European Economic Review 46 (2002) 725–733
EUROPEAN ECONOMIC REVIEW
www.elsevier.com/locate/econbase

Incomplete Contracts

On indescribable contingencies and incomplete contracts☆

Eric Maskin *

Institute for Advanced Study and Princeton University, School of Social Science, Einstein Drive, Princeton, NJ 08540, USA

Abstract

I examine the theoretical foundations underlying the incomplete contracts literature. A common justification for the assumption that contracts are not fully contingent on the state of nature is to point out that some aspects of the state may be unforeseen or indescribable to the contracting partners at the time the contract is written. I argue, however, that as long as risk-averse parties can foresee the probabilities of their possible payoffs, then the fact that they cannot describe the possible physical states does not matter; even with renegotiation, the parties can attain the same welfare as when full description is possible. © 2002 Elsevier Science B.V. All rights reserved.

JEL classification: L14

Keywords: Incomplete contracts; Unforeseen contingencies; Implementation theory

1. Introduction

The literature on incomplete contracts (see Hart (1995) and Tirole (1999) for surveys) has provided significant insight into how assignment of ownership of productive assets bears on economic outcomes. In this short paper I do not intend to quarrel with any of the literature's conceptual accomplishments, which I regard as valuable and interesting. Nevertheless, I wish to take issue with some of the theoretical foundations underlying this line of work.

The points that I will make are not new; most have been made in one form or another in Maskin and Tirole (1999) or Maskin and Moore (1999).[1] However, it is

☆ This paper is based on a talk given at the European Economic Association 2001 meeting in Lausanne.
* Tel.: 609/734-8309; fax: 609/951-4457.
E-mail address: maskin@ias.edu (E. Maskin).

[1] The latter paper is not concerned with incomplete contracts per se, but develops techniques that I use in the mechanism of Section 4.

0014-2921/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0014-2921(01)00209-4

fair to say that the discussion so far has taken place at a rather technical level. Herein, I will try to present some of the main ideas in a reasonably informal way.

2. Incomplete contracts

If I am to criticize the incomplete contracts literature, I must first say what an incomplete contract is. Rather than attempting a precise definition – although such attempts have been made – I will consider a contract to be "incomplete" if it is not as fully contingent on the "state of the world" (the resolution of uncertainty about the future) as the parties to the contract might like it to be. [2]

Imagine, for example, that two agents plan to trade at some time in the future. Before this happens, they must know the characteristics of the good to be exchanged. Suppose that these characteristics are still undetermined at the time the parties negotiate their contract. Then different states of the world will correspond to different characteristic specifications. And if the terms of trade in the contract do not depend on the state, the contract might reasonably be called "incomplete".

The literature offers three main reasons for contractual incompleteness:

(1) Some aspects of the state of the world may not be *common knowledge* or commonly observable; in particular, whoever is responsible for enforcing the contract (e.g., the court) may not be able to ascertain these aspects (in which case, we say that the aspects are "unverifiable");

(2) Some aspects of the state may be *unforeseen* or *indescribable* by the parties in advance (perhaps because there is simply too vast a range of possibilities to think about);

(3) Even if certain aspects are foreseen, writing them into a contract may be too costly.

Let me put aside reason (3) right away. This is not to deny that it has validity, but only that we have not yet discovered a widely accepted principle for gauging the cost of specifying contingencies (see, however, the interesting work by Anderlini and Felli (1994) and MacLeod (1995)). This issue seems intimately related to that of "bounded rationality", a topic that I touch on in the conclusion.

As for reason (2), I will argue that, for a broad range of models used in the incomplete contracts literature, unforeseeability or indescribability "does not matter". More specifically, I will illustrate why the following theorem (sometimes called the Irrelevance Theorem and stated here rather loosely) holds:

Theorem 1. *If parties can assign a probability distribution to their possible future* payoffs, *then the fact that they cannot describe the possible* physical states (*e.g., the possible characteristics of the good to be traded*) *in advance is* irrelevant *to welfare. That is, the parties can devise a contract that leaves them no worse off than were they able to describe the physical states ex ante.*

I should stress that, this theorem does not imply that the parties can do as well as though they had fully contingent contracts, because reason (1) for incompleteness may

[2] This definition is so broad that it covers many contracts in the literature that are not normally considered "incomplete", e.g., insurance contracts with adverse selection. But for my purposes, I need not refine it further.

still pertain (see footnote 5 for an illustration of this point). I claim only that agents should not care that states are indescribable. [3]

3. An illustrative model

Rather than trying to make the above statement of the Irrelevance Theorem more precise at a general level, I will merely invoke an example that is typical of many models studied in the incomplete contracts literature. [4]

Suppose that agents 1 and 2 contemplate exchanging a single indivisible good that agent 1 will produce and agent 2 will consume. There are three dates.

