

THE ‘REFERENCE POINT’ APPROACH TO THE THEORY OF THE FIRM: AN INTRODUCTION

Paul Walker[†]

Abstract

There is a small but growing literature on the theory of the firm based around the “reference point” theory of (incomplete) contracts formulated in Hart and Moore (2008). The reference point approach has been applied to the theory of the firm in Hart (2008, 2009), Hart and Moore (2007) and Hart and Holmström (2010). This introduction reviews the main theoretical aspects of each of these papers. It then discusses the relationship between the reference point approach to the firm, the transaction cost approach and the property rights approach. Here it is argued that the reference point approach is a step back towards ex post inefficiencies, away from reliance on ex ante inefficiencies.

JEL Classification: D23, D86, L22.

Keywords: contracts, theory of the firm, reference points.

[†]psw1937@gmail.com

1 introduction

Since the early 1970s the mainstream literature on the theory of the firm has developed largely around the idea of incomplete contracts.¹ At the forefront of this development have been the transaction cost and the property rights approaches to the firm.² For the transaction cost theory the importance of contractual incompleteness is that it gives rise to ex post inefficiencies, while for the property rights approach it gives rise to ex ante inefficiencies.

The reliance of the property rights approach on a theory of incomplete contracts that assumes unforeseen or indescribable states of the world as a way of generating incompleteness has led to a number of criticisms, the most significant of which is the Maskin and Tirole (1999) critique. Maskin and Tirole argue that information which is observable to the contracting parties can be made verifiable (to a third party) by the use of ingenious revelation mechanisms. The contracting parties write into their contract a game which when played gives the appropriate incentives for them to truthfully reveal their private information in equilibrium.³ This undermines the non-verifiability approach to incomplete contracts.

To deal with the Maskin and Tirole critique, Hart and Moore (2008) developed the “reference point” approach to contracts. Hart and Moore utilise a theory organised around the concepts of “aggrievement” and “shading” to develop an approach to incomplete contracts that is not only free of the Maskin and Tirole critique but which can, when applied to the theory of the firm, overcome other weaknesses – such as the reliance on noncontractible, relationship-specific investments and the assumption that parties always bargain to to an ex post efficient outcome – of the property rights approach to the firm.

The basic ideas underlying the “reference point” approach to incomplete contracts and its application to the theory of the firm can be gathered from a simple example which is a slightly modified version of an example taken from the second section of Hart (2008). Hart assumes that a seller, S , can provide a good or service, costing 10, to a buyer, B , who is willing to pay 20. Let us assume that we are talking about a public lecture on some aspect of microeconomics which B is organising and which B wants S to give. A successful lecture is worth 20 to B and it costs 10 for S to give the lecture.

At this stage Hart ignores the fact that B could engage other economists or that S could

¹For surveys of the theory of the firm see Tirole (1988: 15-61), Hart (1989), Holmström and Tirole (1989), Foss, Lando and Thomsen (2000), Gibbons (2005) and Aghion and Holden (2011).

²See Williamson (1975, 1985, 1996), Grossman and Hart (1986), Hart and Moore (1990) and Hart (1995).

³See Maskin (2002) and Aghion and Holden (2011: 190-3) for non-technical discussions and Maskin and Tirole (1999) for a formal exposition.

give lectures elsewhere. While trade could proceed smoothly, it is also possible that it will not. We will assume that B and S each have some discretion over the ‘quality’ of performance. For example, S could give a witty, lively, entertaining lecture or a very boring one. B on the other hand could treat S well, give her a nice dinner and pay quickly, or treat her badly.

In the language of Hart and Moore (2008) each party is able to provide basic (perfunctory) or exemplary (consummate) performance. It is further assumed that only the perfunctory level of performance can be legally enforced: consummate performance is entirely discretionary. It is assumed that each party is more or less indifferent between providing each level of performance – exemplary performance costs only a little more than basic or may even be slightly more pleasurable – and will provide exemplary performance if they feel they are being ‘well treated’ but not if they feel they are been ‘badly treated’. Cutting back on exemplary performance is called ‘shading’. Such behaviour cannot be observed or punished by an outsider. Shading hurts the other party.

Hart, following on from Hart and Moore (2008), emphasises that each party will feel ‘well treated’ if they receive what they think they are entitled to; that a contract between the parties is a ‘reference point’ for perceived entitlements; and that should there be no reference point, then entitlements can diverge, wildly in some cases.

