ISSN 1936-5349 (print)
ISSN 1936-5357 (online)

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

MORE IS LESS: WHY PARTIES MAY DELIBERATELY WRITE INCOMPLETE CONTRACTS

Maija Halonen-Akatwijuka
Oliver Hart

Discussion Paper No. 748

05/2013

Harvard Law School
Cambridge, MA 02138

More is Less: Why Parties May Deliberately Write Incomplete Contracts

Maija Halonen-Akatwijuka and Oliver Hart

April 2013

Abstract

Why are contracts incomplete? Transaction costs and bounded rationality cannot be a total explanation since states of the world are often describable, foreseeable, and yet are not mentioned in a contract. Asymmetric information theories also have limitations. We offer an explanation based on "contracts as reference points". Including a contingency of the form, "The buyer will require a good in event E", has a benefit and a cost. The benefit is that if E occurs there is less to argue about; the cost is that the additional reference point provided by the outcome in E can hinder (re)negotiation in states outside E. We show that if parties agree about a reasonable division of surplus, an incomplete contract can be strictly superior to a contingent contract.

Halonen-Akatwijuka : Department of Economics, Bristol University (email: Maija.Halonen@bristol.ac.uk); Hart: Department of Economics, Harvard University (email: ohart@harvard.edu). We would like to thank Philippe Aghion, Ola Kvaløy, Evagelos Pafilis, Klaus Schmidt, and particularly Kathy Spier for very helpful comments. We gratefully acknowledge financial support from the U.S. National Science Foundation through the National Bureau of Economic Research and the first author from the Norwegian Research Council.

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)). Indeed Oyer (2004) argues that high tech firms grant stock options to employees to avoid quits. Yet the practice does not seem a common one overall[1]. Similarly, in the recent 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.

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., Dye (1985), Shavell (1980)). 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 recent financial crisis may have been unlikely ex ante, now that it has happened the possibility of future crises seems only too real. Moreover, finding verifiable ways to describe a crisis does not seem to be beyond the capability of contracting parties. Thus one might

[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).

expect parties to rush to index contracts on future crises. 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.

In this paper we offer an alternative and complementary explanation for why verifiable contingencies are omitted based on recent theoretical work on contracts as reference points (see Hart and Moore (2008))[4]. 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 paper 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 one party—the seller—feels entitled to a fraction α of the surplus from renegotiation and the other party—the buyer—feels entitled to a fraction $(1 - \beta)$ of the surplus, where $\alpha \geq ½ \geq \beta$. However, there may be disagreement about the reference point for the evaluation of surplus. 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

[3] For related work, see Aghion and Bolton (1987), Ayres and Gertner (1989,1992) and Aghion and Hermalin (1990).
[4] 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 paper can be seen as one attempt to model such factors.

another verifiable contingency. Thus having contractual outcomes in several contingencies can complicate the renegotiation process in contingencies not covered by the contract.

The problem arises here because there are multiple reference points and the parties may disagree about which is the right one. In the model below 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 could be obtained even if each party randomized over the reference points.

Our approach seems consistent with lawyers' views about contract interpretation. In a recent paper Schwartz and Scott (2010) argue that judicial interpretation should be made on a limited evidentiary basis, the most important element of which is the contract itself. Although Schwartz and Scott do not consider the issue of contingent clauses it seems inevitable that a court that focuses on a contract will find a clause governing one contingency relevant for adjudicating another contingency. If the parties do not want this to happen it may be better to leave the contingency out. This is similar to our idea that contracting parties may want to leave a contingency out to reduce argument among themselves.

The following somewhat "homey" example illustrates our approach[5]. Suppose that you hire a Nanny to work Monday-Friday from 9am-5pm for $600 per week ($15 per hour). There is a chance that you will get stuck in traffic and will be late. Should you include a late fee of, say, $30 per hour in the Nanny's contract? (Being late is a verifiable contingency.) Including the late fee could prevent bad feelings later on about how much the Nanny should be paid when you are late. But if you include the late fee, it may create some expectation by the Nanny concerning what she should receive if, say, you need her to work on the weekend. (There may be several reasons for you to want her to work on the weekend—some business, some pleasure— and it may be difficult to distinguish between these in advance.) She might feel that $30 per hour is the appropriate reference point for such an arrangement, whereas you might feel that $15 per hour is. If you and the Nanny have similar views about what is reasonable absent a reference point, it may be better to leave the late fee out and renegotiate as needed.

Our analysis operates a little differently from this example since we suppose that parties have views about a reasonable division of surplus rather than a reasonable price. Given this we will show that it will be desirable to exclude a verifiable contingency from a contract only if there is some ambiguity about the contingency, in the sense that it refers to several states of the world rather than just one.

We apply our framework to analyze (the relatively infrequent use of) wage indexation. Consider an employment contract that is "at-will" – either party can quit the relationship. For trade to occur the wage must lie between the worker's productivity and her opportunity cost. Suppose that a signal is informative about the worker's value and her opportunity cost but not completely so: other factors also influence these variables. Then indexing wages on the signal may have the disadvantage that this creates a reference point that may make it harder for the parties to agree about what is reasonable in a state of the world where the signal is a poor guide to productivity and opportunity cost. Again, it may

[5] We are grateful to Kathy Spier for suggesting this example.

be better for the parties not to index on the signal at all[6]. This result may shed light on why wage indexation, although observed in some situations (see Oyer (2004)), is not more common.

Developing a general model that pins down these ideas is not easy, and so instead we will present two separate models, each of which is more in the nature of an extended example. In the first model a buyer wants a particular good or service most of the time but with some probability may require an "add-on" or "extra". 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. (So this model is in the spirit of the Nanny example.) In the second model we consider an at-will employment relationship where a verifiable signal is available that provides information about the worker's productivity and opportunity cost, and ask whether the wage should be indexed on this signal.

It should be noted that the Hart-Moore (2008) model, as it stands, cannot explain why easy-to-contract-on contingencies are left out of contracts. The reason is that in Hart-Moore it is supposed that each party feels entitled to the best outcome in each state. Hence a provision in one state has no effect on what a party feels entitled to in another state[7]. Thus it is important as a first step to generalize the Hart-Moore model: we do this by supposing that $1 > \alpha \geq ½ \geq \beta > 0$.

Our paper is related to a number of contributions in the literature. Bernheim and Whinston (1998) show that it can be optimal not to contract on some verifiable aspects of performance to improve unverifiable performance. For example, a buyer and a seller may contract on price but leave quantity unspecified. This partial incompleteness can give the seller an incentive to provide good (unverifiable) quality given that the buyer's demand is increasing in quality. Bernheim and Whinston (1998) focus on verifiable and unverifiable actions and show how discretionary action can discipline unverifiable action. Our model focuses on states rather than actions and shows how an additional contingency can lead to more divergent entitlements and greater shading in some unverifiable states.