At the first date – date 0 – the two agents negotiate the terms on which the good will be produced and exchanged. The outcome of the negotiation is a contract.

At date 1, agent 1 undertakes R&D that determines the set of characteristics of the good and hence the value v of the good to agent 2. At the same time, agent 2 invests in the development of an intermediate input that will facilitate the production process for agent 1 and therefore lower his production cost c. As one would expect, v is an increasing function $v(e_1)$ of agent 1's R&D expenditure e_1 and c is a decreasing function $c(e_2)$ of agent 2's investment e_2 (the more investment, the better the properties of the intermediate good, and so the lower the production cost).

Finally, at date 2, the characteristics of the good and the properties of the intermediate input are realized; agent 1 produces this good using that input and delivers it to agent 2; and agent 2 pays agent 1 as prescribed by the contract.

Agent 1's payoff can be expressed as

$$u_1(p - c(e_2) - e_1),$$

and agent 2's payoff is

$$u_2(v(e_1) - p - e_2),$$

where p is the price of the good and u_1 and u_2 are von Neumann–Morgenstern utility functions.

I shall assume, as is standard, that the investments, e_1 and e_2, and the private benefits and costs, v and c, cannot be verified by the contract enforcer, although they are commonly known by the two agents at date 2.

In this model, a *state of the world* corresponds to the characteristics of the good together with the properties of the intermediate input. Following the literature, I will assume that the state is verifiable by the contract enforcer ex post, i.e., at date 2. Let e_1^* and e_2^* denote the efficient investment levels for agents 1 and 2, respectively, assuming

[3] This is putting the claim too strongly. Maskin and Tirole (1999) state the hypotheses of the Irrelevance Theorem more carefully and show that, if they are violated, indescribability *can* matter. However, these hypotheses are nearly always satisfied by models in the literature.

[4] In some respects this example is closest to the model of Che and Hausch (1999), in which, as here, an agent may enhance the other party's payoff by his own investment. Che and Hausch, however, do not focus on the issue of indescribability.

that trade takes place; i.e., e_1^* and e_2^* are the levels that solve

$$\max_{e_1} v(e_1) - e_1$$

and

$$\min_{e_2} c(e_2) + e_2.$$

Let us assume that $v(e_1^*) - e_1^* > c(e_2^*) + e_2^*$, so that production and trade are indeed desirable. If the parties could foresee the state of nature corresponding to (e_1^*, e_2^*) they could simply write this into the contract. That is, in the contract, they could describe the properties of the intermediate input generated by e_2^* and specify that, should agent 2 fail to deliver on these properties, he must pay a penalty to agent 1. Similarly, if agent 1 failed to produce a good with the characteristics corresponding to e_1^*, he would be liable for a fine payable to agent 2. If these fines were sufficiently big, they would induce the agents to make the efficient investments (e_1^*, e_2^*), and so an optimal outcome would be induced by the contract. [5]

4. An optimal mechanism

It is the assumption that such a "complete" contract is unavailable that motivates the consideration of ownership rights in the incomplete contracts literature. I will argue, however, that even if a complete contract cannot be written – because of the impossibility of describing the characteristics and properties in advance – it should still be possible to reach the optimum through a suitably designed "mechanism" (at least, if parties are risk-averse). The idea is to exploit techniques from the implementation literature (the mechanism I exhibit is inspired by Moore and Repullo (1988); see Moore (1992), Palfrey (2001) and Maskin and Sjöström (2001) for surveys of the literature) to induce the agents to *reveal* the values of v and c and then to use this information in place of that about physical characteristics.

Here is a possible mechanism/contract [6] that the agents could sign at date 0 and execute at date 2:

Stage (i): Agent 1 announces $\hat{c}$ and agent 2 announces $\hat{v}$ (where the hats denote the possibility that the agents may not announce truthfully, i.e., we may have $\hat{c} \neq c(e_2)$ or $\hat{v} \neq v(e_1)$).