To return to the example above. First, we will add a time line, see Figure 2.1. The time line tells us that B and S will write a contract some months before the lecture is given, date 0, rather than at the last minute, date 1. One reason for this is that each party has more options earlier on. In fact it is assumed that there is a competitive market for sellers, at date 0.

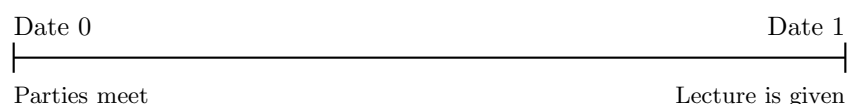


Figure 2.1. Time line.

Assume, further, that although B and S sign a contract at date 0, they leave the question of how much B will pay S open until the night before the lecture, date 1. This may seem a bad idea, and later it will be shown that it is. If no price is specified, then any p between 10 and 20 is possible. What might each party feel entitled to?

Hart and Moore (2008) take the view that entitlements can diverge. S may feel that the whole success of the talk will be due to her giving it and thus she feels entitled to $p = 20$. On the other hand B may have a somewhat different view of S ’s abilities and likely contribution and thus feel that S is worth much less, say, $p = 10$.

Even though they disagree as to what p should be, they are rational enough to arrive at a compromise, say $p = 15$. According to Hart and Moore (2008) each party will feel short-changed and therefore ‘aggrieved’. Since B is aggrieved by 5, (15-10), B shades to the point where S ’s payoff falls by 5θ , where θ is the constant of proportionality. And since S is also aggrieved by 5, (20-15), S shades to the point where the payoff for B falls by 5θ .

The end result of this is that if S and B leave the determination of the price until the night before the lecture, there will be a deadweight loss of 10θ due to the shading activities of each party. This reduces the value of the relationship between S and B from 10 to $10 - 10\theta = 10(1 - \theta)$.

Next Hart asks the question: Can anything be done to avoid this deadweight loss? His answer is yes. But first note an answer that doesn’t do the job. Ex post Coasian bargaining at date 1 doesn’t work. The reason is that shading is not contractible and thus an agreement not to shade is not enforceable. Or to put this another way, if B offers to pay S more to reduce her shading, say B offers to pay $p = 16$ to S rather than 15, then this will indeed reduce S ’s shading, from 5θ to 4θ , since S will now feel less aggrieved, but it will also increase B ’s shading from 5θ to 6θ , since he now feels more aggrieved. Total deadweight loss does not change, it remains at 10θ . However there is a simple solution; the parties just put the price in the contract at date 0. Since it has been assumed that the market for lectures is competitive at date 0, B will be able to hire S for a price $p = 10$. With this price specified in the contract, there is nothing for B and S to disagree about at date 1. The fact that B and S may disagree about the contribution that S makes to the success of the lecture on longer matters. B and S have agreed that B will pay S 10, and neither B nor S will be disappointed or aggrieved when that happens. Importantly, agreeing in advance, at date 0, to a payment of 10 eliminates ex post argument and aggrievement, and thus both parties will be willing to provide exemplary performance. Here we have the first best being achieved and zero deadweight losses as a result. This does raise an obvious question: What changes between dates 0 and 1? Why does a date 0 contract that fixes p avoid aggrievement, whereas a date 1 contract that fixes p does not? The crucial point here is the role of the ex ante market at date 0. This market gives an objective measure of what B and S bring to the relationship. Given the assumption of a competitive date 0 market, there are many sellers willing to supply at $p = 10$ and thus S accepts that she cannot expect to receive more than 10, while B understands that he can’t expect to pay less, as no one would be willing to give the lecture for less. Thus, neither party is aggrieved by $p = 10$. This gives us a model of the contractual relationship between B and S , but, as Hart explains, we need one further ingredient to create a theory of the firm.

Now let us add a little more realism by assuming that not all of the details of the lecture can be anticipated at date 0. To keep things simple we will assume that two different lectures can be given. Table 4.1 gives the payoffs and costs of each lecture.