The literature on the interaction of explicit and relational contracts is also related (see, e.g., Baker et al. (1994) and Schmidt and Schnitzer (1995)). In this literature an explicit contract determines the default position after reneging and can undermine the relational contract governing the relationship if the default position is too attractive. In our approach additional contingencies may hinder renegotiation in an unverifiable state. Kvaløy and Olsen (2009) allow for the parties to improve verifiability of the explicit contract by investing in contract design and show how an inferior explicit breach remedy can strengthen the relational contract by limiting the default position.

---

[6] This is in contrast to a result in Hart (2009), which shows that indexation is beneficial. However, Hart (2009) supposes that (as in Hart-Moore (2008)) each party feels entitled to the best outcome permitted by the contract in each state, whereas in the current paper we relax this assumption. As we will note shortly this relaxation is crucial for our results.

[7] Section V of Hart-Moore considers the case where external contracts or prices can influence entitlements in a particular state but not where contractual provisions in one state affect entitlements in another.

Bénabou and Tirole (2003, 2006) and Herold (2010) find that a principal may choose to rely completely on intrinsic motivation if explicit incentives would backfire by signaling some adverse information, e.g., about the principal's view of the agent's ability, true motivation for good deeds or distrust. Our interest is in explaining whether or not an additional contingency should be added to a contract.

Bounded rationality can also lead to incomplete contracts. In Tirole (2009) agents are aware of their cognitive limitations, in the sense that they know that they may not be aware of the best design for the traded good. The agents can invest in finding out about alternative designs. If agents invest little, contracts are incomplete and there is a high probability that the contract has to be renegotiated. However, contracts may also be too complete if too many resources are spent on search to avoid a vulnerable position in renegotiation. In Bolton and Faure-Grimaud (2010) the agents may postpone thinking about unlikely states until later and instead assign control rights, particularly if the agents have aligned interests. Al-Najjar et al. (2006) derive incomplete contracts from undescribable events.

Kukharskyy (2013), in applying contracts as reference points to global sourcing, posits that in an egalitarian country a manager's entitlement is related to his contribution to the relationship (his parameter in a Cobb-Douglas production function). Therefore contractual flexibility is higher in egalitarian countries compared to selfish countries where managers feel entitled to the best outcome within the contract. This makes egalitarian countries more attractive for foreign direct investment. In our model the results depend importantly on the similarity of the views the parties have about the division of surplus as both parties have an opportunity to shade.

Finally, our approach is quite closely related to Herweg and Schmidt (2012). Their work also depends on the idea that a contract can provide a reference point that may hinder renegotiation. However, they rely on loss aversion rather than aggrievement, and they do not focus specifically on the absence of contingencies in a contract. They also examine price indexation and find that although indexing price on a verifiable signal on its own is never optimal, price indexation combined with the buyer having the right to choose the quantity ex post can increase surplus.

The paper is organized as follows. We present the two models in Sections 2 and 3. Section 4 contains a discussion of the results and some conclusions.

2. A model of variable requirements

Throughout the paper we consider a buyer B and a seller S who meet at date 0 and can trade at date 1. We assume a perfectly competitive market at date 0 but that, possibly because of (unmodelled) relationship-specific investments, B and S face bilateral monopoly at date 1. There is symmetric information throughout. B and S are risk neutral and do not face wealth constraints, and there is no discounting.

In this section we suppose that B and S always want to trade a basic widget, but in some states they want an additional component – an "add-on". Both the basic widget and the augmented widget (the basic widget plus the add-on) are ex ante contractible and specific performance is possible (in contrast to Hart and Moore (2008)). What this means is that B and S can at date 0 write contracts of the form, "We will trade the basic widget" or "We will trade the augmented widget", and these will be enforced at date 1: either party can be assessed a large penalty for failing to comply.

We shall suppose that it is always efficient for the parties to trade the basic widget, but that it is only sometimes efficient to have the add-on. Specifically, there are four states of the world. (Three states are not enough for the model to work as will become clear below.) In s1, which is the most likely state, only the basic widget is needed: the cost exceeds the benefit of the add-on. In s2 – s4 the add-on is efficient. We shall suppose that none of the individual states is verifiable, but the event "s3 or s4" is. Call this event E. As we will see below it is important that there is some ambiguity about the verifiable contingency. The payoffs are illustrated in Figure 1.

State	Probability	Payoff from add-on
s1	π_1	(v_1,c_1)
s2	π_2	(v_2,c_2)
s3	π_3	(v_3,c_3)
s4	π_4	(v_4,c_4)

Figure 1

In Figure 1 B's payoff appears first: v_i is the value of the add-on and c_i its cost in state i.

We assume that $\pi_i > 0$ for all i.

Let $G_i \equiv v_i - c_i$. As mentioned, we assume

(2.1) $G_1 < 0$, $G_i > 0$, i = 2,3,4.

We also suppose that states 3 and 4 can be ranked: one state (without loss of generality, state 4) has a higher value and a higher cost:

(2.2) $v_4 > v_3$, $c_4 > c_3$.

The role of (2.2) will become clear later.

There are three leading contracts:

"Always trade the basic widget";

"Trade the basic widget except in event E where the add-on is included at an extra charge p";

"Always trade the augmented widget".

Call these contracts 1, 2, and 3, respectively[8].

In contrast to Hart and Moore (2008) we allow for renegotiation once the parties learn the state.

Since s1 is likely and the add-on is inefficient in this state, contract 3 is unlikely to be optimal. Hence we will focus on contracts 1 and 2. Our particular interest is whether the more complete contract 2 (in the sense that it includes more contingencies) is superior to the less complete contract 1.

Let us turn now to the issue of entitlements, aggrievement, and shading.

Following Hart and Moore (2008), we suppose that the initial contract is regarded as "fair" since it is negotiated under competitive conditions. However, parties may disagree about what is reasonable within the contract or if it is renegotiated. A party who does not receive what he is entitled to is aggrieved and shades: he performs within the letter rather than the spirit of the contract in a way that hurts the other party. (Shading is noncontractible.) To be more precise, suppose that a party's payoff is y and he feels entitled to x. Then his aggrievement is $(x – y)$ and he shades to the point where the other party's payoff is reduced by $\theta(x – y)$, where $0 < \theta < 1$ is exogenous. Both B and S can shade. Shading does not affect the payoff of the person doing the shading: it simply reduces the payoff of the other party.