[5] One simplifying feature of this example is that the agents' investments give rise to a *deterministic* state of the world. If instead they generated a nondegenerate probability distribution over states then we would have to include the investments themselves as part of a state of the world. Because, however, investment is assumed to be unverifiable, agents would face a double moral hazard problem – and hence would ordinarily be limited to a second-best outcome, even if indescribability were not an issue.

[6] I do not wish to suggest that this contract – which I am proposing for pedagogical reasons – resembles mechanisms that are used in practice. However, more "realistic" institutions, such as auctions and options, often embody much the same kind of logic. Furthermore, to the extent that they do not replicate the performance of my mechanism, one must ask why the market for institutions has not stepped into the breach, an important unsolved question.

Stage (*ii*): Agent 1 can "challenge" agent 2's announcement. If the challenge is made,
(a) agent 2 must pay a fine f to agent 1, and then
(b) agent 1 offers agent 2 the choice between

$$(q^*, p^*) \text{ and } (q^{**}, p^{**}),$$

where

$$q^*, q^{**} \in \{0, 1\}^7$$

and

$$(*) \quad q^*\hat{v} - p^* > q^{**}\hat{v} - p^{**}.$$

Note that $(*)$ implies that agent 2 will choose (q^*, p^*) if he has been truthful (i.e., $\hat{v} = v(e_1)$).

The challenge "succeeds" if agent 2 chooses (q^{**}, p^{**}) (since agent 2 is then shown to have lied), in which case (q^{**}, p^{**}) is implemented. That is, agent 1 produces and delivers q^{**} units of the good (with characteristics corresponding to the realized state of the world, assumed to be verifiable[8]) for price p^{**}. In this case, the mechanism concludes at this point.

The challenge "fails" if agent 2 chooses (q^*, p^*) (since agent 2 is then shown to have told the truth), in which case (q^*, p^*) is implemented, i.e., agent 1 delivers q^* units of the good with characteristics corresponding to the realized state and receives price p^*. Furthermore, agent 1 must pay a fine of $2f$ for having challenged unsuccessfully. In this case, the mechanism concludes at this point.

If agent 1 does not make a challenge, then the mechanism moves to Stage (iii).

Stage (*iii*): Agent 2 can challenge agent 1's announcement. Such a challenge is handled completely symmetrically to that of Stage (ii). And if it occurs, the mechanism then concludes.

If neither agent makes a challenge, then the mechanism moves to Stage (iv).

Stage (*iv*): Agent 2 delivers the input with properties corresponding to the realized state. Agent 1 produces and delivers a unit of the good with characteristics corresponding to the realized state and receives price $p(\hat{v}, \hat{c})$, where

$$p(\hat{v}, \hat{c}) = \hat{v} + \hat{c} + k$$

and k is a constant. The mechanism then concludes.

[7] It may appear that I have "stacked the deck" in favor of an optimal mechanism's existing by supposing that parties can determine in advance that they will trade 0 or 1 unit of the good. But the very concept of a "unit" is ill-defined until the characteristics of the good have been determined.

[8] For simplicity, I am assuming that the parties need not specify before delivery the characteristics corresponding to the realized state; the court could verify for itself what characteristics are appropriate if agent 2 complained about the good he received. More generally, a mechanism could include an announcement/challenge scheme for characteristics similar to that for v and c (see Maskin and Tirole, 1999).

The first thing to notice about this mechanism is that, provided that $\hat{v}$ is untruthful (i.e., $\hat{v} \neq v(e_1)$), it is obvious that we can find (q^*, p^*) and (q^{**}, p^{**}) satisfying $(*)$ such that

$$(**) \quad q^* v(e_1) - p^* < q^{**} v(e_1) - p^{**}.$$

Hence, agent 1 can successfully challenge agent 2 if and only if 2 has been untruthful (from $(**)$, agent 2 would choose (q^{**}, p^{**})). Furthermore, if f is big enough, agent 1 has the incentive to challenge successfully, because he is then paid f by agent 2. Conversely, he will never make a challenge if agent 2 has been truthful; in that case, $(*)$ and $(**)$ cannot simultaneously be satisfied, and so agent 1 would expect any challenge to fail – meaning that, although he would still collect f from agent 2, he would have to pay a penalty of $2f$.

