Overcoming Contractual Incompleteness: The Role of Guiding Principles

by

David Frydlinger and Oliver Hart*

January 2019

Preliminary

Abstract: We develop a model where a buyer and seller contract over a service. The contract encourages the seller to invest and provides a reference point for the transaction. In normal times the contract works well. But with some probability an abnormal state occurs and the service must be modified. This puts the parties below their reference payoffs and may cause costly disagreement. We discuss why neither classical mechanisms nor lawyers' standard approaches adequately deal with this issue. The adoption by the parties of guiding principles such as loyalty and equity as part of their contract can help.

*Frydlinger: Managing Partner, Cirio law firm (email: David.Frydlinger@cirio.se); Hart: Department of Economics, Harvard University (email: ohart@harvard.edu). We have benefited from the comments of participants at the Mini-Conference on Relational Contact Management, Copenhagen Business School, October 2018.

1. Introduction

Contractual incompleteness lies at the heart of several areas of economics, including the theory of the firm, corporate finance and governance, and the choice between public and private ownership. However, a key question that has not been fully resolved is, why are contracts incomplete? Answers have been attempted based on contracting costs, bounded rationality, or the inability to verify certain states of the world, but none has been entirely satisfactory.

The purpose of this paper is two-fold. First, we provide a model of contractual incompleteness that is, we believe, novel (although, of course, it shares elements with a number of previous models). Second, and just as important, we discuss how both in the model and in practice parties can overcome some of the problems of contractual incompleteness. Our approach involves the parties' adoption of what we call "guiding principles," such as loyalty and equity. We suggest that these principles can be 'activated' through ex-ante and ex-post communication. We argue that a contract that specifies standard elements such as the price and the nature of the good or service to be traded can, in combination with the adoption of guiding principles, perform better than either a standard contract alone; or the adoption of guiding principles alone. The prescription for dealing with contractual incompleteness is based on the experience of one of the authors (Frydlinger).

In the model, Buyer B and Seller S meet at date 0, and plan to trade at date 1. To fix ideas, we imagine that B is putting on a concert and that S is providing band services. A contract specifying the price and (many aspects of) the quality of the music is written at date 0. B's value for the concert is taken to be non-verifiable (it could even be non-monetary) but S's costs are verifiable. However, there is moral hazard: S can affect costs (but not B's value) through her (unobservable) effort choice. For this reason it is efficient to make S the residual claimant on costs. Given this assignment, a first-best effort choice is achieved and there is no hold-up problem (in contrast to most of the incomplete contracting models in the literature).

Much of the time (with probability $1-\pi$), this contract delivers a desirable outcome. However, with probability π an abnormal (bad) state occurs. Again to fix ideas, imagine that this is a state where the usual trumpeter that S was planning to hire for the band, Eve, breaks her finger and is unavailable. (The abnormal state covers many possibilities like this and that is why we will suppose that it cannot be contracted on.) So a replacement must be found at relatively short notice. Under non-integration (where B and S are separate entities), the replacement is S's choice (she has residual decision rights over the band composition). Imagine that two trumpeters are available, Adam and George. Both are of worse quality than Eve. However, George is cheaper but worse than Adam. Under standard assumptions of rationality and self-interest, S would threaten to choose George (since S is the residual claimant on costs), but if Adam is more efficient the parties will renegotiate to Adam. Since we will assume that the incremental costs of George and Adam relative to Eve are not affected by S's effort and that there is symmetric information, the first-best is achieved.

In Section 3 we explain why we do not believe that things will work as smoothly as this in practice. The reason is that the contract creates reference payoffs. Since the abnormal state is unusual each party's reference payoff will be biased toward his or her payoff in the normal state. If the abnormal state

occurs, each party will be disappointed and will want to minimize their losses, possibly at the expense of the other party. In particular, each party will expect the other party to behave "fairly". However, in the absence of communication the parties may have different views of what is fair ("self-serving biases"). An implication of this is that B may be upset and angry if S threatens to choose George: his reasoning will be that the least S could do in this unfortunate eventuality is to mitigate B's losses by choosing Adam. B may be even more upset if S asks B to pay S to choose Adam rather than George ("extortion"). We follow Hart and Moore (2008) in supposing that an aggrieved party will retaliate by withholding (noncontractible) cooperation ("shading"). This creates deadweight costs.

In Section 5 we argue that there is a way to reduce deadweight costs. Ideally the parties would change their reference points so that each party's reference payoff in the abnormal state equals their payoff in that state, that is, they fully anticipate and accept the self-interested behavior of the other party. There would then be no aggrievement or shading. We do not believe that this is feasible. Instead we argue that the parties can at least partially align their preferences by adopting "guiding principles". In our formal analysis we focus on the principle of loyalty. By this we mean a principle that obliges each party to treat the other party's interest as having a comparable importance to their own. We formalize this by supposing that through communication and discussion at dates 0 and 1 the parties can transform their payoffs so that each party puts weight λ on the other party's payoff, where $0 < \lambda < 1$. However, communication and discussion are not costless. The higher λ is, the higher are the costs that must be incurred in discussion and communication.

We solve for the optimal contract under communication. Since communication helps only in the abnormal state, our analysis yields the plausible result that communication is worthwhile only if the abnormal state is somewhat likely. To put it another way, in routine situations where unexpected events are rare, a standard contract may suffice.

The paper is organized as follows. In Section 2, we set up the model and describe the optimal contract under standard assumptions of rationality and self-interest. In Section 3 we explain why we do not think that the standard solution will work in practice, given the parties' behavioral biases. In Section 4 we describe the lawyers' standard approach to writing a contract and explain why we do not think this will work in our setting either. In Section 5 we analyze how adoption of a loyalty principle can help. In Section 6 we consider how things change if B purchases S's operations and turns S into an employee. In Section 7 we discuss the use of complementary principles, such as equity. Finally, Section 8 concludes.

2. Model

Buyer B is putting on a concert. Seller S is providing the band. B and S contract at date 0 and the concert takes place at date 1. There is no discounting, and the parties are risk neutral and wealth unconstrained.

B's revenue from the concert is (in the normal state) v, which is observable but not verifiable (and nontransferable). S's cost is (in the normal state) c, which is verifiable (and transferable). However, c

depends on S's effort e, which has a private cost equal to e and which is observable only to S ("moral hazard"). We make the standard assumption that $c(e)$ is decreasing in e and exhibits diminishing returns: $c' < 0,\ c'' > 0,\ c'(0) = -\infty,\ \lim_{e\to\infty} c'(e) > -1$.

In the normal state N which occurs with probability $1-\pi$, the usual trumpeter, Eve, plays in the band. B's revenue $= v$ and S's cost $= c(e)$. With probability π an abnormal state A occurs. The leading case is where π is fairly small, but we also allow for the case where π is large (abnormal is then a misnomer). In the abnormal state, Eve is unavailable because she has broken her finger. There are two alternative trumpeters, Adam and George. Since these are late replacements, it is reasonable to suppose that they more expensive than Eve and of worse quality. Denoting incremental cost and value by $\Delta c, \Delta v$, respectively, we have:

	Cost relative to Eve	Value relative to Eve
Adam	$\Delta\, c_a > 0$	$\Delta\, v_a < 0$
George	$\Delta\, c_g > 0$	$\Delta\, v_g < 0$

An important assumption we will make is that $\Delta c_a,\ \Delta c_g, \Delta v_a,\ \Delta v_g$ are independent of e. As a result, as we will see, the hold-up problem is easily avoided in this model. We also assume that it is still worth going ahead with the concert in the abnormal state: $v - c(0) + Max(\Delta v_a - \Delta c_a,\ \Delta v_g - \Delta c_g) > 0$.

We will suppose that Δv and Δc are observable but not verifiable; however, $c + \Delta c$ (like c in the normal state) is verifiable. In other words, we suppose that final cost in the abnormal state is verifiable but that the incremental cost of hiring Adam or George rather than Eve is not. One interpretation is that the final terms of Eve's contract have not yet been settled when she breaks her finger, and so the cost of having her (that is, what S's cost would have been if it was not necessary to replace Eve) can never be established. A similar assumption about the non-verifiability of incremental costs is made by Bajari and Tadelis (2001).

It is worth saying a few words about why we assume that v is non-verifiable. First, in many situations it is difficult to verify the value of a service to a buyer either because the buyer purchases many services and it is hard to keep them apart, or because the benefit is non-monetary. Second, the assumption can be seen as a simplifying one. It would be easy to generalize the analysis to the case where value is verifiable but there is moral hazard on the buyer's side: value is influenced by B's effort.

In contrast, we assume that cost is verifiable, but subject to moral hazard on the seller's side, because this seems realistic.

Without loss of generality assume that George is cheaper than Adam: $\Delta c_g < \Delta c_a$. The interesting case is where there is a conflict of interest: George is a worse trumpeter than Adam, that is, $\Delta v_g < \Delta v_a$.

We will allow both for the situation where Adam is more efficient than George ($\Delta v_a - \Delta c_a > \Delta v_g - \Delta c_g$) and for the situation where George is more efficient than Adam

($\Delta v_g - \Delta c_g > \Delta v_a - \Delta c_a$).

The first-best

A social planner chooses Eve's replacement in state A in an efficient manner, and chooses e to maximize expected net surplus. Ex-post surplus is given by

(2.1) $v - c(e)$ in state N,

and by

(2.2) $v - c(e) + Max\left(\Delta v_a - \Delta c_a\ , \Delta v_g - \Delta c_g\right)$ in state A,

and so expected net surplus is

(2.3) $v - c(e) - e + \pi\ Max\left(\Delta v_a - \Delta c_a\ , \Delta v_g - \Delta c_g\right)$.