In this paper we will consider only contracts that specify a single trading outcome in each state. Thus there is no aggrievement with respect to the initial contract. However, in contrast to Hart and Moore (2008), we allow for renegotiation. We also generalize Hart and Moore (2008) by supposing that a party does not necessarily feel entitled to the best possible outcome if the contract is renegotiated. In particular, suppose that there are gains from renegotiation equal to G. We assume that S feels entitled

[8] Another leading contract used in practice is a cost plus contract: the buyer can require the seller to supply the add-on but must pay the seller's incremental cost (possibly marked up); for discussions see Bajari and Tadelis (2001) and Chakravarty and MacLeod (2009). In this paper we assume that the seller's incremental cost is not verifiable; this assumption may be plausible in the Nanny example. We also do not consider contracts that grant the buyer the option to buy, or the seller the option to sell, the add-on at a pre-specified price. Such contracts may be useful in some situations but they have their own costs. For example, consider a contract that specifies that the add-on will be provided in event E at price p; and sets a price p' at which the add-on can be traded as long as both parties agree if E does not occur, where $v_2 > p' > c_2$. Such a contract ensures trade of the add-on in s2 where it is efficient but not in s1 where it is inefficient (one party will refuse to trade since $v_1 < c_1$).The problem with this contract is that, if $v_1 > v_2$ or $c_2 > c_1$, the buyer or the seller will be aggrieved when the add-on is not traded in s1, and will shade with respect to the basic widget (which is traded), creating deadweight losses.

to a fraction α of the gains, and B feels entitled to a fraction $(1 - \beta)$ (to put it another way, B feels that S is entitled to a fraction β of the gains) , where

(2.3) $$1 > \alpha \geq ½ \geq \beta > 0.$$

(Hart and Moore (2008) can be regarded as the limiting case where $\alpha = 1$, $\beta = 0$.) We suppose that the parties have equal bargaining power, and so they compromise on a 50:50 split, but each is aggrieved. S's aggrievement is $(\alpha - ½)G$ and B's aggrievement is $(½ - \beta)G$. Thus S will shade to the point where B's payoff falls by $\theta(\alpha - ½)G$ and B will shade to the point where S's payoff falls by $\theta(½ - \beta)G$. Total deadweight losses from shading equal

(2.4) $$\theta(\alpha - \beta)G.$$[9]

We can now analyze the deadweight losses under contracts 1 and 2.

Contract 1: Always trade the basic widget at some agreed-on price

In s1 the outcome mandated by the contract is efficient, and so there will be no renegotiation. Since the contract specifies a single outcome, there is nothing to be aggrieved about: each party gets what he feels entitled to, and so there are no deadweight losses from shading.

Consider next state i, i > 1. Here there are gains from renegotiation given by G_i. From (2.4) the deadweight losses from renegotiation in state i are $\theta(\alpha - \beta)G_i$. Thus we can write the expected deadweight losses from contract 1 as

(2.5) $$L_1 = \sum_{i=2}^{4} \pi_i\theta(\alpha - \beta)G_i.$$

Contract 2: Trade the basic widget except in event E where the add-on is included at an extra charge p

Contract 2 introduces additional reference points, and as we shall see this may cause problems.

Under contract 2 there is no aggrievement in s1, s3 or s4 since the contract mandates a single outcome that is efficient.[10] Thus the only problem state is s2. In s2 the contract mandates the basic widget, but the augmented widget is efficient. Renegotiation will occur but now there are three reference points. "Trade the basic widget" is one (as above), but so are "trade the augmented widget at extra charge p in

[9] Note that we assume that renegotiation does not cause parties to reassess the fairness of the initial contract for the basic widget. For some experimental evidence consistent with this, see Fehr et al. (2012).

[10] Note that one party may make losses in s3 or s4. Yet there is no aggrievement.

s3" and "trade the augmented widget at extra charge p in s4". We will take the position that each party is self-serving in his or her choice of reference point. In particular, S can justify any of the following prices as reasonable for the add-on in s2:

(2.6) $p_S' = c_2 + \alpha G_2$,

(2.7) $p_S'' = c_2 + p - c_3 + \alpha(G_2 - G_3)$,

(2.8) $p_S''' = c_2 + p - c_4 + \alpha(G_2 - G_4)$.

The first price p_S' is based on the idea that S feels entitled to a fraction α of the gains from renegotiation taking "trade of the basic widget" as the status quo. The second price p_S'' is based on the idea that S regards "trade of the add-on at the extra price p in state s3" as the status quo. The argument S can make is that such an outcome is reasonable since the contract admits it in s3. Relative to the s3 outcome S's cost is c_2 rather than c_3 and the gains from trade are G_2 rather than G_3. Since S feels entitled to a fraction α of the change in surplus, this yields (2.7). (2.8) follows from the same calculation for the case where S uses "trade of the add-on at the extra price p in state s4" as the status quo.

Note that because event E consists of two states S can use either of them as a reference point. This will be important for what follows.

As noted we assume that S adopts the most favorable interpretation for her, i.e., she feels entitled to a price equal to

(2.9) $\text{Max}(p_S', p_S'', p_S''')$.

However, following Hart and Moore (2008) we suppose that S recognizes that B will never pay more than v_2 for the add-on in s2. Hence S's entitlement is capped by v_2.

We can simplify things a little. It is easy to see that (2.2) implies that

(2.10) $p_S'' > p_S'''$.

Hence we can write S's entitlement as

(2.11) $\text{Min}(v_2, \text{Max}(p_S', p_S''))$.

By a similar logic B feels entitled to pay

(2.12) $\text{Min}(p_B', p_B'', p_B''')$,

where

(2.13) $p_B' = c_2 + \beta G_2$,

(2.14) $p_B'' = c_2 + p - c_3 + \beta(G_2 - G_3)$, (2.15) $p_B''' = c_2 + p - c_4 + \beta(G_2 - G_4)$.

Again B's entitlement is bounded below by c_2 – he realizes that S will never supply the add-on for less than this. Moreover, we can again use (2.2) to show that $p_B'' > p_B'''$. Hence we can write B's entitlement as

(2.16) $\mathrm{Max}(c_2, \mathrm{Min}(p_B', p_B'''))$.

We may conclude that the expected deadweight losses from contract 2 – incurred entirely in s2 – are

(2.17) $L_2 = \pi_2\theta[\mathrm{Min}(v_2, \mathrm{Max}(p_S', p_S'')) - \mathrm{Max}(c_2, \mathrm{Min}(p_B', p_B'''))]$.

Given $\alpha < 1$, $\beta > 0$, $v_2 > p_S'$ and $p_B' > c_2$ and so

(2.18) $\mathrm{Min}(v_2, \mathrm{Max}(p_S', p_S'')) - \mathrm{Max}(c_2, \mathrm{Min}(p_B', p_B''')) \geq p_S' - p_B' = (\alpha - \beta)G_2$.

That is, deadweight losses in s2 under contract 2 are at least as great as under contract 1.

It is optimal to set p to minimize L_2. Lemma 1 provides necessary and sufficient conditions for (2.18) to hold with strict inequality at the optimum.

<u>Lemma 1.</u> $L_2 > \pi_2\theta(\alpha - \beta)G_2$ if and only if $\alpha v_3 + (1 - \alpha)c_3 - \beta v_4 - (1 - \beta)c_4 < 0$.

<u>Proof:</u> Given $v_2 > p_S'$ and $p_B' > c_2$, (2.18) must hold with strict inequality if either $p_S' < p_S''$ or $p_B' > p_B'''$. Suppose the contrary: $p_S' \geq p_S''$ and $p_B' \leq p_B'''$. Then

(2.19) $p \leq c_3 + \alpha G_3$, $p \geq c_4 + \beta G_4$.

It is possible to find p such that (2.19) holds if and only if

(2.20) $\alpha G_3 - \beta G_4 + c_3 - c_4 = \alpha v_3 + (1 - \alpha)c_3 - \beta v_4 - (1 - \beta)c_4 \geq 0$.