Hence, agent 2 will expect to be challenged and fined if and only if he announces untruthfully. He therefore has the incentive to set $\hat{v} = v(e_1)$. Similarly, agent 1 has the incentive to set $\hat{c} = c(e_2)$.

To show that agents will be willing to participate in our mechanism and that it induces an optimal allocation, we need only show that, for $i = 1, 2$, agent i wishes to set $e_i = e_i^*$ at date 1 and obtains a nonnegative payoff by doing so. But since agents will be truthful at date 2, agent 1's date 1 maximand is

$$p(v(e_1), c(e_2)) - c(e_2) - e_1,$$

which by definition of $p(v, c)$, equals

$$(***) \quad v(e_1) + k - e_1.$$

Notice that, regardless of e_2 and k, the maximizing choice of e_1 in $(***)$ is e_1^*. A symmetric argument applies to agent 2 and e_2^*. Furthermore, because $v(e_1^*) - e_1^* - c(e_2^*) - e_2^* > 0$, we can choose k so that both agents get positive payoffs in equilibrium, completing the argument.

This mechanism may appear to contradict the conclusions of Segal (1999) and Hart and Moore (1999). These authors present models in which mechanisms such as mine accomplish little,[9] rather than implement the optimum (as mine purports to do). The Segal/Hart–Moore arguments rely, however, on the premise that parties cannot commit themselves to refrain from renegotiating their contract if, at some point, it is to their mutual advantage to do so.

Maskin and Tirole (1999) contend that this premise is debatable, that there are ways in which parties who are determined to prevent renegotiation can succeed in doing so. However, I will not take issue with renegotiation here. Instead, I will show that, in addition, the Segal/Hart–Moore logic rests critically on the assumption that parties are risk-neutral.

To see the powerful role that renegotiation can play, consider the fine of $2f$ that agent 1 must pay in the mechanism above if his challenge is unsuccessful. Note that this fine cannot be paid to agent 2; otherwise, the latter would have the incentive

[9] In fact, the Segal (1999) and Hart and Moore (1999) models introduce sufficient "complexity" (in the form of a vast multiplicity of possible states) so that mechanisms accomplish literally nothing-parties are no better off with a mechanism than with no contract at all.

to make the challenge fail (i.e., to select (q^*, p^*)), even if he had not been truthful. Hence, the fine must be paid to a third party. But, in that case, the agents have the incentive to renegotiate just before the fine is paid. That is, rather than giving away money to an outsider, the agents are better off if they split the fine between themselves. Yet, if agent 2 gets a significant portion of the split, he may again have the incentive to see that a valid challenge fails.

This difficulty with the fine illustrates a general problem created by the possibility of renegotiation, viz., that it may be hard to punish one party for "misbehaving" (e.g., invalidly challenging) without simultaneously rewarding the other party (and thereby distorting the latter's incentives).

This is where risk-aversion can help. Suppose that agent 1 is risk-averse and (to keep matters simple) agent 2 is risk-neutral. Rather than having agent 1 pay a fine $2f$ if his challenge fails, have him pay

$$g, \quad \text{with probability } \tfrac{1}{2}$$

and

$$-g, \quad \text{with probability } \tfrac{1}{2},$$

where agent 2 receives the fine. By making g big enough, then in view of agent 1's risk-aversion, we can make this stochastic fine as harsh as we like (in particular, we can make it as bad as a deterministic fine of $2f$). Notice, however, that the fine does not constitute a reward to agent 2 since its mean is zero.

If the realization of the random fine is determined as soon as a challenge fails, [10] then renegotiation will not be possible after agent 2 has made his choice. What, though, about renegotiation *beforehand*?

If agent 1 has *invalidly* challenged, then the parties will anticipate that the challenge should fail, and so indeed *will* want to renegotiate the random fine beforehand. Nevertheless, provided that agent 2's share of the surplus from renegotiation is bounded away from zero, agent 1's payoff even after renegotiation can still be driven as low as we like by making g big enough. Thus for a suitably big value of g, agent 1 will be deterred from making invalid challenges.