	Lecture 1	Lecture 2
Value	20	14
Cost	10	8
Surplus	10	6

Table 4.1. Payoffs to lectures

Lecture 1 - say, a theory of the firm lecture - is the same as above, with a value of 20 and costs of 10. Lecture 2 - say, a microeconometrics lecture - yields value of 14 and costs of 8. Note that lecture 1 is more efficient in that it generates a greater surplus. Assume that the lectures can not be specified in the date 0 contract, since thinking about econometrics is sooooo boring that no one can stay awake long enough to write the contract! At date 1, however, the choice between them becomes clear.

Now we have to compare two organisational forms: an employment contract and an independent contractor. First, let B and S fix the price of the good at date 0, at say 10, and let B and S agree at date 0 that S will be an independent contractor. This is, in other words, a market exchange between two separate economics agents. Independent contractor means here that S gets to pick which lecture to give.⁴

Hart then asks, What will S do? Given that the price has been fixed by the date 0 contract, S will pick lecture 2, since it is cheaper for her. But note this is inefficient. B will be aggrieved because S didn't choose lecture 1, which B feels entitled to; B is short-changed by 6 (20-14), and he will therefore shade enough to reduce S 's payoff by 6θ . Total surplus in this case will be $6 - 6\theta$.

The second organisation form to be considered is an employment contract. B and S agree at date 0 that S is an employee of B . This we take to mean that S will work for B at a fixed wage, again assume 10. B , being the employer, has the right to decide on which lecture is to be given. As the wage is fixed B will choose lecture 1, as this gives him the greater value. This is efficient. S will be aggrieved since lecture 2 wasn't chosen, but S 's aggrievement is only 2. This induces S to shade by enough to reduce B 's payoff by 2θ . Total surplus is therefore $10 - 2\theta$.

Under the conditions specified, the employment contract is better. This is true for two related

⁴That is S has the residual control rights.

reasons. First, the lecture matters more to B than to S . B will lose $20 - 14 = 6$ if his favoured lecture is not chosen while S only loses $10 - 8 = 2$ if her favoured lecture is not chosen. This means it is efficient for B to choose the lecture. Second, and related, S 's aggrievement is low since she doesn't care very much.

Hart now changes the numbers in Table 4.1 to create Table 5.1.

	Lecture 1	Lecture 2
Value	20	14
Cost	10	2
Surplus	10	12

Table 5.1. Modified payoffs to lectures

Keep lecture 1 as it is, but change lecture 2 so while it still yields 14, it now costs only 2. Lecture 2 is now the more efficient (12 v's 10). Under employment, lecture 1 will be chosen, yielding a total surplus of $10 - 8\theta$. If S is an independent contractor, lecture 2 will be chosen resulting in a total surplus of $12 - 6\theta$.

What this suggests is that employment is good if the lecture matters more to B than to S , while independent contracting is good if the lecture matters more to S than to B .

Hart goes on to say that “[o]ne point worth emphasizing is that in neither of the above examples is the following contract optimal: to leave the choice of price and method until date 1, i.e. to rely on unconstrained Coasian bargaining. This would always yield the efficient method, but the aggrievement costs would be high. In [Table 4.1] the parties would agree on method 1; however, since there are 10 dollars of surplus to argue over, shading costs equal 10θ : net surplus = $10(1 - \theta)$, which is less than that obtained under the employment contract. In [Table 5.1] there are 12 dollars of surplus to argue over and net surplus = $12(1 - \theta)$, which is less than that obtained under independent contracting.” (Hart 2008: 409).

Clearly the examples above are toy ones, but Hart argues they contain the basic ingredients of a theory of the firm in that they consider the choice between carrying out a transaction in the market, using an independent contractor, and ‘inside the firm’, via an employment contract.

The basic ingredients of the reference point approach to contracts have been applied to the theory of the firm more fully in papers by Hart and Moore (2007), Hart (2009) and Hart and Holmström (2010). In the next three sections of this paper we look at the relevant theoretical components of each of these three papers. The Hart and Moore (2007) paper attempts to move the property rights literature away from its reliance on ex post efficiency and distortions in ex

ante investments, Hart (2009) looks at the hold-up problem and asset ownership while Hart and Holmström (2010) models the boundaries of the firm incorporating strategic decisions that are taken in the absence of *ex post* bargaining. This diverse series of papers deals with a number of issues central to the theory of the firm but within a unified framework based around the reference point approach to contracts. While providing a detailed synthesis of such a small and recent literature seems premature at this juncture, in Section 5 we do consider the relationship between the reference point approach, the transaction cost approach and the property rights approach to the firm. Section 6 is the conclusion.