In this case p can be chosen so that $L_2 = \pi_2\theta(\alpha - \beta)G_2$.

However, it is not possible to find p such that (2.19) holds if and only if

(2.21) $\alpha v_3 + (1 - \alpha)c_3 - \beta v_4 - (1 - \beta)c_4 < 0$.

Then $L_2 > \pi_2\theta(\alpha - \beta)G_2$.

Q.E.D.

It may be possible to make the additional reference points provided by "trade the add-on in E" redundant by choosing p appropriately. This is the case when $(p_S' - p_B') \geq (p_S'' - p_B''')$ as p increases p_S''

and p_B''' by equal amount. However, when $(p_S' - p_B') < (p_S'' - p_B''')$ – or $\alpha v_3 + (1 - \alpha)c_3 - \beta v_4 - (1 - \beta)c_4 < 0$ – at least one party will use the additional reference point irrespective of the value of p, increasing aggrievement. This condition holds if v_4, c_4 are much bigger than v_3 , c_3 since the additional reference points based on different states are very divergent. Also importantly, the condition holds if α, β are close as then $p_S' \approx p_B'$, but the additional reference points remain divergent as they are based on different states. Lemma 1 tells us that in this case the multiple reference points hinder the renegotiation process.

Proposition 1 follows immediately from Lemma 1.

<u>Proposition 1</u>

(1) If $(\alpha - \beta)$ is sufficiently small, contract 1 is strictly superior to contract 2.

(2) If $(\pi_3 + \pi_4)$ is sufficiently small and $\alpha v_3 + (1 - \alpha)c_3 - \beta v_4 - (1 - \beta)c_4 < 0$, contract 1 is strictly superior to contract 2.

(3) If $(\alpha - \beta)$ is sufficiently large, contract 2 is strictly superior to contract 1.

<u>Proof</u>

(1) Consider the case $\alpha = \beta = ½$. Obviously the deadweight losses from contract 1 are zero. Consider contract 2. From (2.2), $\alpha v_3 + (1 - \alpha)c_3 - \beta v_4 - (1 - \beta)c_4 = ½(v_3 + c_3 - v_4 - c_4) < 0$ and so by Lemma 1 $L_2 > 0$. Hence the deadweight losses from contract 2 are strictly positive. Therefore Proposition 1(1) is true for $\alpha = \beta = ½$. By continuity it must also be true for α, β close to ½.

(2) As $(\pi_3 + \pi_4) \to 0$, the deadweight losses under contract 1 converge to $\theta(\alpha - \beta)G_2$. By Lemma 1 the deadweight losses under contract 2 converge to L_2 which is strictly above this, given $\alpha G_3 - \beta G_4 + c_3 - c_4 < 0$. Hence contract 1 is strictly superior to 2 in the limit, and, by continuity for $(\pi_3 + \pi_4)$ sufficiently close to zero.

(3) When α is close to 1 and β is close to zero, the deadweight losses from contract 1 are approximately $L_1 = \sum_{i=2}^{4} \pi_i \theta G_i$. From (2.19) the deadweight losses from contract 2 are approximately $L_2 = \pi_2\theta(v_2 - c_2) = \pi_2\theta G_2$, and so $L_2 < L_1$ given that $\pi_3,\pi_4 > 0$. Q.E.D.

Proposition 1 can be understood as follows. If B and S have similar views about what is a reasonable division of surplus then it is efficient to contract only on the basic widget and leave the add-on for later (part (1)). The reason is that renegotiation will proceed smoothly if the add-on is required. In contrast if the parties contract on the add-on in certain states then renegotiation in other states becomes problematic because the presence of additional reference points hinders it. At the other extreme if B and S have very different views about what is reasonable then additional reference points do not hinder the renegotiation process – it is already as bad as it gets – and so contracting on whatever is possible is desirable (part (3)). Finally, part 2 says that contracting on unlikely events is undesirable since the benefits are small whereas the hindering effect of the additional reference points on renegotiation in other states is large.

We can see why it is important for our results that there is some ambiguity about the verifiable contingency: E must contain (at least) two states. If $v_4 = v_3$ and $c_4= c_3$, $\alpha v_3 + (1 - \alpha)c_3 - \beta v_4 - (1 - \beta)c_4 \geq 0$ and so, by Lemma 1, $L_2 = \pi_2\theta(\alpha - \beta)G_2$. Since contract 2 yields zero deadweight losses in event E, contract 2 then dominates contract 1. It is worth noting that in this case setting p to divide the gains from trade evenly in E achieves $L_2 = \pi_2\theta(\alpha - \beta)G_2$. Indeed this result is general. If every verifiable contingency where the add-on is efficient is a single state the price for the add-on in that contingency can be chosen to divide the surplus evenly in that state; and then in a non-verifiable contingency none of the reference points will hinder renegotiation given that $\alpha \geq ½ \geq \beta$.

One proviso should be noted. We have assumed that when there are multiple reference points each party will choose the one most favorable to him or her (although this assumption could be relaxed—similar conclusions would be reached if reference points were chosen randomly). One could, however, argue that parties who have similar views about how the surplus is divided – parties for whom $(\alpha - \beta)$ is small – will also agree about what is a reasonable reference point. Only empirical work or experiments can determine whether we are right to suppose that these two dimensions – division of the surplus and choice of reference point – are independent[11]. For further discussion about the possible determinants of $(\alpha - \beta)$, see Section 4.

3. A model of price variation

In the first model the issue is which kind of widget should be traded. We now consider a situation where the same kind of widget should always be traded but there has to be a variation in price. The simplest way to justify the latter is to suppose that ex post trade is voluntary; that is, trade will occur if

[11] Our intuition is that feelings about surplus division and reference points are distinct. Suppose that a contract says that S will supply B with a widget except if state s occurs. In actuality state s', similar to but different from s, occurs. S might argue that since s' is similar to s she should be excused from supplying. B might argue that precisely because s was mentioned but s' wasn't S should *not* be excused. Such a disagreement seems to have little to do with differences in α, β.

and only if the price lies between the buyer B's value v and the seller S's cost c. (This could be because if trade does not occur it is not clear who is responsible and so specific performance cannot be enforced.) One example of this is an at-will employment contract where B is the employer and S is the employee and either party can quit ex post. (S's cost may then be an opportunity cost or a reservation wage.) Voluntary trade takes place only at a price p such that $v \geq p \geq c$.[12] The problem is that, if v, c are stochastic, it may be impossible to find a singe price p such that $v \geq p \geq c$ whenever $v > c$.

The question we are interested in is: under what conditions will the parties find it desirable to index their contract to a (possibly noisy) signal of the state of the world? (As will be seen shortly we will actually ask a slightly different question.)

We know from Hart and Moore (2008) that in a voluntary trade model the first-best can be achieved if only one of v, c is stochastic. To make things simple we suppose that v, c are independent and each can take only two values: $v = v_H$ or v_L, $c = c_H$ or c_L, where $v_H > v_L > 0$, $c_H > c_L > 0$. Hence there are four possible outcomes; see Figure 2.