The planner maximizes (2.3) with respect to e. The first-order condition is

(2.4) $c'\,(e) = -1$.

Denote the solution by e_{FB}. Note that since the abnormal state affects value and cost independently of e, the probability π of the abnormal state does not affect the first-best effort level.

Substituting e_{FB} in (2.3), we can write the first-best level of net surplus as

(2.5) $v - c(e_{FB}) - e_{FB} + \pi\ Max\left(\Delta v_a - \Delta c_a\ , \Delta v_g - \Delta c_g\right)$.

A simple contract that achieves the first-best under classical assumptions

In what follows we will suppose that state A cannot be contracted on: Eve's unavailability is only one of many things that can go wrong. We will also assume that B and S are separate entities ("non-integration"), and that S has the right to choose Eve's replacement (this qualifies as a "residual control right" or "residual decision right" in the sense of Grossman-Hart (1986)). Later we will consider the case where B purchases S's operations, acquiring residual control rights, and S becomes an employee.

In spite of the fact that the parties cannot contract on state A, there is a simple way to achieve the first-best under "classical" assumptions that the parties are self-interested and rational.

Let B and S agree on a price p for the band and that S will bear all of the costs, $c(e)$ in state N and $c(e) + \Delta c$ in state A, which recall are verifiable. In state N everything proceeds smoothly: Eve plays in the band and B pays p. The payoffs are given by:

(2.6) B's ex-post payoff in state $N = v - p$, S's ex-post payoff in state $N = p - c(e)$.

In state A, S must replace Eve. S's incentive is to choose George since he is cheaper. However, George is worse for B. If Adam is more efficient than George ($\Delta v_a - \Delta c_a > \Delta v_g - \Delta c_g$), the parties will renegotiate (since Δv and Δc are observable, there is symmetric information). Assume that the parties have equal bargaining power. Then in state A, the payoffs are given by:

(2.7) B's ex-post payoff in state $\mathrm{A} = v - p + \Delta v_g + \frac{1}{2}\, Max\big(\Delta v_a - \Delta c_a - (\Delta v_g - \Delta c_g), 0\big)$,

(2.8) S's ex-post payoff in state $A = p - c(e) - \Delta c_g + \frac{1}{2}\, Max\big(\Delta v_a - \Delta c_a - (\Delta v_g - \Delta c_g), 0\big)$.

Turning to ex-ante payoffs at date 0, these are given by

(2.9) B's expected payoff $= v - p + \pi \left[\Delta v_g + \frac{1}{2}\, Max\big(\Delta v_a - \Delta c_a - \Delta v_g + \Delta c_g, 0\big)\right]$,

(2.10) S's expected payoff $= p - c(e) - e + \pi \left[-\Delta c_g + \frac{1}{2}\, Max\big(\Delta v_a - \Delta c_a - \Delta v_g + \Delta c_g, 0\big)\right]$.

Under rational expectations, S will choose e to maximize (2.10), which yields the first-order condition

(2.11) $c'(e) = -1$.

The first-best is achieved.

Assume that there is a competitive market for sellers at date 0, and that the market-clearing expected return for a seller is $\overline{U}$. Then price p will satisfy

(2.12) $p - c(e_{FB}) - e_{FB} + \pi \left[-\Delta c_g + \frac{1}{2}\, Max\big(\Delta v_a - \Delta c_a - \Delta v_g + \Delta c_g, 0\big)\right] = \overline{U}$,

and so B's expected payoff equals

(2.13) $v - c(e_{FB}) - e_{FB} + \pi\, Max\big(\Delta v_a - \Delta c_a,\ \Delta v_g - \Delta c_g\big) - \overline{U}$.

In other words, B receives the first-best level of surplus minus $\overline{U}$, and S receives $\overline{U}$.

At the risk of belaboring the point, note that, in contrast to the standard literature, although the above contract is incomplete (state A cannot be contracted on), there is no underinvestment in e. The reason is that the incremental payoffs $\Delta c, \Delta v$ are independent of e. Hence, the contract for band services at price p avoids the hold-up problem and provides S with socially optimal investment incentives (choice of e). Note that if no contract at all were written at date 0 there *would* be a hold-up problem. At date 1 B and S would bargain over the gains from trade $v - c(e)$ in state N and $v - c(e) + Max\ (\Delta v_a - \Delta c_a,\ \Delta v_g - \Delta c_g)$ in state A. If they split these 50:50, S's ex-ante payoff will equal

(2.14) $\frac{1}{2}v - \frac{1}{2}\,c(e) + \frac{1}{2}\,\pi\, Max\big(\Delta v_a - \Delta c_a\,,\ \Delta v_g - \Delta\, c_g\big) - e,$

and so e will satisfy the first-order condition

(2.15) $\frac{1}{2}c'(e) = -1,$

which implies $e < e_{FB}$.

A simple example

Suppose that c(e) $= (1/80)e^2$. Then, solving (2.4) yields $e_{FB} = 40, c(e_{FB}) = 20.$ Let $v = 150,\ \Delta c_g = 6, \Delta v_g = -14,\ \Delta c_a = 8,\ \Delta v_a = -10.$ Then Adam is more efficient than George, and the first-best level of surplus is $150 - 20 - 40 - 18\pi = 90 - 18\pi$. The last term reflects the fact that relative to Eve Adam costs S 6 more and reduces B's value by 12. Finally, let $\overline{U} = 45 - 9\pi$. Then $p = 105 - 4\pi$ and each party's payoff will equal $45 - 9\pi$.

3. Why we think that the classical contracting solution will not work in practice

In this section we explain why the simple contract of Section 2 is unlikely to work in practice. In our view a key factor is that the contract creates reference points. The creation of reference points has both positive and negative elements. On the positive side, the contract nails things down in the normal state and avoids disagreement there. But, on the negative side, the contract-created reference points may cause problems in the abnormal state.

Specifically, we suppose that each party has a reference payoff based on the probability distribution of payoffs under the contract (as in Kőszegi and Rabin (2006)). When a party is below his reference payoff he will become sensitive to the actions of the other party, expecting the other party to behave "fairly", and being disappointed or aggrieved if this does not happen. Aggrievement will lead to the withdrawal of cooperation, or shading in the language of Hart and Moore (2008), with consequent deadweight losses.

In our model, what puts a party below his reference payoff is the occurrence of the abnormal state A. We should emphasize that we suppose that the event "Eve breaks her finger" is exogenous and not something that either party could have controlled or prepared for. Thus, we take the view that neither party will blame the other for the fact that A has occurred. However, each party expects the other party to behave in a fair manner in such an event and will shade if he feels that the other party does not.

Our basic position is that absent communication at date 0 the parties may have different views of what is fair. In Section 5 we will discuss how communication can be used to align the parties' views or preferences.

It is useful to compare our approach with that of Hart and Moore (2008). In both papers contracts are reference points. However, in Hart and Moore, each party is aggrieved and shades if he does not receive the best outcome allowed by the contract. In contrast, in our model, a party is aggrieved and shades only if he is below his reference payoff.

We model shading as in Hart and Moore (2008). We suppose that a party who feels entitled to a payoff s but receives a payoff $s' < s$, will be aggrieved by $s - s'$ and will retaliate by shading on performance in such a way that the other party's payoff falls by $\theta(s - s')$, where $0 < \theta < 1$ is an exogenous parameter ("negative reciprocity"). The party doing the shading neither gains nor loses from shading: the only effect is on the party who is the recipient of the shading. Shading is noncontractible.

In our context, examples of shading might be S refusing to play an encore, being rude to customers, or turning up late; or B not providing beer for the band members during the interval, not cleaning or heating the changing room adequately, or paying late[1].

In what follows the contract will continue to provide S with first-best effort incentives and so we will set $e = e_{FB}$ and write $c(\ e_{FB}) = c$.

Since A is an unusual state, the parties' reference payoffs will be weighted toward their payoffs in the normal state N. To simplify suppose that the parties' reference payoffs *equal* their ex-post payoffs in state N:

(3.1) B's reference payoff $= \ v - p$, S's reference payoff $= \ p - c$.

(We relax this assumption below.) Then if A occurs both parties are below their reference payoffs, whether Adam or George replaces Eve.

It will be useful to conduct our discussion using the numerical example at the end of the last section. In that example, $\Delta c_g = 6$, $\Delta v_g = -14$, $\Delta c_a = 8$, $\Delta v_a \ = -10$. According to the simple contract, S threatens to replace Eve with George since he is cheaper, but the parties will renegotiate to Adam. The selection of George makes B's payoff $v - p - 14$ and S's $p - c - 6$. The gain in efficiency from moving from George to Adam is 2, which the parties split 50:50, and so B's final payoff equals $v - p - 13$, and S's equals $p - c - 5$. So B ends up 13 below his reference point, and S 5 below her reference point.

This final allocation of the losses from Eve breaking her finger is quite asymmetric. B is likely to feel that S should have chosen Adam in the first place. This would have put B 10 below his reference point and S 8 below hers. Suppose that B does indeed think that this is the appropriate outcome. Then, B will be aggrieved by 3, the difference between the loss he bears (13) and the loss that he thinks he should have borne (10), and will shade by 3θ.

But the situation may be worse than this. B may regard S's threat to hire George as coercive, an attempt to force B to pay an extra amount for an outcome (Adam) to which he feels entitled anyway Thus B may refuse to renegotiate out of principle, leaving the outcome as George. In this case the deadweight loss is

[1] Although some of these actions could in principle be contracted on at least ex-post, in practice they are often discretionary.