On the other hand, if agent 1 has made a valid challenge (i.e., a challenge satisfying $(*)$ and $(**)$ above [11]), then, provided that the random fine is *not* renegotiated beforehand, agent 2 has no incentive to make the challenge fail (since his expected payment would then be zero). Indeed, if the challenge is expected to succeed, there is no value to renegotiating the random fine, because, this fine will not be expected to arise anyway. Moreover, provided that g is big enough, agent 1 would *veto* any proposal for renegotiating the fine (presumably, both agents must consent for renegotiation to take

[10] This could be arranged by having agent 2 report his choice between (q^*, p^*) and (q^{**}, p^{**}) by depressing the "f" or "s" key, respectively, on a computer keyboard. The computer would be set up so that depression of the "f" key (for "failed challenge") would instantly generate a realization of the randomization between g and $-g$.

[11] Once renegotiation is possible, then (q^*, p^*) or (q^{**}, p^{**}) might itself be renegotiated if it is not already efficient. But it is not hard to verify that, as long as $\hat{v} \neq v(e_1)$ and $v(e_1) > c(e_2)$, agent 1 can still devise a successful challenge.

place), as eliminating the randomness might give agent 2 the incentive to make the valid challenge fail, in which case, despite the renegotiation, agent 1 would be worse off than if it succeeded (as in the preceding paragraph, g can be chosen big enough to ensure this). Thus, the possibility of renegotiation beforehand does not interfere with *valid* challenges either.

I have argued that a well-chosen random fine will deter agent 1, if risk-averse, from challenging invalidly and will not prevent his valid challenges from succeeding. Symmetrically, the same is true for agent 2. Thus, if both agents are risk-averse, my mechanism above, amended to incorporate the randomness, will attain the optimum even in the face of renegotiation.

5. Conclusion

I have suggested that attributing the incompleteness of contracts to the indescribability of contingencies is theoretically problematic. But as I stated at the outset, this foundational difficulty does not imply that we should ignore the valuable contributions that the incomplete contracts literature has made. In my view, we need not wait for completely rigorous foundations to explore the implications of incompleteness, as long as we recognize the potential tentativeness of the conclusions.

At the same time, I am certainly in favor of more work directed at foundations, difficult though such an enterprise may be. One can certainly argue that I have relied exceedingly heavily on agents' abilities to foresee future payoffs in my treatment of the optimal mechanism above. It may well be too much to expect that agents can make these forecasts in reality. This suggests that "bounded rationality" could be a potentially fruitful explanation of incompleteness. Unfortunately, a useful model in which agents' forecasting abilities are plausibly and realistically circumscribed seems yet to be developed.

Acknowledgements

I would like to thank M. Dewatripont, J. Tirole, and J. Weibull for their reactions to the oral presentation and M. Dewatripont, B. MacLeod and J. Tirole for helpful comments on an earlier draft. I gratefully acknowledge research support from the NSF.

References

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

Che, Y.-K., Hausch, D., 1999. Cooperative investments and the value of contracting. American Economic Review 89, 125–147.

Hart, O., 1995. Firms, Contracts and Financial Structure. Oxford University Press, Oxford.

Hart, O., Moore, J., 1999. Foundations of incomplete contracts. Review of Economic Studies 66, 115–138.

MacLeod, B., 1995. Incentives in organizations: An overview of some of the evidence and theory. In: Siebert, H. (Ed.), Trends in Business Organization, The Kiel Institute of World Economics.

Maskin, E., Moore, J., 1999. Implementation and renegotiation. Review of Economic Studies 66, 39–56.
Maskin, E., Sjöström, T., 2001. Implementation theory. In: Arrow, K., Sen, A., Suzumura, K. (Eds.), Handbook of Social Choice and Welfare, North-Holland, Amsterdam, to appear.
Maskin, E., Tirole, J., 1999. Unforeseen contingencies and incomplete contracts. Review of Economic Studies 66, 83–114.
Moore, J., Repullo, R., 1988. Subgame perfect implementation. Econometrica 56, 1191–1220.
Moore, J., 1992. Implementation, contracts and renegotiation in environments with complete information. In: Laffont, J.-J. (Ed.), Advances in Economic Theory, Vol. 1. Cambridge University Press, Cambridge, pp. 182–282.
Palfrey, T., 2001. Implementation theory. In: Aumann, R., Hart, S. (Eds.), Handbook of Game Theory, Vol. 3. North-Holland, Amsterdam.
Segal, I., 1999. Complexity and renegotiation: A foundation for incomplete contracts. Review of Economic Studies 66, 57–82.
Tirole, J., 1999. Incomplete contracts: Where do we stand? Econometrica 67, 741–781.