State	Probability	Payoff
s1	π_1	(v_H,c_H)
s2	π_2	(v_H,c_L)
s3	π_3	(v_L,c_H)
s4	π_4	(v_L,c_L)

Figure 2

We assume that $\pi_i > 0$ for all i.

There are several possible configurations of payoffs where there is a single price p such that $v > p > c$ whenever $v > c$, in which the first-best can be achieved. For future reference we call these "the first-best cases".[13] The only case where the first-best cannot be achieved is $v_H > c_H > v_L > c_L$; see Figure 3.

[12] We normalize the no trade price to be zero. We also suppose that lump sum transfers can be used to divide up the surplus in a way that reflects market competition at date 0. Finally, we follow Hart and Moore (2008) in assuming that there is no shading if there is no trade.

[13] If $v_L > c_H$, the first-best can be achieved with a contract that specifies $v_L > p > c_H$. (This guarantees trade in all states.) If $v_H > c_H > c_L > v_L$, the first-best can be achieved with a contract that specifies $v_H > p > c_H$. (This guarantees trade in s1 and s2.) If $c_H > v_H > v_L > c_L$, the first-best can be achieved with a contract that specifies $v_L > p > c_L$. (This guarantees trade in s2 and s4.) If $c_H > v_H > c_L > v_L$, the first-best can be achieved with a contract that specifies $v_H > p$

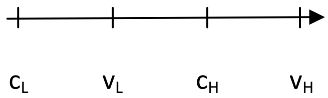

Figure 3

Here there is no price p such that $v_H \geq p \geq c_H$ and $v_L \geq p \geq c_L$.[14]

We ask the following question. Given that the first-best cannot be achieved, is it better for the parties to choose some (single price) contract or no contract at all? While this may seem at first sight different from the indexation question posed above it is in fact closely related. Suppose that the only thing that can be verified is that one of the four states s1,s2,s3,s4 has occurred. Then indexing on E is equivalent to writing a contract and non-indexing to writing no contract[15].

<u>No contract</u>

If no contract is written the parties will negotiate from scratch at date 1. As in Section 2 this yields expected deadweight losses equal to

$L = \pi_1\theta(\alpha - \beta)(v_H - c_H) + \pi_2\theta(\alpha - \beta)(v_H - c_L) + \pi_4\theta(\alpha - \beta)(v_L - c_L).$

Recall that since $v_L < c_H$ trade does not occur in s3.

$> c_L$. (This guarantees trade in s2.) If $c_H > c_L > v_H > v_L$, any contract achieves first-best since there are no gains from trade.

[14] As Hart and Moore (2008) show, a contract specifying a price range also cannot achieve the first-best in this case. Price ranges do not change our basic results and in what follows we ignore them.

[15] There is a qualification. In the typical indexation situation the question is whether it is better to have a non-indexed contract or an indexed contract, whereas we compare no contract to some contract. We could include the possibility of some contract by assuming an additional state 0 that occurs with high probability; it will typically be optimal for the parties to specify a price p that guarantees trade in this state unless α and β are extremely close. The question then is whether the parties choose a different price conditional on E. Note that one difference between the indexation case and the one in the text is that if the parties do not index on E "trade in state 0 at price p" will serve as a reference point for renegotiation in E. In our case, if the parties choose no contract, then the only reference point for renegotiation in E is no trade.

Contract with price p

We will not provide a full analysis of the optimal contract, but will compare "a contract" with "no contract" in each state of the world.

There are various possibilities according to the level of p.

Case 1: $v_H \geq p \geq c_H$

There is no aggrievement or shading in s1 or s2 since trade takes place at the price p.

Consider s4. Here renegotiation is required for trade to occur. As in Section 2 there are multiple reference points for renegotiation: no trade, trade in s1 at price p, and trade in s2 at price p. This means that S can justify any of the following prices in s4:

(3.1) $p_S = c_L + \alpha(v_L - c_L)$,

(3.2) $\hat{p}_S = c_L + p - c_H + \alpha(v_L - c_L - v_H + c_H)$,

(3.3) $\tilde{p}_S = p + \alpha(v_L - v_H)$.

However, S recognizes that B will never pay more than v_L in s4. It is easy to see that $\tilde{p}_S > \hat{p}_S$. Hence, S's entitlement is

(3.4) $\mathrm{Min}(v_L, \mathrm{Max}(p_S, \tilde{p}_S))$.

Similarly B can justify any of

(3.5) $p_B = c_L + \beta(v_L - c_L)$,

(3.6) $\hat{p}_B = c_L + p - c_H + \beta(v_L - c_L - v_H + c_H)$,

(3.7) $\tilde{p}_B = p + \beta(v_L - v_H)$,

but recognizes that S will never accept less than c_L. Since $\hat{p}_B < \tilde{p}_B$, B's entitlement is

(3.8) $\mathrm{Max}(c_L, \mathrm{Min}(p_B, \hat{p}_B))$.

It follows that deadweight losses in s4 are given by

(3.9) $L_4' = \theta[\mathrm{Min}(v_L, \mathrm{Max}(p_S, \tilde{p}_S)) - \mathrm{Max}(c_L, \mathrm{Min}(p_B, \hat{p}_B))]$.

Under no contract deadweight losses in s4 are $\theta(\alpha - \beta)(v_L - c_L)$. Note that, given $\alpha < 1$, $\beta > 0$, $v_L > p_S$ and $p_B > c_L$. It is then easy to see that

(3.10) $\text{Min}(v_L,\text{Max}(p_S, \tilde{p}_S)) - \text{Max}(c_L,\text{Min}(p_B, \hat{p}_B)) \geq p_S - p_B = (\alpha - \beta)(v_L - c_L)$,

i.e., deadweight losses in s4 are at least as great as under no contract. Lemma 2 provides a sufficient condition for strict inequality in (3.10).

Lemma 2. A sufficient condition for

$L_4' > \theta(\alpha - \beta)(v_L - c_L)$ is that

$[(1 - \beta)c_H + \beta v_H,(1 - \alpha)c_L + \alpha v_H] \cap [c_H,v_H] = \varnothing$.

Proof: Given $v_L > p_S$, $p_B > c_L$, (3.10) must hold with strict inequality if either $p_S < \tilde{p}_S$ or $p_B > \hat{p}_B$. Suppose the contrary: $p_S \geq \tilde{p}_S$ and $p_B \leq \hat{p}_B$. Then

(3.11) $p \leq (1 - \alpha)c_L + \alpha v_H$, $p \geq (1 - \beta)c_H + \beta v_H$

However, $v_H \geq p \geq c_H$. Hence

$[(1 - \beta)c_H + \beta v_H,(1 - \alpha)c_L + \alpha v_H] \cap [c_H,v_H] \neq \varnothing$,

which is a contradiction.

Q.E.D.

Lemma 2 can be understood as follows. The sufficient condition will be satisfied if α, β are close. In this case the multiple reference points associated with the Case 1 contract will hinder the renegotiation process in s4. On the other hand the sufficient condition will not be satisfied if α, β are far apart. In this case there are many prices satisfying (3.11) that lie in the range $[c_H, v_H]$ and so the multiple reference points do not increase aggrievement in s4.