$2 + 4\,\theta$, where the first term reflects the fact that George is 2 less efficient than Adam and the second term that B's final payoff $= v - p - 14$, whereas he feels entitled to $v - p - 10$[2].

Another element may be important. The parties may feel entitled to a change in price to make up for their losses. Indeed, this could be true even if George is the only replacement for Eve. B may think that it is unfair that he loses 14 because Eve has broken her finger and S loses only 6. He may feel that S should reduce the price by 4 to equalize matters, and may shade if this does not happen. Recall, however, that Δv and Δc are not verifiable, which may make B's argument for a price change more challenging (v could even be non-monetary if we imagine that B is organizing a concert for his pleasure). There are many situations where people are reluctant to give or even ask for a monetary payment to offset a non-verifiable loss[3]. At the same time, in reality, although not in the model, there may be verifiable aspects of quality. Perhaps Eve has received glowing reviews or won an award whereas George has not, and B can use this as an argument for a price change.

S faces a similar challenge concerning a price increase. Reverse the numbers, so that George costs S 14 more than Eve and reduces B's value by 6. S may feel that the price should rise so that B shares some of the losses. B may feel differently, in particular that cost increases are S's problem. B may be particularly unwilling to accept a price increase given that S's increased costs are not verifiable. In this situation, if there is no price change, S may be aggrieved and may shade.

There are many forces at play here, all of which seem realistic and interesting. To make progress we focus on only a few of them. First, in the rest of this section, we develop a benchmark model for the case where the parties have not communicated in such a way as to align their views of fairness (communication is the subject of Section 5). Second, we suppose that a party below his (or her) reference point has an extreme self-serving bias (as in Hart and Moore (2008)): he feels entitled to his maximum payoff under the contract, and is aggrieved and shades if the other party does not enable this to happen. However, we suppose that neither party expects a price change. The one exception is if B needs to pay S to choose an action other than her preferred one. In this case B will make the minimum payment necessary (he cannot get away with less and is not expected to pay more). Finally, we assume that B is willing to renegotiate: he does not regard this as coercive.

One implication of the above is that there is no aggrievement or shading in the case where only George, say, is available to replace Eve (neither party expects a price change). A second implication is that, if Adam and George are both available, B will be willing to pay exactly 2 to persuade S to choose Adam rather than George (he cannot get away with less and is not expected to pay more).

---

[2] There is a countervailing force. Given B's shading, S's final payoff under George is $p - c - 6 - 4\theta$, whereas if she picked Adam her payoff would be $p - c - 8$ since B will not shade in this case. Thus if $\theta > 1/2$, S will choose Adam. However, in this case, S may feel that she was "forced" to choose Adam. Given that her payoff is $p - c - 8$ under Adam, but would have been $p - c - 6$ under George, she will shade by $2\,\theta$. For more on this, see Hart and Moore (2008, Section 4).

[3] See, e.g., Roth (2007).

Using these simplifications, let us move beyond the numerical example to the general case presented in Section 2. For the moment we stick with the idea that the parties' reference points are given by their payoffs in the normal state. Consider first the case where Adam is more efficient than George.

Case 1: $\Delta v_a - \Delta c_a > \Delta v_g - \Delta c_g$

In state A, since S is below her reference payoff, she feels entitled to George since he is cheaper. B is also below his reference point and so feels entitled to Adam since he is the better trumpeter. S has the right to choose and so absent renegotiation will pick George as long as the gain from picking George $\Delta c_a - \Delta c_g$ exceeds the shading costs imposed by B as a result of this choice, $\theta(\Delta v_a - \Delta v_g)$. However, George is inefficient and so renegotiation will take place. B will offer S $\Delta c_a - \Delta c_g$ to switch to Adam, which S will accept. B is aggrieved about having to pay $\Delta c_a - \Delta c_g$ since he thinks that S should have chosen Adam in the first place.

The bottom line is that the efficient outcome Adam will occur but B will be aggrieved and will shade by $\theta(\Delta c_a - \Delta c_g)$.

Note that if $\Delta c_a - \Delta c_g < \theta(\Delta v_a - \Delta v_g)$, that is, the gain from picking George is less than the shading costs imposed by B, S will pick Adam even without renegotiation. However, S will be aggrieved that she has been forced into this choice and will shade by $\theta(\Delta c_a - \Delta c_g)$ (see also footnote 2). Hence, in this situation too, shading costs equal $\theta(\Delta c_a - \Delta c_g)$.

For the rest of the paper we will simplify matters by supposing that $\Delta c_a - \Delta c_g > \theta(\Delta v_a - \Delta v_g)$.

Case 2: $\Delta v_a - \Delta c_a < \Delta v_g - \Delta c_g$

Now S chooses George and there is no renegotiation since George is efficient. B feels entitled to Adam and is aggrieved by $\Delta v_a - \Delta v_g$. Shading equals $\theta(\Delta v_a - \Delta v_g)$.

We can combine Cases 1 and 2:

(3.2) In state A, aggrievement equals $Min(\Delta c_a - \Delta c_g, \Delta v_a - \Delta v_g)$, shading equals $\theta Min(\Delta c_a - \Delta c_g, \Delta v_a - \Delta v_g)$, B's payoff $= v - p + Max(\Delta v_g, \Delta v_a - \Delta c_a + \Delta c_g)$ and S's ex-post payoff $= p - c - \Delta c_g - \theta Min(\Delta c_a - \Delta c_g, \Delta v_a - \Delta v_g)$.

In our numerical example, aggrievement $= 2$, shading $= 2\theta$, B's payoff $= 150 - p - 12$, S's payoff $= p - 20 - 6 - 2\theta$.

We can use (3.2) to compute overall expected payoffs of the parties. S's expected payoff net of effort costs equals

(3.3) $p - c(e) - e - \pi \Delta c_g - \pi\theta Min(\Delta c_a - \Delta c_g, \Delta v_a - \Delta v_g)$.

Since S maximizes (3.3), we can confirm that S will choose $e = e_{FB}$.

The price p adjusts so that S's expected payment equals $\overline{U}$. Thus

(3.4) $p - c(e_{FB}) - e_{FB} - \pi\Delta c_g - \pi\theta Min(\Delta c_a - \Delta c_g,\ \Delta v_a - \Delta v_g) = \overline{U}$.

Finally, B's expected payoff equals

(3.5) $v - p + \pi Max(\Delta v_g, \Delta v_a - \Delta c_a + \Delta c_g) =$

$$v - c(e_{FB}) - e_{FB} + \pi\ Max\left(\Delta v_a - \Delta c_a\ ,\ \Delta v_g - \Delta\ c_g\right) - \pi\theta Min(\Delta c_a - \Delta c_g,\ \Delta v_a - \Delta v_g) - \ \overline{U}.$$

Expected deadweight losses are $\pi\theta\min(\Delta c_a - \Delta c_g,\ \Delta v_a - \Delta v_g)$.

In our numerical example, B's payoff equals $90 - 18\,\pi - 2\,\pi\theta - \overline{U}$, S's payoff $= \overline{U}$, and expected deadweight losses $= 2\,\pi\theta$.

Let us now reconsider the issue of reference payoffs. Since the parties at date 0 have rational expectations they know that the adverse state can occur and so it seems strong to suppose that their reference payoffs do not incorporate this information. In the spirit of Kőszegi and Rabin (2006), let us therefore now suppose that each party's reference payoff is given by his or her *expected* payoff[4].

Note that since each party does better in state N than in state A, his payoff in state A will still be below his reference payoff (which is the average of the two payoffs). Thus, it cannot be an equilibrium for S to pick Adam (B's choice) directly in state A (since she is below her reference payoff she will want to move in the direction of George)[5]. So in equilibrium S will pick George in the first instance, although, in Case 1, B will pay her to pick Adam instead, as above.

Thus, having reference points given by expected payoffs does not change the final outcome. What does change is B's aggrievement. B's reference payoff is given by

(3.6) $v - p + \pi Max(\Delta v_g, \Delta v_a - \Delta c_a + \Delta c_g)$,

while his payoff in state A equals

(3.7) $v - p + \ Max(\Delta v_g, \Delta v_a - \Delta c_a + \Delta c_g)$.

Thus, in state A he is below his reference payoff by

(3.8) $-(1 - \pi)Max(\Delta v_g, \Delta v_a - \Delta c_a + \Delta c_g)$.

[4] Of course, a critical question is why each party's reference payoff in state A does not equal his payoff in state A. In this case there would no aggrievement or shading and the first-best would be achieved. It is a basic assumption of our analysis that a perfect state-contingent adjustment of reference points is (psychologically) infeasible.

[5] To make this argument precise, we need to allow convex combinations of Adam and George. Doing so does not change the analysis.

Since B feels aggrieved only when he is below his reference point, B's aggrievement is capped by (3.8). Hence B's aggrievement, which was previously $Min(\Delta c_a - \Delta c_g, \Delta v_a - \Delta v_g)$, is now given by

(3.9) $Min(\Delta c_a - \Delta c_g, \Delta v_a - \Delta v_g, -(1-\pi)Max(\Delta v_g, \Delta v_a - \Delta c_a + \Delta c_g))$,

and shading equals

(3.10) $\theta Min(\Delta c_a - \Delta c_g, \Delta v_a - \Delta v_g, -(1-\pi)Max(\Delta v_g, \Delta v_a - \Delta c_a + \Delta c_g))$[6].

We may now compute the overall expected payoffs of the parties. S's expected payoff net of effort costs is given by

(3.11) $p - c(e) - e - \pi\Delta c_g - \pi\,\theta Min(\Delta c_a - \Delta c_g, \Delta v_a - \Delta v_g, -(1-\pi)Max(\Delta v_g, \Delta v_a - \Delta c_a + \Delta c_g))$.