2 Hart and Moore (2007)

Hart and Moore (2007) (HM) is an attempt to broaden the property rights literature in such a manner as to move this literature away from its use of renegotiation of an incomplete contract to achieve ex post efficiency and its focus on distortions in ex ante investments. HM open their paper by discussing a simple version of the standard property rights approach to the firm. HM show that hold-up can occur and that the standard approach to dealing with it is via the allocation of non-human assets among the contracting parties. HM also point out weaknesses with the standard model including the Maskin and Tirole (1999) critique; the model's reliance on noncontractible, relationship-specific investments which are by their very nature problematic in that they are hard to measure empirically; and the assumption that parties always bargain to to an ex post efficient outcome. HM argue that the use of side payments to achieve such an efficient outcome is a poor description of what goes on inside real firms.

All of this suggests to HM that it is worth trying to develop an alternative model. They begin to do this by noting that Hart and Moore (2008) develop a theory of incomplete contracts based on the idea that a contract can act as a reference point for the contracting parties' feelings of entitlement and that such feelings of entitlement can affect contractual performance. The basic ideas underlying the theory can be outlined as follows. Consider a situation where a buyer B wants a good or service from a seller at some future date 1. Assume the good is a homogeneous widget. Also assume that there is a 'fundamental transformation' between dates 0 and 1, that is, what starts as a situation of perfect competition at date 0 evolves into one of bilateral monopoly by date 1.

The parties meet and contract at date 0 but there is uncertainty about the state of the world at this time. This uncertainty is resolved shortly before date 1. There is symmetric information

thought-out but the state of the world is not verifiable. A date 0 contract can be thought of as specifying a set of possible price-quantity pairs which form the set of possible outcomes of B and S 's date 1 transaction. Note that the outcomes cannot be state contingent since the state itself is not verifiable. A mechanism for choosing from among the set of possible outcomes may also be included as part of the date 0 contract.

Importantly for the HM story, the date 0 contract acts of a reference point for the contracting parties feelings of entitlement at date 1. Neither party feels entitled to an outcome not included in the contract. The contract is seen as “fair” since it was negotiated under the competitive conditions prevailing at date 0.⁵ Problems can arise, however, when choosing among the different outcomes allowed under the contract. HM suppose that each party feels entitled to the best possible outcome allowed under the contract.⁶ This means that it is likely that at least one of the parties, if not both, will feel disappointed or aggrieved by the actual outcome.

A second important assumption built into the HM theory is that outcome is not perfectly contractible even at date 1. Each party has the freedom to choose between perfunctory performance and consummate performance but only perfunctory can be contracted on. Consummate performance will be provided if the party feels well treated but each party will shade if they feel aggrieved.

To make matters a little more precise, suppose that if the outcome that is chosen from those available under the contract causes S to feel aggrieved by $\$k$, that is, S 's actual payment is $\$k$ less than the best possible outcome, then S will shade on her performance to such a degree that B 's payoff falls by θk . θ being an exogenously given parameter where $0 < \theta \leq 1$. A similar situation with regard to shading pertains to B . There a symmetry here, both B and S can shade and θ is the same for both. Shading can not occur if there is no trade.

Assume further that B 's value of the widget at date 1 is v and that S 's cost of production is zero but there is an opportunity cost of r . That is, trade with B means S foregoes an alternative income of r . Thus, trade is efficient if and only if $v \geq r$. At date 0, v and r are random variables with a probability distribution which is common knowledge. Also no third party is able to tell who is at fault if trade does not take place at date 1. This means that trade is voluntary. Given these assumptions Hart and Moore (2008) are able to show that it is only the difference between the trade price and the no-trade price that matters and that is possible to normalise the no-trade price to be zero. HM also assume that lump-sum transfers can be used to carryout any

⁵This is an important point for all the papers considered in this survey. See Fehr, Hart and Zehnder (2011) for experimental evidence related to this issue.

⁶Such an extreme assumption not necessary for the analysis but does simplify the workings.

redistribution of surplus at date 0.

The simplest case to consider is that where there is no uncertainty as to the value of v and r , that is, they are constants, and $v > r$. In this situation the first-best can be achieved. All that has to happen is that the parties agree, at date 0, that S will supply the widget to B at date 1 for a given price, p , where $r < p < v$. This contract would ensure trade and would result in no aggrievement because both parties receive the best outcome permitted under the contract. Note that the contract only specifies one outcome, trade at price p .