Case 2: $c_H > p > v_L$

Now trade takes place only in s2 under the contract. Renegotiation will occur in s1 and s4. Consider s1. B and S have two reference points: no trade and trade in s2 at p. Letting

(3.12) $p_S = c_H + \alpha(v_H - c_H)$,

(3.13) $\hat{p}_S = p + c_H - c_L + \alpha(c_L - c_H) = p + (1 - \alpha)(c_H - c_L)$,

we can write S's entitlement as

(3.14) $\text{Min}(v_H,\text{Max}(p_S, \hat{p}_S))$.

Similarly, letting

(3.15) $p_B = c_H + \beta(v_H - c_H)$,

(3.16) $\hat{p}_B = p + (1 - \beta)(c_H - c_L)$,

we can write B's entitlement as

(3.17) $\text{Max}(c_H,\text{Min}(p_B, \hat{p}_B))$.

Deadweight loss in s1 is given by

(3.18) $L_1' = \theta[\text{Min}(v_H,\text{Max}(p_S, \hat{p}_S)) - \text{Max}(c_H,\text{Min}(p_B, \hat{p}_B))]$.

By the same logic as in Case 1 L_1' cannot be lower than $\theta(p_S - p_B) = \theta(\alpha - \beta)(v_H - c_H)$. Lemma 3 provides a sufficient condition for it to be higher.

Lemma 3. A sufficient condition for

$L_1' > \theta(\alpha - \beta)(v_H - c_H)$ is that

$[(1 - \beta)c_L + \beta v_H,(1 - \alpha)c_L + \alpha v_H] \cap [v_L,c_H] = \varnothing$.

Proof: As in Lemma 2 it is sufficient to show that either $p_S < \hat{p}_S$ or $p_B > \hat{p}_B$. Suppose the contrary: $p_S \geq \hat{p}_S$ and $p_B \leq \hat{p}_B$. Then

(3.19) $p \leq (1 - \alpha)c_L + \alpha v_H$, $p \geq (1 - \beta)c_L + \beta v_H$.

But $c_H > p > v_L$. Hence

$[(1 - \beta)c_L + \beta v_H,(1 - \alpha)c_L + \alpha v_H] \cap [v_L,c_H] \neq \varnothing$,

which is a contradiction.

Q.E.D.

Some insight into Lemma 3 can be gained as follows. In case 2 there are only two reference points for renegotiation in s1: no trade at price zero and trade in s2 at price p. If p divides the gains from trade in s2 evenly then this second reference point will lead the parties to feel entitled to a price in s1 that is between $c_H + \beta(v_H - c_H)$ and $c_H + \alpha(v_H - c_H)$ (this follows from (3.13), (3.16)); thus the second reference point will not hinder renegotiation. Hence if the average of c_L, v_H lies in $[v_L,c_H]$ (required for Case 2), we have found a contract that achieves the same deadweight losses in s1 as no contract. (Note that the average of c_L, v_H lying in $[v_L,c_H]$ is sufficient, although not necessary, for $[(1-\beta)c_L + \beta v_H,(1-\alpha)c_L + \alpha v_H] \cap [v_L,c_H] \neq \varnothing$.)

Now consider s4. Let

(3.20) $p_S = c_L + \alpha(v_L - c_L)$,

(3.21) $\hat{p}_S = p + \alpha(v_L - v_H)$.

Then S's entitlement is

(3.22) $\mathrm{Min}(v_L,\mathrm{Max}(p_S, \hat{p}_S))$.

Similarly, B's entitlement is

(3.22) $\mathrm{Max}(c_L,\mathrm{Min}(p_B, \hat{p}_B))$,

where

(3.24) $p_B = c_L + \beta(v_L - c_L)$,

(3.25) $\hat{p}_B = p + \beta(v_L - v_H)$.

Let L_4' be the deadweight losses from the above contract in s4. By the same logic as in Case 1 L_4' cannot be lower than the deadweight losses from no contract. A necessary condition for them to be the same is

(3.26) $p_S \geq \hat{p}_S$, $p_B \leq \hat{p}_B$,

i.e.,

(3.27) $p \leq (1-\alpha)c_L + \alpha v_H$, $p \geq (1-\beta)c_L + \beta v_H$.

Hence we obtain a similar result to Lemma 3.

<u>Lemma 4.</u> A sufficient condition for

$L_4' > \theta(\alpha - \beta)(v_L - c_L)$ is that

$[(1-\beta)c_L + \beta v_H,(1-\alpha)c_L + \alpha v_H] \cap [v_L,c_H] = \varnothing$.

Lemma 4 can be understood in the same way as Lemma 3.

Case 3: $v_L \geq p \geq c_L$

Now trade takes place in s2, s4 under the contract. Renegotiation is required in s1. Let

(3.28) $p_S = c_H + \alpha(v_H - c_H)$,

(3.29) $\hat{p}_S = c_H + p - c_L + \alpha(v_H - c_H - v_L + c_L)$,

(3.30) $\tilde{p}_S = c_H + p - c_L + \alpha(c_L - c_H)$.

It is easy to see that $\tilde{p}_S < \hat{p}_S$. Hence S's entitlement is

(3.31) $\mathrm{Min}(v_H, \mathrm{Max}(p_S, \hat{p}_S))$.

Similarly B's entitlement is

(3.32) $\mathrm{Max}(c_H, \mathrm{Min}(p_B, \tilde{p}_B))$,

where

(3.33) $p_B = c_H + \beta(v_H - c_H)$,

(3.34) $\tilde{p}_B = c_H + p - c_L + \beta(c_L - c_H)$.

Let L_1' be the deadweight losses from the above contract in s1. By the same logic as in Case 1 L_1' cannot be lower than the deadweight losses from no contract. A necessary condition for them to be the same is

(3.35) $p_S \geq \hat{p}_S$, $p_B \leq \tilde{p}_B$,

which yields

(3.36) $p \leq (1 - \alpha)c_L + \alpha v_L$, $p \geq (1 - \beta)c_L + \beta v_H$.

This establishes

Lemma 5. A sufficient condition for

$L_1' > \theta(\alpha - \beta)(v_H - c_H)$ is that

$[(1 - \beta)c_L + \beta v_H,(1 - \alpha)c_L + \alpha v_L] \cap [c_L,v_L] = \varnothing$.

Lemma 5 can be understood in the same way as Lemma 2.

This completes our analysis of the case where B and S write an initial contract.

Proposition 2 is analogous to Proposition 1.

Proposition 2. Assume $v_H > c_H > v_L > c_L$.

(1) Suppose $\frac{1}{2}c_L + \frac{1}{2}v_H \notin [v_L,c_H]$. Then for $(\alpha - \beta)$ sufficiently small, no contract is strictly superior to a contract.

(2) If $(\alpha - \beta)$ is sufficiently large, some contract is strictly superior to no contract.

Proof.