As before, S will choose $e = e_{FB}$.

The price p adjusts so that S's expected payment equals $\overline{U}$. Thus

(3.12) $p - c(e_{FB}) - e_{FB} - \pi\Delta c_g - \pi\,\theta Min(\Delta c_a - \Delta c_g, \Delta v_a - \Delta v_g, -(1-\pi)Max(\Delta v_g, \Delta v_a - \Delta c_a + \Delta c_g)) = \overline{U}$.

Finally, B's expected payoff equals

(3.13) $v - p + \pi Max(\Delta v_g, \Delta v_a - \Delta c_a + \Delta c_g) =$

$$v - c(e_{FB}) - e_{FB} + \pi\, Max\,(\Delta v_a - \Delta c_a,\ \Delta v_g - \Delta\, c_g) - \pi\theta \mathrm{Min}(\Delta c_a - \Delta c_g, \Delta v_a - \Delta v_g, -(1-\pi)Max(\Delta v_g, \Delta v_a - \Delta c_a + \Delta c_g)) - \overline{U}.$$

In our numerical example, B's payoff $= 90 - 18\,\pi - \pi\theta Min(2{,}12(1-\pi)) - \overline{U}$, while S's payoff equals $\overline{U}$. The expected deadweight losses are $\pi\theta Min(2{,}12(1-\pi))$.

One useful observation is that, as long as π is not too high, the more sophisticated treatment of reference points does not change anything. In our numerical example, deadweight losses will still be $2\pi\theta$ as long as $\pi < 5/6$. More generally, deadweight losses will be $\pi\theta Min(\Delta c_a - \Delta c_g, \Delta v_a - \Delta v_g)$ as long as

[6] This leads to the intuitive conclusion that if π is close to 1, shading in state A will be very low: basically, the abnormal state becomes the reference point. It can be checked that expected shading (shading weighted by the probability π) is also low when π is close to 0. Indeed, expected shading equals zero when $\pi = 0$ or 1.

(3.14) $\pi < -\Delta v_a / \, Max(-\Delta v_g, -\Delta v_a + \Delta c_a - \Delta c_g)$.

We may summarize this section as follows. We argue that the classical solution does not work in practice because the parties do not adjust their reference points in the adverse state. As a result, B is aggrieved and shades. From (3.13), we see that, if reference points are given by expected payoffs, there are deadweight losses equal to

(3.15) $\pi\theta Min\left(\Delta c_a - \Delta c_g, \Delta v_a - \Delta v_g, -(1-\pi)Max\left(\Delta v_g, \Delta v_a - \Delta c_a + \Delta c_g\right)\right)$.

Note that these deadweight losses can be large. Suppose that $\Delta v_a - \Delta c_a = \Delta v_g - \Delta c_g$=$-\alpha$, where α>0. Then if $\Delta v_g = -\alpha, \Delta c_g = 0, \Delta v_a = 0, \Delta c_a = \alpha$(thus making conflicts of interest large), it is easy to see that the expression in (3.15) equals α/4 if $\pi = \frac{1}{2}$ and $\theta = 1$, and this can be as high as (1/4) $(v - c(0))$ (recall that we have assumed that $v - c(0) + Max(\Delta v_a - \Delta c_a, \ \Delta v_g - \Delta c_g) > 0$).

In Section 5 we will discuss how these deadweight losses can be mitigated.

4. Why the lawyers' traditional approach is unlikely to work here either

A contracting lawyer approaching the situation between B and S has a different perspective from the economist. Typically, the lawyer would be engaged by either B or S to protect their interests in the negotiation and formulation of the written contract. But as we will see, in practice, many of the problems will be the same. And not least, the conclusion is the same that a contract based on classical assumptions will not work.

For simplicity, we assume that B uses a lawyer to negotiate and draft the contract with S. B would have some more or less formal business case in mind and would want the contract to be written to maximize the likelihood of this business case working out. The business case would involve expected revenues, v, and a price paid to S, p. B would also have expectations as regards the service provided by S, i.e., a certain kind of music (rock, pop, jazz, classical, etc.) and a certain quality. Both will be important for the expected revenues since playing jazz in a rock club will decrease revenues and using poor trumpeters in a jazz club will have the same effect.

An important factor for B and his lawyer is *time*. If the concert is at date 1, it will be harder for B to find a band that fits B's needs and profit expectations the closer date 1 becomes. In other words, the more B will be in a hurry, the lower will be his bargaining power. But time will be important for S as well. If S has agreed to play at date 1, she will have said no to other bookings on that date and as date 1 draws closer, it will be harder for S to find another remunerative gig. B and B's lawyer will know this.

Given the above, B's lawyer has to write a contract including three basic components to protect B's business case: a description of the services and their quality, the price, and incentives for S to comply with her contractual obligations. A simple contract could include a brief description of the band service, for example that S shall play high quality jazz from the traditional repertoire at date 1 with a band of, say, six members at a price p. The contract can be elaborated on a great deal, but the difficulty in

writing contractual obligations on how to perform music cannot be completely overcome. As regards the incentive mechanisms, there are two that B can use, both involving forms of power. First, B can induce S to perform through state power, i.e., the threat of enforcing the contract in court, either before the concert or after the concert by claiming damages from S for failure to perform. Second, B can include a right to terminate the contract in advance, regardless of whether a breach has occurred, i.e., a termination for convenience clause. By using the threat to terminate the contract and go to the market to find another band, B can induce S to comply with the contract. This second form of power we therefore call market power.

Now suppose state A occurs and Eve has broken her finger, and B and S negotiate whether George or Adam should be used. Assume for simplicity that Adam is efficient (as in our numerical example). S is entitled to choose, having the residual control rights, but she will have to consider B's reaction. B is, through the contract, entitled to a certain quality and price. George is less expensive for S but a worse player than Adam. Here, the weakness of the traditional contract appears. The termination for convenience clause, entitling B to terminate the contract, will have low influence on S since S knows that B most likely will not terminate close to date 1 given that it is very hard for B to find alternative bands at that point. So B's market power is low. B and S could, of course, use a "knife-edge" contract. That is, they could choose the price p so that B is indifferent between having Adam and not having a concert at all in state A ($v - p + \Delta v_a = 0$). In this case a termination for convenience clause would cause S to choose Adam. But notice that the slightest uncertainty about v would lead to inefficient terminations under these conditions[7].

As regards B's state power, B could of course use it to reject a worse trumpeter than Eve. But B's revenues v are only observable and not verifiable. B would therefore have a great difficulty in proving a cause-and-effect relationship between S's choosing George and lower revenues for B. In effect, B's state power is low as well and, therefore, S's incentive to choose Adam is low since the risks from choosing George are low.

Hence, in this scenario, the contract will probably not stop S from choosing the less efficient George.

The underlying problem is, of course, that the contract is incomplete. It is not possible to specify with accuracy the quality of the services, nor to prove the relationship between performance quality and revenues. And the parties' shading behavior cannot be contracted on as such. This means that the parties are not, with a traditional contract, prepared to deal with uncontracted-for contingencies such as Eve breaking her finger. In short, they have no good mechanisms to deal with the unexpected.

5. How communication leading to the adoption of guiding principles can help

B's lawyer could, however, approach the situation from another perspective, using an alternative approach to the contract. While still being important to safeguard investments (in this case, the seller's

[7] For more on the fragility of knife-edge contracts, see Hart and Moore (2008, Section 4).

choice of effort, e), the contract can be written to ensure alignment of reference points, thereby avoiding shading behavior. As a matter of fact, in many situations, such alignment is the best and maybe the only way to ensure good performance. As the already simple example above shows, both state and market power (as we defined these terms) are often weak incentives for performance, forcing the parties to rely on other mechanisms.

B and S could write a contract that specifies the service, quality and price but that also include a number of guiding principles of fairness that the parties commit to apply in case of unexpected events. For simplicity, in this section, we will focus on one such principle, a principle of loyalty. In our general discussion in Section 7, however, we will consider also the role of other principles.

By a principle of loyalty, we mean a principle that obliges each party to treat the other party's interest as having a comparable importance to their own. Applying this principle could mean that a party should refrain from taking an action that costs the other party more than it benefits the first party. It could also mean that a party should bear a risk if that party is in the best position to avoid or mitigate it.

The loyalty principle, as we use this term, is a widely shared social norm, which can be 'activated' to alter the parties' payoff preferences. To activate such a principle will require an ex-ante communication process and to use it in case of unexpected events ex-post will similarly require communication. Ex-ante, this can be done in practice by the parties entering into discussions about what the principle means, how it applies in different private and professional situations, and whether using the principle would be beneficial for B and S in their relationship. Once the principle is 'activated', it can be incorporated into the formal contract as well, which is important in order to make it part of the parties' reference points.

Assume that B and S agree on a band service and price as before; but they also agree that, in case of unexpected events not dealt with under the contract, they will meet, discuss and apply the principle of loyalty (as defined above). What would happen, under these conditions, if Eve breaks her finger and George and Adam are the available alternatives?

As noted above, the loyalty principle asks each party to treat the other party's interest as having a comparable importance to their own. One way to formalize this is to suppose that each party's payoff becomes their own private payoff plus $\lambda \text{ times the}$ other party's payoff, where $0 < \lambda < 1$. Note that under these conditions S will not be prepared to reduce the price (unless $\lambda = 1$) since she prefers a dollar in her pocket to a dollar in B's, and for the same reason B will not agree to a price increase. Consistent with this, we maintain our assumption that neither party expects or feels entitled to a price change.