While this form of contract achieves the first-best not all contracts, even in this no uncertainty world, do so. For example, consider a contract that specifies that the trade price can be anything in the range $[r, v]$ and that B will choose the price. Here B will choose the lowest price possible, $p = r$, at date 1. Note however that this will cause S to be aggrieved since the best possible price for her, $p = v$, was not chosen and thus she will shade resulting in a deadweight loss of $\theta(v - r)$.⁷

Things are more interesting, however, if v and r are uncertain. In this case any contract which specifies a single trading price, p , will ensure trade if and only if $v \geq p \geq r$, that is, if and only if both parties gain from trade. The problem is that as v and r are now stochastic, it can not be guaranteed that it is possible to find a single p that lies between v and r whenever $v > r$. HM point out that under such conditions a contract that specifies a range of trading prices $[\underline{p}, \bar{p}]$ can be superior to a single price contract. Hart and Moore (2008) show that it is not necessary to go beyond a contract which specifies a no-trade price of zero, as above, a trading price range of $[\underline{p}, \bar{p}]$ and lets B choose the price. The advantage of the large price range is that it makes it more likely that B can find a price between v and r whenever $v \geq r$. The cost is that there are typically many feasible prices between v and r when $v \geq r$ and B will pick the lowest price. This means that S will feel aggrieved that B did not pick the highest price and will therefore shade, resulting in a deadweight loss. The optimal contract will trade off these two effects.

Thus far one important issue has been ignored. If $v > r$ but there is no price in the range $[\underline{p}, \bar{p}]$ such that $v \geq p \geq r$, it would be expected that the parties would renegotiate their contract. But renegotiation doesn't change the analysis in any fundamental ways as is shown in Hart and Moore (2008).

Next HM turn to the issue of ownership. Up to this point HM have implicitly assumed that B and S are separate entities, that is, they are "nonintegrated". Now suppose that B acquires S 's firm (S 's non-human assets) at date zero. This is interpreted as "integration". It amounts to

⁷ S 's best outcome is $p = v$ while the actual outcome is $p = r$ and thus S is aggrieved by the amount $v - r$ which means she shades, thereby lowering B 's payoff, by an amount $\theta(v - r)$. This is the deadweight loss. B does not shade as she receives her best outcome.

saying that B now owns and possesses the widget. HM take this to mean that B can get someone other than S to produce the widget, at zero cost, at date 1. It is assumed that S 's human capital, along with the widget, is still needed to realise the opportunity cost, r .⁸ Effectively for S to earn r , B must sell the widget back to S . If no trade occurs B earns v since she already owns the widget, while if trade does occur, S earns r but pays p . Trade is now efficient if and only if $r \geq v$; trade is still voluntary. In this situation a contract consists of a zero no-trade price and a range of trading prices $[\underline{p}, \bar{p}]$, with S choosing the price. S will choose the smallest price such that $r \geq p \geq v$, whenever $r > v$.

In place of a complete analysis of nonintegration versus integration HM makes a number of observations on the difference between them. Assume that $v > r$ with probability 1. Then as was noted above, it may be impossible to achieve the first-best. The reason being that in order to ensure trade with probability 1 it may be necessary to have a range of trading prices, but this results in aggrievement and shading whenever there is more than one price satisfying $v \geq p \geq r$. On the other hand integration can achieve first-best because the status-quo has been transformed into one where B owns the widget, which is the efficient outcome. S is irrelevant and does not or cannot shade.

The situation is reversed if $v < r$ with probability 1: now integration is inefficient as a range of prices is required to ensure that B trades the widget to S . This results in aggrievement and thus shading whenever there is a number of feasible prices in the range while for nonintegration the status-quo point has S possessing the widget which is efficient and does not give rise to shading.

The HM model can be thought of as capturing the idea that integration is useful for ensuring input supply in an uncertain world. When $v > r$ but v and r vary, nonintegration is usually inefficient, that is, either trade will not take place when it should or there will be shading, while integration results in the first-best outcome.