(1) Consider the case $\alpha = \beta = \frac{1}{2}$. Obviously the deadweight losses from no contract are zero. Consider any contract. If Case 1 applies ($v_H \geq p \geq c_H$), then by Lemma 1 $L_4' > 0$ (the interval $[\frac{1}{2}c_H + \frac{1}{2}v_H, \frac{1}{2}c_L + \frac{1}{2}v_H]$ is degenerate). If Case 2 applies ($c_H > p > v_L$), then by Lemmas 2 and 3 L_1' and $L_4' > 0$ (since by assumption $\frac{1}{2}c_L + \frac{1}{2}v_H \notin [v_L,c_H]$). If Case 3 applies ($v_L \geq p \geq c_L$), $L_1' > 0$ by Lemma 4 (the interval $[\frac{1}{2}c_L + \frac{1}{2}v_H, \frac{1}{2}c_L + \frac{1}{2}v_L]$ is degenerate). Hence the expected deadweight losses are strictly positive under any contract. That is, a contract is dominated by "no contract". By continuity this will also be true for α, β close to ½.

(2) When α is close to 1 and β is close to zero, the deadweight losses from no contract are approximately $L = \pi_1\theta(v_H - c_H) + \pi_2\theta(v_H - c_L) + \pi_4\theta(v_L - c_L)$. Consider a contract with $v_H \geq p \geq c_H$ (as in Case 1 above). Then by (3.9) the deadweight losses from this contract are approximately $L_c = \pi_4\theta(v_L - c_L)$. Hence $L_c < L$ given that $\pi_1, \pi_2 > 0$.

Q.E.D.

Proposition 2 in combination with the "first-best cases" considered earlier in this section provide an answer to the question: Given that the value-cost range is subject to a shock, is it better to write an ex ante contract or to wait and renegotiate ex post? The first-best cases tell us that if it is possible to find a single price p that lies in the value-cost range whenever value exceeds cost then a contract is better. On the other hand Proposition 2 tells us that if the first-best cannot be achieved then (subject to a proviso) a contract is better if parties' views about the appropriate division of surplus are very different but worse if they are similar.

The proviso is the condition $\frac{1}{2}c_L + \frac{1}{2}v_H \notin [v_L, c_H]$. How important is this condition? Let us give an informal argument as to why the answer is not very. The four state example considered in this section is very special. Suppose we add more states—in fact make v,c continuous (but stick with the assumption that v,c are independent). Then if $v > c$ with probability 1 it will be possible to find a price p such that $v > p > c$ with probability 1 (choose p between the maximum c and minimum v). In this case the first-best can be achieved. However, if $v < c$ with positive probability, the scenario where, for some p, there is a single state where $v > p > c$ will never occur. (This corresponds to Case 2 in the analysis above.) But this is the scenario where the condition $\frac{1}{2}c_L + \frac{1}{2}v_H \notin [v_L, c_H]$ is needed for Proposition 2. This suggests that in a more general example the condition will not be needed.

4. Summary and conclusions

In this paper we have investigated when and why parties will deliberately write incomplete contracts even when contract-writing costs are zero. We have argued that adding a contingency of the form, "The buyer will require an extra good or service in event E", has a benefit and a cost. The benefit is that there is less to argue about in event E; the cost is that the reference point provided by the extra service in event E may increase argument costs in states outside E. Similarly indexing a price or wage to an exogenous variable has the benefit that if this variable tracks the buyer's value and seller's cost closely then breakdown in trade can be avoided; but the cost that if the index does not track value and cost closely the reference point provided by the indexation may make renegotiation harder when trade does break down.

Our principal result is that the relative benefit and cost of adding a contingency or indexing will be sensitive to how closely the parties agree about what is a reasonable division of surplus when an incomplete contract is renegotiated. The benefit is likely to exceed the cost when parties have very different views about what is a reasonable division of surplus, but the opposite will be the case if they have shared views. Under the latter conditions an incomplete contract will be strictly optimal. Our results can shed light on why wage indexation, although observed in some situations (see Card (1986) and Oyer (2004)), is not more common.

It is worth considering how our theory's implications differ from those of a theory based on asymmetric information. Consider the Nanny example in the introduction where the question is why a late fee is not introduced. The asymmetric information explanation would be that introducing the late fee might signal to the Nanny that the employer knows that he is unpunctual, which makes the job less attractive. But this problem could be presumably solved through the choice of a high late fee.

Or take the case of wage indexation. If an employee is offered a contract whereby the wage is indexed on some signal, the employee might think that the employer already knows that the signal will be such that the employee's wage is low, making the contract less attractive. But this would suggest that in an optimal contract the wage should not vary much with the index, not that it should not vary at all. Only by introducing costs of contractual clauses (as in Spier (1992)) can one explain a complete lack of indexation.

In contrast in our theory, introducing a late fee or any amount of indexation has a discontinuous effect: it introduces a brand new reference point. We have seen that in some circumstances the cost of doing this outweighs the benefit.

Our theory also has different implications from the asymmetric information one regarding the timing of incompleteness. Signaling favorable private information is particularly important at the beginning of a relationship. In our theory one possible explanation for similar views about the division of surplus is the history of the relationship between the buyer and the seller. If the parties have interacted before they may have grown to know and like each other, with the implication that each will become more generous about sharing surplus (see the social influence theory of Kelman (1958)). Therefore we would expect contracts to become less complete in long-term relationships, but be more complete when such relationships are formed -- in contrast to the asymmetric information theory[16].

Finally, our approach may also be able to explain why parties often use general rather than specific language in contracts. For example, parties negotiating acquisitions frequently include a clause that excuses the buyer if the target seller suffers a "material adverse change" (see Schwartz and Scott (2010)). According to our theory the advantage of a general clause is that it creates a neutral reference point: In terms of the model of Section 2 it is like describing states s2-s4, rather than event E, as a situation where the add-on should be provided. In contrast spelling out particular contingencies that qualify as a material adverse change may complicate renegotiation in other contingencies that are not easily described but where the parties also intended to excuse the buyer. Asymmetric information theories do not seem to have much to say about this issue.

Our results depend importantly on how similar or different views the parties have about the division of surplus. It is natural to ask what determines empirically whether parties' views about the division of surplus are likely to be similar or different. At this point we do not have a very good answer to this question. It seems reasonable that it has something to do with norms, trust, social capital, and empathy. A "dog-eat-dog" world may be one where each party feels entitled to the best outcome possible. A more civilized world may be one where sharing the surplus from renegotiation comes more naturally. The vast empirical and experimental literature on ultimatum, dictator and public goods games

[16] However, a complete analysis would have to incorporate the possibility that parties will anticipate this potential warming at the beginning of their relationship, which would complicate matters considerably.