At the same time each party will be prepared to accept a one dollar loss if this increases the other party's payoff by more than $1/\lambda$ dollars. Consider again S's decision about whether to replace Eve with George or Adam. S will now internalize the effect on B by putting weight λ on B's payoff. Thus if $6 + 14\lambda > 8 + 10\lambda$, i.e., $\lambda > 1/2$, S will choose Adam. In this case loyalty achieves the efficient outcome. On the other hand, if $\lambda < 1/2$, then S will be inclined to choose George, but B will pay S to choose Adam. How much does he have to pay? Just enough for S to be indifferent: an amount m such that $6 +$

$14\,\lambda\, = 8 - m + (10 + m)\,\lambda$, that is, $m = (2 - 4\,\lambda\,)/(1 - \lambda)$. According to B's evaluation, S should have chosen Adam in the first place since $6\,\lambda\, + 14\, > 8\,\lambda\, + 10$. Since B puts weight λ on S's payoff, paying m to S causes B to be aggrieved by $m(1 - \lambda)$= $(2 - 4\,\lambda)$. B will therefore shade by $(2 - 4\,\lambda)\theta$. Note that, even though shading occurs, it is less than in the no loyalty case: $2 - 4\,\lambda < 2$.

We now provide some formal analysis that goes beyond our numerical example. As above, we will assume that communication leads each party to put weight λ on the payoff of the other party, where $0 < \lambda < 1$. So now in state A,

(5.1) B's ex-post payoff $= v + \Delta v - p + \lambda(p - c - \Delta c)$,

(5.2) S's ex-post payoff $= p - c - \Delta c + \lambda(v + \Delta v - p)$.

Communication is, of course, not costless. We suppose that achieving the weight λ costs an amount (in time and energy) at date 0 equal to $g(\lambda)$ where $g(0) = 0,\ g' \geq 0,\ g'' > 0, g'(0) = 0,\ \lim_{\lambda \to 1} g\ (\lambda) = \infty$. (There could be a fixed cost of communication but for simplicity we do not consider this.) The assumption that $\lim_{\lambda \to 1} g\,(\lambda) = \infty$ captures the idea that it is prohibitively costly to make a party fully internalize another party's preferences. For simplicity, assume that the communication cost is borne entirely by B.

We will also suppose that the level of communication, and hence the choice of λ, can be specified in the date 0 contract, and that loyalty preferences are not activated until after communication takes place. Hence each party evaluates the contract according to their pre-loyalty preferences.

As we have emphasized above, in practice, aligning preferences is also likely to require communication (and the incurring of communication costs) ex-post at date 1. For simplicity, we ignore this.

Start by assuming that reference payoffs are payoffs in the normal state. Consider state A. Suppose that Adam is more socially efficient than George. S's inclination previously was to choose George, while B wanted Adam. Now, however, with the payoff in (5.2), S will choose Adam directly as long as

(5.3) $\lambda\Delta v_a - \Delta c_a > \lambda\Delta v_g - \Delta\, c_g$,

that is,

(5.4) $\lambda > (\Delta c_a - \Delta\, c_g)/\left(\Delta v_a - \Delta v_g\right)$.

Under these conditions, renegotiation will not be required and there will be no aggrievement by B. On the other hand, if (5.4) is not satisfied, renegotiation will take place. B will have to pay S an amount m where $\lambda(\Delta v_a - m) - (\Delta c_a - m) = \lambda\Delta v_g - \Delta\, c_g$, that is, $m = \ (\lambda\Delta v_g - \Delta\, c_g - \ (\lambda\Delta v_a - \Delta c_a))/(1 - \lambda)$. Since B puts weight λ on S's payoff, paying m to S causes B to be aggrieved by $m(1 - \lambda) = (\lambda\Delta v_g - \Delta\, c_g - \ (\lambda\Delta v_a - \Delta c_a))$ and B will shade by $\theta(\lambda\Delta v_g - \Delta\, c_g - (\lambda\Delta v_a - \Delta c_a))$.

We can combine the cases where (5.4) does and does not hold: shading in state A equals

(5.5) $\theta\, Max(\lambda\Delta v_g - \Delta c_g - (\lambda\Delta v_a - \Delta c_a), 0)$.

Continuing with the case where Adam is more efficient than George, let us now (as in Section 3) allow for reference payoffs to equal expected payoffs over the normal and abnormal states rather than payoffs in the normal state.

B's ex-post payoff in state A equals

(5.6) $v - p + \Delta v_a - Max\big(\lambda\Delta v_g - \Delta c_g - (\lambda\Delta v_a - \Delta c_a), 0\big) + \lambda(p - c - \Delta c_a)$,

and so his expected ex-post payoff is

(5.7) $v - p + \lambda(p - c) + \pi\ (\Delta v_a - \lambda\Delta c_a - Max(\lambda\Delta v_g - \Delta c_g - (\lambda\Delta v_a - \Delta c_a), 0))$.

It follows that in state A B is below his reference payoff by the amount

(5.8) $-(1 - \pi)\ (\Delta v_a - \lambda\Delta c_a - Max(\lambda\Delta v_g - \Delta c_g - (\lambda\Delta v_a - \Delta c_a), 0))$.

Since B's aggrievement is capped by the expression in (5.8), shading in state A is given by the smaller of (5.5) and θ times (5.8):

(5.9) $\theta\, Min[Max(\lambda\Delta v_g - \Delta c_g - (\lambda\Delta v_a - \Delta c_a), 0), -(1 - \pi)\ (\Delta v_a - \lambda\Delta c_a - Max(\lambda\Delta v_g - \Delta c_g - (\lambda\Delta v_a - \Delta c_a), 0))]$,

and expected shading is given by

(5.10) $\pi\theta Min\ [Max(\lambda\Delta v_g - \Delta c_g - (\lambda\Delta v_a - \Delta c_a), 0), -(1 - \pi)\ (\Delta v_a - \lambda\Delta c_a - Max(\lambda\Delta v_g - \Delta c_g - (\lambda\Delta v_a - \Delta c_a), 0))]$.

It is useful to compute ex-ante payoffs. As we have noted, each party evaluates the contract according to their pre-loyalty preferences. Thus , if B anticipates paying m to S at date 1, this gets full weight rather than weight $(1 - \lambda)$, and similarly for S . Hence

(5.11) B's expected payoff= $v - p + \pi\ (\Delta v_a - (\frac{1}{1-\lambda})[Max\big(\lambda\Delta v_g - \Delta c_g - (\lambda\Delta v_a - \Delta c_a), 0\big)]) - g(\lambda)$,

and

(5.12) S's expected payoff= $p - c(e) - e - \pi(\Delta c_a - (\frac{1}{1-\lambda})[Max\big(\lambda\Delta v_g - \Delta c_g - (\lambda\Delta v_a - \Delta c_a), 0\big)])$

$$-\pi\theta Mi\,n\Big[Max\big(\lambda\Delta v_g - \Delta c_g - (\lambda\Delta v_a - \Delta c_a), 0\big), -(1 - \pi)\Big(\Delta v_a - \lambda\Delta c_a - Max\big(\lambda\Delta v_g - \Delta c_g - (\lambda\Delta v_a - \Delta c_a), 0\big)\Big)\Big],$$

from which we can confirm that S chooses $e = e_{FB}$. In a competitive market for sellers, p will adjust so that the right-hand side of (5.12) equals $\overline{U}$.

Suppose in contrast that George is more efficient than Adam. Again, start with the case where reference payoffs are payoffs in the normal state. S will choose George as before, and the only question is whether B is aggrieved. B will not be aggrieved and will not shade as long as B thinks that George is the right choice, that is, according to (5.1), as long as

(5.13) $\Delta v_g - \lambda\Delta c_g > \Delta v_a - \lambda\Delta c_a$,

which we can rewrite as

(5.14) $\lambda > (\Delta v_a - \Delta v_g)/(\Delta c_a - \Delta c_g)$.

Even if (5.14) is not satisfied, B's shading will be reduced. The difference between B's ex-post payoff under Adam and George equals

(5.15) $\Delta v_a - \lambda\Delta c_a - (\Delta v_g - \lambda\Delta c_g)$,

which is decreasing in λ. We can combine the cases where (5.14) does and does not hold: if reference payoffs are given by payoffs in the normal state, shading equals

(5.16) $\theta Max(\Delta v_a - \lambda\Delta c_a - (\Delta v_g - \lambda\Delta c_g),0)$.

Now incorporate the idea that reference payoffs are expected payoffs over the normal and abnormal state. B's ex-post payoff in state A equals

(5.17) $v - p + \Delta v_g - \lambda\Delta c_g$,

and so his expected ex-post payoff equals

(5.18) $v - p + \pi(\Delta v_g - \lambda\Delta c_g)$.

It follows that in state A B is below his reference payoff by the amount

(5.19) $-(1-\pi)(\Delta v_g - \lambda\Delta c_g)$.

Since B's aggrievement in state A is capped by the expression in (5.19), shading in state A is given by the smaller of (5.16) and θ times (5.19):

(5.20) $\theta\min\left[Max(\Delta v_a - \lambda\Delta c_a - (\Delta v_g - \lambda\Delta c_g),0), -(1-\pi))(\Delta v_g - \lambda\Delta c_g)\right]$

and expected shading by

(5.21) $\pi\theta\mathrm{Min}[Max(\Delta v_a - \lambda\Delta c_a - (\Delta v_g - \lambda\Delta c_g),0), -(1-\pi))(\Delta v_g - \lambda\Delta c_g)]$.

We leave it to the reader to write down ex-ante payoffs in this case.

