assumed to be hard to describe, it cannot be an end-point of any of the intervals in which I divides the set of possible actions. It follows that the solution to the agent's maximization problem is 'interior' to these rectangles. Hence, by Eq. (17), relaxing the contract's prescriptions in terms of actions will not change the agent's behaviour. In other words, a new contract $I'(a,y) = I(a^*,y)$ for every $(a,y) \in \mathscr{S}$ which is not contingent on a, also induces $\hat{a} = a^*$ and hence yields to the principal at least as much expected utility as I.

Notice now that there always exists a written contract which does not depend on a and which gives the agent the same remuneration as the solution to Problem (20) for any outcome level y_n in the support of the distribution of outcomes. This is simply because the written contract can be taken to be a step function with easy to describe cut off levels $\bar{y}_n$ each strictly between y_{n-1} and y_n, and giving the same level of remuneration as prescribed by the solution to Problem (20) for outcome y_n over the entire interval $[\bar{y}_n, \bar{y}_{n+1})$. Thus we have shown that in all cases a simple written contract solving Problem (21) does not depend on a. □

We conclude with an observation about our assumption that the distribution of outcomes is discrete and finite. To see its relevance to our result, consider the

[16] Recall that the agent is assumed to be risk-averse.

polar case of a distribution of outcomes with a continuous density. Clearly, the part of the proof of Proposition 4 which claims that given any contract not dependent on actions a written contract can be found which is equal for all outcomes in the support simply does not hold anymore. Proposition 4 as stated is false if the distribution of outcomes admits a continuous density.

It is clear, however, that the argument which guarantees that the optimal simple written contract does not depend on a in the proof of Proposition 4 still holds in this case. It is also clear that with a continuous density over outcomes, a written contract can be made to approximate with a step function any contract (integrable over y) solving the standard principal-agent Problem (20). In this sense Proposition 4 above, can be said to hold for an 'approximating sequence' of written contracts which in the limit again look like the solution to Problem (20).

7. Concluding remarks

This paper explores the consequences of imposing a restriction on the remuneration mapping of the agent in a general principal-agent contracting problem. The restriction is meant to embody the requirement that the contract between the principal and the agent must be a written one. We require the remuneration mapping to be a step function which partitions the possible values of the contracting variables into half-open non degenerate intervals.

If the contracting problem satisfies some (in our opinion not generally justified) continuity properties, the restriction to written contracts is essentially of no consequence to the parties in terms of their expected utility.

On the other hand we have found that, if discontinuities arise in the original contracting problem the effect of the restriction to written contracts can change the parties expected utility quite dramatically. When stochastic outcomes depend discontinuously on the agent's action and the latter's preferences are discontinuous we found that the parties will choose a contract which is the same as the canonical principal-agent contract when the agent's action is not observable. This is so notwithstanding the fact that the agent's action is contractible in principle. The written contracts we have found in two of our examples do not distinguish sufficiently finely among different possible values of the contracting variables.

To conclude, it is useful to comment on some possible generalizations of our model. Roughly speaking, the approximation result hinges on the fact that the writing restriction allows the parties to 'approximate' any contract they want precisely in the same topology (the standard Euclidean topology) in which the contracting problem is continuous. Similarly, the departures from the first best contract in our first and third examples, are due to the fact that the contracting problem is sufficiently discontinuous, precisely in the topology in which the writing restriction allows the parties to approximate any contract they want.

We consider the Euclidean topology an important and interesting example; as we pointed out in Section 4.2 we believe there are economically interesting examples which yield precisely the discontinuities needed for our results. It is clear however, that both the approximation results and the flavour of our three examples could be generalized to a model in which another topology is imposed on the contracting variables.

In other words, if using some other model of describability one obtains that the parties can approximate any contract they want in a given topology, and the contracting problem happens to be 'sufficiently continuous' in such topology, then the equivalent of the approximation result would hold. Similarly, if the contracting problem were to be sufficiently discontinuous in the topology generated by the describability model, then dramatic departures from the first best, analogous to our first and third examples above, would hold in the appropriate cases.

Acknowledgements

We wish to thank two anonymous referees, Philippe Aghion, Giuseppe Bertola, Patrick Bolton, Francesca Cornelli, Christopher Harris, Oliver Hart, Preston McAfee, John Moore, Ben Polak, Klaus Schmidt and seminar participants at the University of Chicago Graduate School of Business and the Institute for Advanced Studies of the Hebrew University of Jerusalem. Any remaining errors are, of course, our own responsibility. A revision of this paper was prepared while the first author was visiting the Cowles Foundation for Research in Economics at Yale University. Their generous hospitality is gratefully acknowledged.

References

Anderlini, L., Felli, L., 1993. Incomplete written contracts: endogenous agency problems. Theoretical Economics Discussion Paper TE/93/267, STICERD, London School of Economics.

Anderlini, L., Felli, L., 1994a. Endogenous agency problems. Economic Theory Discussion Paper 200, Department of Applied Economics, University of Cambridge.

Anderlini, L., Felli, L., 1994b. Incomplete written contracts: undescribable states of nature. Quarterly Journal of Economics 109, 1085–1124.

Calamari, J.D., Perillo, J.M., 1987 Contracts. West Publishing Company, St. Paul, MN.

Cutland, N.J., 1980. Computability: An Introduction to Recursive Function Theory. Cambridge University Press, Cambridge.

Grossman, S.J., Hart, O.D., 1983. An analysis of the principal-agent problem. Econometrica 51, 7–45.

Grossman, S.J., Hart, O.D., 1986. The costs and benefits of ownership: A theory of vertical and lateral integration. Journal of Political Economy 94, 691–719.

Hart, O.D., Holmström, B., 1987. The theory of contracts. In: Bewley, T.F. (Ed.), Advances in Economic Theory. Fifth World Congress, Cambridge University Press, Cambridge.

Hart, O.D., Moore, J., 1988. Incomplete contracts and renegotiation. Econometrica 56, 755–785.

Hart, O.D., Moore, J., 1990. Property rights and the nature of the firm. Journal of Political Economy 98, 1119–1158.

Holmström, B., 1979. Moral hazard and observability. Bell Journal of Economics 10, 74–91.

Matthews, S.A., 1995. Renegotiation of sales contracts. Econometrica 63, 567–6590.