3 Hart (2009)

The aim of Hart's paper to reexamine some of the central themes of the incomplete contracts literature – by which Hart means the hold-up problem and asset ownership – through the lens of contracts as reference points. The model Hart develops considers a (long-term) economic relationship between a buyer and seller where there is initial uncertainty about the buyer's value and/or the seller's cost. Importantly, for the relationship to work out the parties need to cooperate

⁸ S gets zero if she has only their own human capital to work with. To get r , S requires both her human and non-human (or equivalently the widget) assets.

in ways that cannot be specified in an initial contract. The buyer and seller, therefore, face a trade-off between, on the one hand, a flexible contract which attempts to index the terms of trade – price – to the state of the world and on the other hand, a (relatively) rigid contract, e.g., a fixed price contract. A fixed price contract works well in “normal” times because there is nothing for the parties to argue about. It is assumed that in the absence of any argument the parties will be willing to cooperate. However should the value or costs fall outside the normal range, one of the parties will have an incentive to engage in hold-up. That is, one party will threaten to withhold cooperation, unless the contract is renegotiated in their favour. An example would be where the buyer’s valuation is unusually high and the seller would therefore have an incentive to hold-up the buyer to get a higher price. If, on the other hand, costs were surprisingly low, then the buyer would have an incentive to hold-up the seller to force a lower price. Hart assumes that hold-up transforms a friendly relationship into a hostile one. This means that the parties operate within the letter rather than the spirit of their (renegotiated) contract which results in deadweight losses. However, even a hostile relationship is assumed to create more surplus than no trade, and so, if value or cost has moved sufficiently far outside the normal range, hold-up will occur.

Hart shows that there is a price range $[p_L, p_H]$ which depends on the state of the world such that, should the long-term price, p , which is chosen before the state of the world is known, fall within this range there is no hold-up; whereas if p falls outside this range, hold-up does occur. The range $[p_L, p_H]$ is nondegenerate, in fact $p_H > p_L$, because the deadweight costs involved with hold-up introduces frictions. Hart also shows that if there is no *ex ante* uncertainty then it is easy to find a p which lies in $[p_L, p_H]$. Exactly where in this range it falls doesn’t matter since lump-sum transfers can be used to redistribute surplus. However, should the state of the world be uncertain, it may be impossible to find a p that falls within the required range $[p_L, p_H]$ with probability 1. This means that hold-up will sometimes take place.

Hart shows that an appropriate allocation of the ownership of assets can help counter hold-up. In fact, more generally, enhancing a party’s outside option is useful. Hart writes that “[c]eteris paribus, allocating an asset to the buyer is good to the extent that this increases the correlation between the buyer’s outside option and his value from trade. This improves matters because, when value is unusually high, the buyer’s outside option will also be high, which reduces the seller’s ability to hold up the buyer. To put it a bit more formally, the $[p_L, p_H]$ range becomes less sensitive to the buyer’s value. However, there is a countervailing effect on the seller: taking an asset away from the seller causes the $[p_L, p_H]$ range to become more sensitive to the seller’s

but not on all of them. Some actions are never contractible. This means that some modification or renegotiation of the contract is possible at date 1-. Hart also assumes that all the helpful actions are chosen simultaneously by both parties at date 1. The time line is given in Figure 11.1.

Next Hart makes the following assumptions:

- A1. If at date 1 all helpful actions are taken, the value of the widget to B is v and the cost to S is c , where $v > c$. Hence net surplus equals $v - c$ in this case.
- A2. If at date 1 all the contractible, but none of the noncontractible, helpful actions are taken, the value of the widget to B is $v - \frac{1}{2}\lambda(v - c)$ and the cost to S is $c + \frac{1}{2}\lambda(v - c)$, where $0 < \lambda < 1$. Hence net surplus is $(1 - \lambda)(v - c)$ in this case. λ represents the proportion of net surplus lost due to the non-implementation of the noncontractible helpful actions.
- A3. If at date 1 none of the helpful actions (contractible or otherwise) are taken, B 's value is very low (approximately, $-\infty$) and S 's cost is very high (approximately, $+\infty$). In this case each party walks away from the contract (neither party has an incentive to enforce it) and no trade occurs; that is, the parties earn their outside options. The no-trade price has been normalised to zero.

The importance of A2 is that withholding noncontractible helpful actions moves both the value of the widget to B and the cost of the widget to S towards $\frac{1}{2}(v + c)$ ¹⁰ resulting in a net surplus of zero.