So far we have taken λ to be exogenous. Since λ is contractible, in an optimal contract it will be chosen efficiently to minimize the sum of communication and expected shading costs. Thus, if Adam is more efficient than George, λ willbe chosen to minimize (from (5.10))

(5.22) $g(\lambda) + \pi\theta M\text{in}[Max\big(\lambda\Delta v_g - \Delta c_g - (\lambda\Delta v_a - \Delta c_a), 0\big), -(1-\pi)(\Delta v_a - \lambda\Delta c_a - Max(\lambda\Delta v_g - \Delta c_g - (\lambda\Delta v_a - \Delta c_a), 0))]$,

while, if George is more efficient than Adam, λ will be chosen to minimize (from (5.21))

(5.23) $g(\lambda) + \pi\theta Min[Max(\Delta v_a - \lambda\Delta c_a - \big(\Delta v_g - \lambda\Delta c_g\big), 0), -(1-\pi)(\Delta v_g - \lambda\Delta c_g)]$.

As in Section 3, it is easy to show that, when π is small, the second term in the Min expressions of (5.22)-(5.23) is not binding, that is, it makes no difference that reference payoffs are given by expected payoffs rather than normal payoffs. But it is easy to see that the first terms in the Min expressions of (5.22)-(5.23), $Max\ (\lambda\Delta v_g - \Delta c_g - (\lambda\Delta v_a - \Delta c_a), 0)$ and $Max\big(\Delta v_a - \lambda\Delta c_a - \big(\Delta v_g - \lambda\Delta c_g\big), 0\big)$, are decreasing in λ. A standard revealed preference argument yields the implication that for small π the optimal value of λ is increasing in π . Also λ converges to zero as $\pi \to 0$: it does not make sense to engage in significant communication if π is small.

Another case where communication is not valuable is if conflicts of interest are small. Suppose that $\Delta v_a = \Delta v_g$. Then, George is more efficient than Adam, and it is easy to show that the first term in the Min expression of (5.23) is zero for λ positive and so the optimal value of $\lambda = 0$. (We do not consider the case where $\Delta c_a = \Delta c_g$ since we have supposed $\Delta c_a - \Delta c_g > \theta(\Delta v_a - \Delta v_g)$.)

To obtain more insight, it is useful to look at the first-order conditions for (5.22)-(5.23). These yield, respectively, for small π,

(5.24) $g'(\lambda) = \pi\,\theta\,\lambda\big(\Delta v_a - \Delta v_g\big)\ if\ 0 < \lambda < (\Delta c_a - \Delta\, c_g)/\big(\Delta v_a - \Delta v_g\big)$,

(5.25) $g'(\lambda) = \pi\,\theta\,\lambda\big(\Delta c_a - \Delta c_g\big)\ if\ 0 < \lambda < \big(\Delta v_a - \Delta v_g\big)/(\Delta c_a - \Delta\, c_g)$.

We can see from (5.24) that several effects are at play. Suppose that $\Delta v_a - \Delta v_g$ increases, that is, conflicts of interest about quality increase. Then the right-hand side of (5.24) rises, which suggests that λ will rise. However, it is also possible that we are at a corner solution: $\lambda = (\Delta c_a - \Delta\, c_g)/\big(\Delta v_a - \Delta v_g\big)$. In this case λ will fall. The intuition is that an increase in the amount by which B favors Adam makes it more likely that S will choose Adam for a given λ. Thus, it may be possible to reduce λ.

So far we have argued that communication can be a valuable supplement to a contract. An important question to ask is, would it ever make sense to rely on communication and loyalty alone and dispense with a contract altogether?

It is not completely obvious how to analyze the no contract case since we have assumed that reference points and payoffs are determined through the contract. One approach is to follow Hart and Moore

(2008), and assume that absent a contract each party feels entitled to 100% of the ex-post surplus from the transaction. Note that the loyalty principle will not change this since each party prefers a dollar in their pocket to a dollar in the other party's pocket, although other principles, such as equity, which we discuss in Section 7, could be important.

If the parties have equal bargaining power ex-post, they will compromise on a 50:50 split[8]. It follows that each party will be aggrieved by the half of the gains they do not get times $(1-\lambda)$ and so will shade by θ times this amount. In the normal state this means that the deadweight losses from shading will be $\theta(1-\lambda)\ (v-c)$, where c depends on S's choice of effort. Modifying (2.14), we can see that S will choose her effort to maximize:

(5.26) $\left(\frac{1}{2}v-\frac{1}{2}\ c(e)\right)(1+\lambda)(1-\theta)+\frac{1}{2}\ \pi(1+\lambda)(1-\theta)\ Max\left(\Delta v_a-\Delta c_a,\ \Delta v_g-\Delta\ c_g\right)-e,$

yielding

(5.27) $\frac{1}{2}(1+\lambda)(1-\theta)c'(e)=-1.$

Obviously, compared to the case where loyalty is combined with a contract, we have under-investment. But more than this, shading costs are greater. To see this observe that expected shading costs without a contract are given by

(5.28) $L_{NC}=(1-\pi)(1-\lambda)\ \theta(v-c)+\pi(1-\lambda)\ \theta(v-c+Max\left(\Delta v_a-\Delta c_a,\ \Delta v_g-\Delta\ c_g\right))$

$$=\theta(1-\lambda)[(v-c)+\pi\ \ Max\left(\Delta v_a-\Delta c_a,\ \Delta v_g-\Delta\ c_g\right)],$$

where c is evaluated at the e that solves (5.27).

Proposition 1. Assume $v+\Delta v_g > c+\Delta c_a$. Then $L_{NC}>L_C$.

Proof:

Fix $\lambda<1$. Let $S_{NC}=\frac{L_{NC}}{\theta}$ and $S_C=\frac{L_C}{\theta}$ denote shading under a contract and no contract, respectively. It suffices to show that $S_{NC}>S_C$.

Case 1: $\Delta v_a-\Delta c_a>\Delta v_g-\Delta c_g$

From (5.28),

[8] Experiments by Ellingsen and Johannesson (2004) suggest that a fair-minded B might be willing to grant S more than 50% of the ex-post surplus to compensate S for her effort investment. In our setting this effect is likely to be mitigated since B does not observe S's effort.

(5.29) $S_{NC} = (1-\lambda)[(v-\mathrm{c}) + \pi\ (\Delta v_a - \Delta c_a)]$,

which is greater than

(5.30) $(1-\lambda)\pi\ [(v-\mathrm{c}) +\ (\Delta v_a - \Delta c_a)]$.

Also, from (5.10),

(5.31) $S_C \leq \pi Max(\lambda\Delta v_g - \Delta c_g - (\ \lambda\Delta v_a -\ \Delta c_a), 0)$.

Hence, since the expression in (5.30) is positive, it suffices to show (dividing by π) that

(5.32) $(1-\lambda)\ [(v-\mathrm{c}) +\ \ (\Delta v_a - \Delta c_a\ )] >\ \lambda\Delta v_g - \Delta c_g - (\ \lambda\Delta v_a -\ \Delta c_a)$.

But we have assumed $v + \Delta v_g > c + \Delta c_a$. From this it follows that

(5.33) $v + \Delta v_a - \mathrm{c} - \Delta c_a > \Delta c_a - \Delta c_g$.

But it can easily be seen that this implies (5.32).

The argument in Case 2, where $\Delta v_a - \Delta c_a\ \leq \Delta v_g - \Delta c_g$, is analogous. Q.E.D.

Note that to prove the result we need an extra assumption, borrowed from Hart and Moore (2008, Section 4), that the value-cost intervals for Adam and George in the abnormal state overlap: $v + \Delta v_g > c + \Delta c_a$.

The conclusion is that no contract plus loyalty provides less effort *and* higher deadweight costs than a contract plus loyalty and so is dominated by it.

It is worth drawing out a further implication. The no communication-no loyalty model of Section 3 is a special case where $\lambda = 0$. Thus the argument in the proof shows that in that world a contract is better than no contract.

6. Employment

So far we have supposed that S has residual rights of control and bears all the costs $c + \Delta c$. We can think of this as non-integration. Another possibility would be for B to acquire S's operations; B would then possess residual rights of control. We can think of this as vertical integration or employment. In this case it is natural to suppose that B bears 100% of the costs: S is on a fixed wage.

Of course, under both non-integration and vertical integration, cost-sharing is feasible since costs are verifiable. For simplicity we stick to the polar cases where S bears all the costs under non-integration and B bears all the costs under integration.

Under employment there is no conflict of interest ex-post since B bears both the value and cost consequences of replacing Eve. However, under standard assumptions, S has no incentive to work since she bears none of the ex-post costs: $e = 0$. Thus, under the classical assumptions of Section 2, employment (with B bearing all the costs) would never be optimal (recall that we saw in Section 2 that under classical assumptions a simple non-integration contract achieves the first-best).

However, employment can be attractive if the loyalty principle is adopted. Now S will choose e to maximize

(6.1) $\quad p - e + \lambda(v - p + \pi\, Max\big(\Delta v_a - \Delta c_a,\ \Delta v_g - \Delta c_g\big) - c(e))$.

The first-order conditions are

(6.2) $\quad \lambda c'(e) = -1$.

Denote the solution by $e(\lambda)$.

The expected net surplus under employment equals

(6.3) $\quad v - c\big(e(\lambda)\big) - e(\lambda) + \pi\, Max\big(\Delta v_a - \Delta c_a,\ \Delta v_g - \Delta c_g\big) - g(\lambda)$,

where recall that there are no shading costs since there are no ex-post conflicts of interest.

From (6.3) it follows that under employment it is optimal to choose the level of communication λ to minimize

(6.4) $\quad g(\lambda) + c(e(\lambda)) + e(\lambda)$.