(see, e.g., Camerer (2003)) suggests that views of a reasonable division of surplus may vary across countries, societies, etc., in a systematic way. Our theory predicts that one should expect to see less complete contracts in situations where people are more empathetic toward each other and more complete contracts when people are less empathetic. Some guidance about the importance of shared views for building trust can be obtained from the relationship marketing literature (Morgan and Hunt (1994)). Trust has two dimensions: credibility and benevolence. The first is related to ideas formalized in the repeated games literature in economics (see Malcomson (2013) for a recent survey). The second is concerned with shared values as trust develops through interpreting and assessing whether the other party is interested in his partner's best interests. Parties with shared values have a similar definition of what behaviors and policies are appropriate and can therefore better understand what drives the partner's behavior (see the attribution theory of Heider (1958)).

As mentioned in the introduction, there is a sizeable law and economics literature on contractual incompleteness. We have noted that one difference between our paper and this literature is that the literature tends to assume a fixed cost of writing or enforcing contractual clauses. To understand other differences it is useful to make the distinction introduced in Ayres and Gertner (1992) between "obligationally incomplete" and "insufficiently state contingent" contracts. The first refer to contracts that cannot be enforced as they stand or are ambiguous, e.g., a contract might require S to supply a widget to B even in a situation where this is impossible; or might require S to supply a widget by a particular time but not say what should happen if S fails to do this. Some sort of judicial (or outside) interpretation seems required to complete such a contract. The second – insufficiently state contingent – refers to a contract that is fully specified in all circumstances but which does not contain all the contingent clauses that the parties would like. In this case the parties do not require judicial (or outside) intervention (although they might benefit from it).

Our paper is about the second situation rather than the first, whereas much of the law and economics literature is about the first (see, e.g., Ayres and Gertner (1989, 1992), Shavell (1980)). Indeed we have ignored the role of courts (or other outsiders, such as arbitrators) in interpreting contracts. In future work it would be desirable to introduce the courts. A well-functioning judicial system may allow the parties to economize on the number of contingencies they include themselves, thereby reducing the number of reference points. The parties can rely on the courts to tell them what to do in some verifiable states; while in other states renegotiation may proceed smoothly given that the judicial solution may loom less large as a reference point than a party-induced remedy.

An analysis of legal rules in a world where parties write incomplete contracts for the kinds of reasons explored here is an interesting and challenging topic for future research.

REFERENCES

Aghion Philippe and Patrick Bolton. 1987. "Contracts as a Barrier to Entry." American Economic Review 77(3): 388-401.

Aghion, Philippe and Benjamin Hermalin. 1990. "Legal Restrictions on Private Contracts Can Enhance Efficiency." Journal of Law, Economics and Organizations 6(2): 381-409.

Al-Najjar, Nabil I., Luca Anderlini and Leonardo Felli. 2006. "Undescribable events." Review of Economic Studies 73(4):849-868.

Ayres, Ian and Robert Gertner. 1989. "Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules." Yale Law Journal 99(1): 87-130.

Ayres, Ian and Robert Gertner. 1992. "Strategic Contractual Inefficiency and the Optimal Choice of Legal Rules." Yale Law Journal 101(4):729-773.

Bajari, Patrick and Steven Tadelis. 2001. "Incentives versus Transaction Costs: A Theory of Procurement Contracts." Rand Journal of Economics 32(3): 387-407.

Baker, George, Robert Gibbons and Kevin J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts." Quarterly Journal of Economics 109(4): 1125-1156.

Benabou, Roland and Jean Tirole. 2003. "Intrinsic and Extrinsic Motivation." Review of Economic Studies 70(3): 489-520.

Benabou, Roland and Jean Tirole. 2006. "Incentives and Prosocial Behavior." American Economic Review 96(5): 1652-1678.

Bernheim, B. Douglas and Michael D. Whinston. 1998. "Incomplete Contracts and Strategic Ambiguity." American Economic Review 88(4):902-932.

Bolton, Patrick and Antoine Faure-Grimaud. 2010. "Satisficing Contracts." Review of Economic Studies 77(3):937-971.

Camerer, Colin. 2003. *Behavioral Game Theory.* Princeton: Princeton University Press.

Card, David. 1986. "An Empirical Model of Wage Indexation Provisions in Union Contracts." Journal of Political Economy 94(3): S144-S175.

Chakravarty, Surajeet and W. Bentley MacLeod. 2009. “Contracting in the Shadow of the Law.” Rand Journal of Economics 40(3):533-557.

Dye, Ronald. 1985. “Costly Contract Contingencies.” International Economic Review 26(1): 233-250.

Fehr, Ernst, Oliver Hart and Christian Zehnder. 2012. “How Do Informal Agreements and Renegotiation Shape Contractual Reference Points?” mimeo, Harvard University.

Hart, Oliver and John Moore. 2008. “Contracts as Reference Points.” Quarterly Journal of Economics 123(1): 1-48.

Hart, Oliver. 2009. “Hold-up, Asset Ownership, and Reference Points.” Quarterly Journal of Economics 124(1):267-302.

Heider, Fritz. 1958. *The Psychology of Interpersonal Relations.* Hillside (NJ): Erlbaum.

Herold, Florian. 2010. “Contractual Incompleteness as a Signal of Trust.” Games and Economic Behavior 68: 180-191.

Herweg, Fabian and Klaus M. Schmidt. 2012. “Loss Aversion and Ex Post Inefficient Renegotiation.” mimeo, University of Munich.

Kelman, Herbert C. 1958. “Compliance, Identification, and Internalization: Three Processes of Attitude Change.” The Journal of Conflict Resolution 2(1): 51-60.

Kukharskyy, Bohdan. 2013. “Fairness and Global Sourcing.” mimeo, University of Passau.

Kvaløy, Ola and Trond E. Olsen. 2009. “Endogenous verifiability and Relational Contracting.” American Economic Review 99: 2193-2208.

Malcomson, James. 2013. “Relational Incentive Contracts.” In *The Handbook of Organizational Economics*, edited by Robert Gibbons and John Roberts, 1014-1065. Princeton: Princeton University Press.

Morgan, Robert M. and Shelby D. Hunt. 1994. “The Commitment-Trust Theory of Relationship Marketing.” Journal of Marketing 58(3): 20-38.

Oyer, Paul. 2004. “Why Do Firms Use Incentives That Have No Incentive Effects?” The Journal of Finance 59(4):1619-1649.

Posner, Richard. 1986. *The Economic Analysis of Law.* Boston: Little, Brown and Company.

Rothschild, Michael and Joseph Stiglitz. 1976. “Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information.” Quarterly Journal of Economics 90(4): 629-649.

Schmidt, Klaus M. and Monika Schnitzer. 1995. “The Interaction of Explicit and Implicit Contracts.” Economics Letters 48(2): 193-199.

Schwartz, Alan and Robert E. Scott. 2010. "Contract Interpretation Redux." Yale Law Journal 119: 926-964.

Shavell, Steven. 1980. "Damage Measures for Breach of Contract." Bell Journal of Economics 11(2): 466-490.

Spier, Kathryn. 1992. "Incomplete Contracts and Signalling." RAND Journal of Economics 23(3): 432-443.

Tirole, Jean. 2009. "Cognition and Incomplete Contracts." American Economic Review 99(1):265-294.

Weitzman, Martín L. 1984. *The Share Economy: Conquering Stagflation*. Harvard University Press.