1 Introduction

It is generally accepted by both economists and lawyers that almost all contracts are incomplete. It is simply too costly for parties to anticipate the many contingencies that may occur and to write down unambiguously how to deal with them. Contractual incompleteness has been shown to throw light on a number of matters of interest to economists, such as the boundaries of the firm, asset ownership, and the allocation of control and authority.

Yet the million dollar question remains: why are contracts as incomplete as they are? The idea that transaction costs or bounded rationality are a total explanation for this is not convincing. In many situations some states of the world or outcomes are verifiable and easy to describe, appear relevant, and yet are not mentioned in a contract. A leading example is a breach penalty. A contract will usually specify the price the buyer should pay the seller if trade occurs as intended, but may not say what happens if there is a breach or under what conditions breach is justified. Of course, sophisticated parties often do include breach penalties in the form of liquidated damages but this is far from universal.

A second example concerns indexation. Since a worker's marginal product varies with conditions in the industry she works in as well as the economy as a whole, we might expect to see wages being indexed on variables correlated with industry profitability such as share prices or industry or aggregate unemployment, as well as to inflation. Such an arrangement might have large benefits, allowing wages to adjust and avoiding inefficient layoffs and quits of workers (see, e.g., Weitzman [1984] and Oyer [2004]). Yet, the practice does not seem a common one overall.[1] In the 2008 financial crisis many debt contracts were not indexed to the aggregate state of the economy; if they had been the parties might have been able to avoid default, which might have had large benefits both for them and for the economy as a whole. Similarly, in 2020, few contracts had clauses describing what should happen in the event of a pandemic.

How do we explain the omission of contingencies like these from a contract? One possibility is to argue that putting any contingency into a contract is costly – some of these costs may have to do with describing the relevant state of the world in an unambiguous way – and so if a state is unlikely it may not be worth including it (see, e.g., Shavell [1980], Dye [1985]). This is often the position taken in the law and economics literature (see, e.g., Posner [1986], p.82). However, this view is not entirely convincing. First, states of the world such as breach are often not that unlikely and not that difficult to describe.[2] Second, while the financial crisis or a pandemic may have been unlikely ex ante, now

[1] However, see Card (1986) on wage indexation in union contracts in North America.
[2] As argued by Ayres and Gertner (1989), p.128, fn177.

that they have happened the possibility of future crises or pandemics seem only too real. Moreover, finding verifiable ways to describe a crisis or pandemic does not seem to be beyond the capability of contracting parties. Thus, one might expect parties to rush to index contracts on such events. We are not aware of any evidence that this is happening.

A second possibility is to appeal to asymmetric information (see, e.g., Spier [1992]).[3] The idea is that suggesting a contingency for inclusion in a contract may signal some private information and this may have negative repercussions. Such an explanation does not seem very plausible in the case of financial crises – where is the asymmetry of information about the prospects of a global crisis? – but it may apply in other cases. For example, if I suggest a (low) breach penalty you may deduce that breach is likely and this may make you less willing to trade with me. Or if you suggest that my wage should fall if an industry index of costs rises I may think that you are an expert economist who already knows that the index is likely to rise.

Even in these cases asymmetric information does not seem to be a complete answer. Asymmetric information generally implies some distortion in a contract but not that a provision will be completely missing. For example, in the well-known Rothschild-Stiglitz (1976) model, insurance companies offer low-risk types less than full insurance to separate them from high-risk types. But the low-risk types are not shut out of the market altogether – they still obtain some insurance (and the high-risk types receive full insurance). Indeed to explain why a contingency might be omitted from a contract, Spier assumes a fixed cost of writing or enforcing contractual clauses in addition to asymmetric information.[4]

In this Element, we offer an alternative and complementary explanation for why verifiable contingencies are omitted based on the theory of contracts as reference points (see Hart and Moore [2008]).[5] In a nutshell this approach takes the view that a contract circumscribes what parties feel entitled to. Parties do not feel entitled to outcomes outside the contract but may feel entitled to different outcomes within the contract. If a party does not receive what he feels entitled to he is aggrieved and shades on performance, creating deadweight losses.

Hart and Moore (2008) suppose that each party feels entitled to the best outcome permitted by the contract and rule out renegotiation. In this Element

[3] For related work, see Aghion and Bolton (1987), Ayres and Gertner (1989,1992), and Aghion and Hermalin (1990).

[4] However, see Hartman-Glaser and Hebert (2020) for a model of missing provisions that does not depend on a writing cost.

[5] There are no doubt other reasons why contingencies are left out of contracts. Parties may find it distasteful to talk about bad outcomes, such as breach or default, or mentioning them may suggest or breed a lack of trust. These explanations tend to involve psychological factors; our Element can be seen as one attempt to model such factors.

we relax both these assumptions. We confine attention to initial contracts that specify a single (possibly contingent) trading outcome ex post (so there is no aggrievement or shading with respect to the initial contract). Renegotiation occurs ex post if the trading outcome is inefficient in the contingency that arises. We assume that as a result of a self-serving bias each party feels entitled to more than half of the surplus from renegotiation, causing aggrievement and shading. In addition, there may be disagreement about the reference point for the evaluation of surplus, increasing aggrievement further. We show that adding a verifiable contingency to the contract may increase disagreement about the appropriate reference point in contingencies not covered by the contract. An incomplete contract would then reduce the deadweight losses from renegotiation.

In our model, a buyer wants a particular good or service most of the time but with some probability may require an "add-on". Some states of the world in which the add-on is required are verifiable, but others are not. The question we ask is whether it is better to specify that the add-on should be supplied in the verifiable states or whether it is better to specify the basic good and rely on renegotiation in the event that a change is needed.

Suppose that a contingency not covered by the contract occurs. One party may choose what would have occurred in one verifiable contingency to be the reference point for renegotiation whereas the other party may choose what would have occurred in another verifiable contingency. Thus having contractual outcomes in several contingencies can complicate renegotiation in contingencies not covered by the contract. This is particularly an issue if the parties have relatively similar views about a reasonable division of surplus. Then renegotiation would proceed smoothly with an incomplete contract, while additional reference points drawn from a more complete contract may hinder renegotiation. We will show that the renegotiation-hindering effect can be dominant – and an incomplete contract can be superior – even when the parties have different views about surplus division. This is the case when the reference points are drawn from very divergent additional contingencies.

The problem arises here because there are multiple reference points and the parties may disagree about which is the right one. In our main model we will assume that each party chooses the reference point most favorable to him or her, but we do not need to go this far. Similar (although weaker) results can be obtained even if each party randomized over the reference points.[6]

[6] The idea that a contractual provision for another state can affect entitlements in the current state is related to the notion of external reference points in Section V of the Hart–Moore model. In Hart–Moore, comparable transactions – that can be justified as reasonable to outsiders – can influence entitlements in a particular state but Hart–Moore do not analyze the case where contractual provisions in one state affect entitlements in another.