Let λ* be the solution of (6.4). Then the total deadweight costs under employment are

(6.5) $\quad g(\lambda*) + c(e(\lambda*)) + e(\lambda*) - c(e_{FB}) - e_{FB}$.

To see whether employment is better than non-integration we must compare (6.5) with the optimized expressions in (5.22) and (5.23). Some simple conclusions are immediate. Suppose that S's effort does not matter much: $c(e_{FB}) + e_{FB}$ is close to $c(0)$. Then employment with no communication $(\lambda = 0)$ achieves approximately the first-best (whereas non-integration typically does not). On the other hand, suppose that ex-post conflicts of interest are small: Δv_g is approximately equal to Δv_a. Then non-integration achieves approximately the first-best (whereas employment typically does not).

Perhaps more interesting, suppose that we replace S's effort cost e by ke and the cost function $c(e)$ by $kc(e)$, where $k > 0$. Then the first-best effort level e_{FB} and the effort under employment $e(\lambda*)$ remain the same (the left-hand and right-hand sides of (2.4) and (6.2) are both multiplied by k). However, the deadweight cost from an inefficient choice of e is multiplied by k. Thus if k is large the deadweight costs from employment (in (6.5)) become very large, either because $e(\lambda*)$ is bounded away from e_{FB} or because $\lambda*$ is close to 1, in which case $g(\lambda*)$ is large (we have supposed that $\lim_{\lambda\to 1} g(\lambda) = \infty$). Thus, for large k, non-integration will dominate employment.

7. Other Principles

Leaving the formal analysis aside, there are a number of points to discuss about the approach we take. We have focused on the principle of loyalty. But in practice other principles are also very important. One of them is *equity*. By a principle of equity, we mean that costs, benefits and risks should be allocated in proportion to each party's effort (where taking a risk can also be an effort). Applying this principle could mean that the two parties split the benefits of a joint investment in strict proportion to how much each party has contributed to the investment. Equity is therefore not necessarily only about equal splits of revenues or costs. Equity calls for a *proportionate* allocation of benefits, losses, etc.

A principle of equity seems particularly applicable in a situation where, contrary to our maintained assumption, (incremental) value and (incremental) cost are verifiable. In our concert example, equity could come into play in a number of ways. Let us assume, for simplicity, that S and B agree, ex-ante, on only the equity principle. How should the decreased value (Δv) and increased cost (Δc) be allocated? If we assume that it has taken no effort to find and hire Eve's replacement, B and S would most likely agree on a 50:50 split of the losses. In our numerical example, this means that, if George is selected, the parties will reduce the price by 4, so that each bears a loss of 10; while if Adam is selected the price reduction will be 1, so that each bears a loss of 9. Under these conditions there will be unanimity that the more efficient choice, Adam, should be made, and there will be no shading. The first-best is achieved.

If we assume, instead, that finding Adam on short notice required a lot of effort from S, an equitable split would take this into account, allocating a larger proportion of the losses to B.

Equitable splits are unlikely to proceed as smoothly under our assumption that incremental value and cost are observable, but not verifiable, and where subjectivity and self-serving biases can come into play. Yet people are often able to find amicable solutions on how to allocate losses, and compensate for emotional damage and other problems, even if they would not be able to prove their case in court. Thus, the equity principle is probably relevant in our setting too. Incorporating the equity principle into the analysis is an important topic for future research.

There are also other principles that could be taken into consideration, for example a principle of reciprocity, obliging each party to return good actions in kind; or a principle of autonomy, obliging each party not to make threats or be coercive. An autonomy principle would apply not to the allocation of losses, but rather to the *process* for coming to a decision.

In practice, then, B and S could adopt a number of principles ex-ante which would come into play ex-post to assist the parties in achieving outcomes that meet their reference points, thus avoiding shading.

A very important question that our analysis raises is, what is special about principles like loyalty or equity? Why could not the parties use other principles? For example, why couldn't the parties adopt the principle that in an abnormal state they will sort things out using the bargaining protocol underlying the first-best contract described in Section 2 (50:50 bargaining using side-payments). This bargaining

outcome would become the new reference point and neither party would be aggrieved or would shade. The first-best would be achieved. Other possibilities would be that the parties agree that if something unexpected happens B will make a take-it-or leave-it offer to S about how to proceed, or the parties will play a Maskin-Tirole (1999) mechanism to make observable information verifiable.

Our (tentative) answer is that the principles of loyalty and equity are not just ad-hoc principles chosen by the parties. They rest on strong social norms and are thus better described as being 'activated' than 'chosen'. This makes them different from mechanisms such as take-it-or-leave-it offers or Maskin-Tirole revelation games, which have no motivating power in themselves. The motivating power to apply principles such as equity or loyalty does not rest only on the urge to fulfill a promise made. Making a promise to fulfill a social norm has more force than making a promise to apply a principle not based on a social norm, for example to receive a take-it-or-leave-it offer.

There is evidence that social norms affect outcomes in economic transactions. First, it has been convincingly shown in the laboratory and elsewhere that people have not only self-regarding motives but also other-regarding motives.[9] Second, the extent to which they exhibit other-regarding motives will depend on the economic and social context. Fehr and Schmidt (1999) have shown that there is an important interaction between the distribution of preferences in a given population and the strategic environment: in some situations a minority of self-regarding players can hamper collaboration and in other situations a minority of other-regarding players can induce the self-regarding players to cooperate.

Also relevant are studies by Ostrom (1990) and Ellickson (1991). These authors show how groups of people can overcome social dilemmas in situations where people must engage in face-to-face discussions and negotiations on how to solve problems, for example on how to allocate costs and risks. Put in situations where they have to make decisions and argue their case, people are affected by social norms, and are led by those norms to efficient outcomes.

Our analysis in Section 5 rests on the idea that communication can align preferences and induce cooperative behavior. It is a well-established fact that communication, not least face-to-face communication, can in many circumstances improve cooperation and reduce deadweight losses. An area of particular focus has been how communication can mitigate conflicts of interest in social dilemmas, starting with studies by Deutsch (1958, 1960) and Loomis (1959). In a 1995 meta-study, Sally (1995) analyzed over 100 studies and concluded, having tested a number of independent variables, that communication increases cooperation by 40 percent and was the variable having the strongest effect on cooperation. These results were confirmed in a later meta-study by Balliet (2010), who concluded also that face-to-face communication has a stronger effect than written communication.

Not only does communication as such have an effect; the *content* of the messages communicated matters. For example, Charness and Dufwenberg (2006) have shown in the laboratory that a statement of intent or promise can have a particularly strong effect on cooperation. We find this important, since

[9] See Charness and Rabin (2002), Fehr and Schmidt (1999), and Thaler and Dawes (1992).

our model builds on a scenario with ex-ante communication where the parties exchange promises to follow certain guiding principles.

Communication has also been studied within the framework of contracts-as-reference-points. In a study related to Hart and Moore (2008), Brandts et al. (2016) (BCE) tested whether communication affects parties' reference points and thereby shading behavior. As noted by Hart and Moore (2008), an important consequence of the fact that contracts serve as reference points is that there is a tension between contractual rigidity and flexibility. Whereas a flexible contract is generally preferable, it can also lead to increased shading behavior since the flexibility gives more room for conflicting feelings of entitlement. In their experiment, BCE showed that free-form communication significantly reduced shading levels in flexible contracts, making them more profitable for both parties than rigid contracts. In particular, clarification of transfer plans, friendliness and promises helped the subjects align their expectations and resolve ambiguity, thereby reducing shading behavior.

To be sure, we are not claiming that communication is a solution in all situation. Fehr et al. (2015) obtain less optimistic results about communication although they do not allow for free-form communication. Also, in a working paper, Fehr et al. (2017) show that, under certain competitive conditions, communication, rather than being used by buyers for aligning expectations and improving cooperation, was instead abused for the purpose of influencing the activities of the seller. The distinction made by German social philosopher and sociologist Jürgen Habermas between communication in strategic actions, aimed at influencing others, and communicative action, aimed at reaching a common understanding, seems relevant in this context[10]. It seems plausible that the interaction between the distribution of preferences in a given population and the strategic environment, emphasized by Fehr and Schmidt in the context of social preferences, is highly relevant for communication as well. Communication can be a weapon in the pursuit of strictly self-regarding behavior, even though strong evidence also shows that communication can promote and be part of other-regarding behavior.

In summary, both social norms and communication have been shown to improve cooperation and reduce deadweight losses. Based on this, we find it plausible that ex-ante and ex-post communication *about* social norms should have important effects and be capable of aligning reference points. This claim is related to the concept of *principled negotiations*, introduced by Roger Fisher and William Ury (1981) in their classic negotiation book *Getting to Yes,* which was a product of the Harvard Negotiation Project. Among other things, Fisher and Ury recommend that negotiating parties avoid imposing their wills on one another and instead apply standards of fairness, market practice or scientific merit, which both parties recognize as valid. While disagreement is still possible, the parties avoid damaging conflicts by agreeing on common criteria.

Fisher and Ury did not frame their recommendations in the language of reference points. However, it could easily be argued that the reason that following their recommendations has proved to be valuable for so many organizations negotiating agreements is related to reference points: by agreeing on common standards, the parties will have their expectations (reference points) tied to outcomes based

[10] Habermas (1984).

on those standards and even though the actual application of the standard may not in all cases lead to the optimal solution for one party, that party will still not be disappointed since the standard adopted by that party was followed.

Fisher and Ury saw principles of fairness as just one among many possible objective criteria. While we acknowledge that there can be several principles playing a similar role, in this paper we emphasize the importance of principles based on social norms, which have the motivating force discussed above to induce cooperation.

The practical experience of one of us (Frydlinger) from working with several large corporations and their suppliers shows that a critical factor determining how effective communication is in contracting situations concerns the *order* in which matters are communicated. An early focus in the process on price and quality tends to set the contracting parties against one other, inducing self-regarding behavior and increasing friction. If, instead, the parties discuss the hard factors of the contract late in the process and begin, as suggested in this paper, by discussing and 'activating' social norms such as loyalty and equity, the parties tend to adopt a more other-regarding attitude. This helps them to find fair solutions when negotiating hard factors such as price and quality, and when dealing with unexpected events and other matters that could lead to conflicts of interest. Frydlinger has found this approach to work in practice both in situations where there is ex-ante competition and in situations where there is bilateral monopoly.

8. Conclusions

A common explanation for the idea that contracts are incomplete is that there are too many future contingencies for parties to contract on. However, as the literature has noted, even though writing down all the contingencies may be impossible, anticipating the payoff consequences of these contingencies should be feasible and the question then is, why do the parties not contract on these payoff consequences instead of the contingencies themselves, via mechanisms?

Our paper provides an answer to this question, and also proposes an alternative approach to dealing with contractual incompleteness. We have considered a setting where a buyer and a seller write a contract for two reasons. First, the contract provides incentives for the parties to make efficient relationship-specific investments (effort decisions). Second, the contract aligns the parties' reference points so that in normal times there will be no disagreement about what is considered a fair outcome. However, the contract has a minus. If an abnormally bad state occurs, the parties will be below their reference payoffs and this can lead to disagreement about what is fair. The result is that parties will become aggrieved and will withdraw cooperation (shade), causing deadweight losses.

We have suggested that one way to improve the situation is for the parties to commit as part of the contract to apply a number of guiding principles such as loyalty or equity in case of unexpected events, and to build in communication processes that will enable the application of these principles (we have focused on the principle of loyalty). These principles rest on strong social norms and are thus better

described as being 'activated' than 'chosen'. This makes them different from mechanisms such as take-it-or-leave-it offers or revelation games, which have no motivating power in themselves, and which we believe will not work as well[11].

There is obviously an overlap between the ideas that we have presented and those that form the basis of the vast literature on relational contracts in that both highlight the importance of trust and norms (see Malcomson (2013) for a survey). However, the approaches are also importantly different. First, the relational contracting literature considers relationships that are repeated whereas our model is one-shot (although we are not in any way denying the importance of long-term interactions). Second, relational contracting models are plagued by multiple equilibria whereas our model has a unique equilibrium. Third, formal contracts are often a negative in that literature since they can make informal relationships harder to sustain (see Baker et al. (1994)); in contrast, in our model formal contracts and guiding principles are complements. Finally, we stress the importance of communication to align reference points and notions of fairness, something that is not a major feature of the relational contracting literature.

The concert example, which forms the center-piece of our paper, is quite special in the sense that the deal (realistically) has low complexity and a relatively low likelihood of an abnormal state. However, we believe, and the practical experience of one of us supports the idea, that our model is applicable to a broad range of more complex deals and contracts in areas such as outsourcing, supply chain management, public-private partnerships, franchises, application development, and shareholder agreements. To take just one example, outsourcing deals regarding IT often exhibit a great deal of complexity and uncertainty because of the difficulty of completely specifying the customer's needs and because the needs of the customer change over time. Outsourcing contracts are therefore always to a large extent incomplete. Also, to be perceived as successful, outsourcing deals often require the generation of innovative ideas and other forms of cooperative behavior, which are more or less non-contractible. The other side of the coin is that there is a lot of room for shading behavior and thus deadweight losses.

Experience and studies[12] show that traditional contracting practices often fail to address these challenges. For example, customers often complain about lack of proactivity and innovation from the supplier, and suppliers often complain about the customer's micromanaging of the supplier, who in theory is supposed to be the expert on the work. And both customers and suppliers complain about the time and transaction costs spent arguing about what is inside and outside the scope of the contract. Customers often try to deal with the latter problem through so-called scope-sweeper clauses, using contract language that "sweeps in" most possible services as part of the supplier's responsibility, but experience shows that such clauses often fail to work well.

Our model can help to identify where the real problem lies in outsourcing and similar deals, namely in often unclear and misaligned reference points and a lack of effective mechanisms to deal with unexpected events. Based on our model, we would argue that any attempt to deal with these problems through an even more rigid contract with more clauses will often fail and that it is instead more important to address the root cause of the problems, namely the misaligned reference points.

[11] Aghion et al. (2018) find that revelation games do not work well in the lab.
[12] See Deloitte (2016) and Vitasek at al. (2013).

Interestingly enough, people on the customer and supplier sides often show a tendency to ignore the complexity of the deal and the likelihood of unexpected events and instead put their belief in business cases that show overly-optimistic savings, quality increases and margins. Humans' bias towards over-simplification[13] seems to mean that the parties fail to include real-life complexity in their reference points, making them vulnerable to disappointment during the contract term.

In addition to identifying the real problem in many complex deals such as outsourcing, we believe that our model can show a way to deal with this problem. Using a less rigid written contract in conjunction with guiding principles such as loyalty and equity can, we think, be an effective way to keep the parties' reference points met and aligned even in deals where the probability of something going wrong is quite high.[14]

Obviously, although we have argued for our approach both in relatively simple situations such as a concert or in more complex situations such as the outsourcing of IT, much more work needs to be done to validate our ideas. We believe that theory, empirical work, and experiments are all promising directions to pursue to clarify the role of guiding principles in overcoming contractual incompleteness.

[13] See Kahneman (2011).

[14] Our model has the feature that, if the probability of an abnormal state π is close to 1, aggrievement and shading are low because the abnormal state becomes the new reference point (see footnote 6). However, we have supposed that payoffs in the abnormal state are always the same. The conclusion will not hold If there is uncertainty about payoffs in the abnormal state, which is very likely to be the case.

REFERENCES

Aghion, Philippe, Ernst Fehr, Richard Holden, and Tom Wilkening. 2018. The Role of Bounded Rationality and Imperfect Information in Subgame Perfect Implementation—An Empirical Investigation. *Journal of the European Economic Association* 16(1): 232–274.

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

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

Balliet, Daniel. 2010. Communication and Cooperation in Social Dilemmas: A Meta-Analytic Review. *Journal of Conflict Resolution* 54(1): 39-57.

Brandts, Jordi, Matthew Ellman, and Gary Charness. 2016. Let's Talk: How Communication Affects Contract Design. *Journal of the European Economics Association* 14(4): 943-974.

Charness, Gary and Martin Dufwenberg. 2006. Promises and Partnerships. *Econometrica* 74(6): 1579–1601.

Charness, Gary and Rabin, Matthew. 2002.Understanding Social Preferences with Simple Tests. *The Quarterly Journal of Economics* 117(3): 817-869.

Deloitte, Deloitte's 2016 Global Outsourcing Survey, May 2016.

Deutsch, Morton. 1958. Trust and Suspicion. *Journal of Conflict Resolution* 2: 256-79.

Deutsch, Morton. 1960. The Effect of Motivational Orientation Upon Trust and Suspicion. *Human Relations* 13: 123-139.

Ellickson, Robert C. 1991. *Order without Law: How Neighbors Settle Disputes.* Cambridge, Mass: Harvard University Press. Pg.71.

Ellingsen, Tore and Magnus Johannesson. 2004. Is There a Hold-Up Problem? *Scandinavian Journal of Economics* 106(3): 475-494.

Fehr, Ernst, Oliver Hart and Christian Zehnder. 2015. How Do Informal Agreements and Renegotiation Shape Contractual Reference Points? *Journal of the European Economic Association* 13(1): 1-28.

Fehr, Ernst, Oliver Hart, and Christian Zehnder. 2017. Contracts, Conflicts and Communication. Harvard working paper.

Fehr, Ernst and Klaus M. Schmidt. 1999. A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics* 114(3): 817-868.

Fisher, Roger and William Ury. 1981. *Getting to Yes*. Penguin Books.

Grossman, Sanford J. and Oliver Hart. 1986. The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy* 94(4): 691-719.

Habermas, Jürgen. 1984. The Theory of Communicative Action. Volume One: *Reason and Rationalization of society*. Beacon Press.

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

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. Farrar, Straus and Giroux.

Kőszegi, Botond and Matthew Rabin. 2006. A Model of Reference-Dependent Preferences. *Quarterly Journal of Economics* 121(4): 1133-1165.

Loomis, James L. 1959. Communication, the Development of Trust and Cooperative Behavior. *Human Relations* 12(4): 305-315.

Malcomson, James. 2013. Relational Incentive Contracts. In Robert Gibbons and John Roberts (eds.), *The Handbook of Organizational Economics*, chapter 25, pp. 1014-1065. Princeton University Press.

Maskin Eric and Jean Tirole. 1999. Unforeseen Contingencies and Incomplete Contracts. *Review of Economic Studies*. 66: 83-114.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.

Roth, Alvin E. 2007. Repugnance as a Constraint on Markets. *Journal of Economic Perspectives* 21(3): 37-58.

Sally, David. 1995. Conversation and Cooperation in Social Dilemmas: A Meta-analysis of Experiments from 1958 to 1992. *Rationality and Society* 7(1): 58-92.

Thaler, Richard and Robyn Dawes. 1992. Cooperation. In Thaler, Richard. *The Winner's Curse*. Princeton University Press.

Vitasek, Kate, Mike Ledyard, and Karl Manrodt. 2013. *Vested Outsourcing*. Palgrave Macmillan. 2nd edition.