The next question is, What determines whether a party is helpful? That is, When do the parties act cooperatively? Following Hart and Moore (2008), Hart supposes that each party is indifferent between being helpful and not. Given this indifference, a party will be helpful if they feel they have been “well treated” by the other party, but not otherwise.

Hart starts his analysis by considering the case where the parties write a “simple” date 0 contract, that is, a contract which simply specifies a single trading price p . When date 1- is reached and the uncertainty is resolved each of the parties has a choice, either they can stick to the contract as agreed or they can try to force the other party to renegotiate the contract. This renegotiation is interpreted here as “hold-up”.

First Hart deals with the scenario where the parties stick to the contract. In such a situation each party feels well-treated by the other since they are getting exactly what the contract said

¹⁰As $\lambda \rightarrow 1$, $v - \frac{1}{2}\lambda(v - c) \rightarrow \frac{1}{2}(v + c)$ and $c + \frac{1}{2}\lambda(v - c) \rightarrow \frac{1}{2}(v + c)$.

they would: the contract specifies a single trading price p . Given that each feels well-treated, they are willing to be helpful and all cooperative actions are taken.

The buyer's and seller's payoffs are given by

$$U_b = v - p \quad (1) \qquad U_s = p - c \quad (2)$$

The second scenario considered is where one party engages in hold-up. This they do by trying to force the other party to renegotiate the contract by threatening to withhold all helpful actions unless given a sidepayment. Such an action is considered, by the other party, to be a breach of the spirit of the date 0 contract and leads to the end of cooperation. The result is a Nash equilibrium where neither party cooperates.¹¹ This is the no-trade outcome of A3, with payoffs to B and S of r_b and r_s .

But renegotiation is possible and thus even if the relationship is soured the parties can, and will, consent to carry out the **contractible** helpful actions at date 1. But there is a cost to renegotiation in that neither party will carry out the noncontractible helpful actions. Their relationship will be impersonal, if correct. The result is that renegotiation yields a surplus of $(1 - \lambda)(v - c)$ by A2. Informally, λ is the proportion of surplus generated by the non-contractible actions.

Thus whenever

$$A4. \quad (1 - \lambda)(v - c) > r_b + r_s$$

the parties will renegotiate their contract away from the no-trade outcome. A4 is assumed to hold in what follows.

The important point to take from this is that hold-up leads to a loss of surplus of $\lambda(v - c)$.

Hart assumes that there will be a 50:50 split of the surplus from renegotiation so after hold-up the parties' payoffs are given by

$$U_b = r_b + \frac{1}{2}G \quad (3)$$

$$U_s = r_s + \frac{1}{2}G \quad (4)$$

$$\text{where} \quad G = (1 - \lambda)(v - c) - r_b - r_s \quad (5)$$

From this it is possible to determine when hold-up occurs. Define p_L to be the price such that S is indifferent between receiving p , i.e. being helpful, and holding B up and let p_H be the

¹¹This is a Nash equilibrium because each party is indifferent between cooperating and not doing so.

price at which B is indifferent between paying p and holding S up. Then from (1)–(4) we get

$$p_L - c = r_s + \frac{1}{2}G \quad (6)$$

$$v - p_H = r_b + \frac{1}{2}G \quad (7)$$

and thus

$$p_L = c + r_s + \frac{1}{2}G \quad (8)$$

$$p_H = v - r_b - \frac{1}{2}G \quad (9)$$

Note that

$$p_H - p_L = \lambda(v - c) > 0. \quad (10)$$

Which means that the difference between the two hold-up prices equals the loss of surplus due to hold-up. Equation (10) reflects the fact that there are frictions in the renegotiation process. If the renegotiation could be achieved without acrimony, that is if $\lambda = 0$, then $p_H = p_L$. This implies that $p_H = p_L$ when all helpful actions are contractible and thus hold-up is not possible. However since hold-up causes some dissipation of surplus, $p_H > p_L$: the price at which B is just willing to hold-up S is strictly greater than the price at which S is just willing to hold-up B . Informally, “low” prices suit buyers but not sellers and thus sellers are more likely to be disgruntled at “low” prices and therefore have an incentive to engage in hold-up at lower prices than buyers.

Note that since S is indifferent between holding B up and not doing so at a price $p = p_L$, S will strictly prefer to hold up B when $p < p_L$. Similarly, if $p > p_H$ then B will strictly prefer to hold up S . Thus hold-up is avoided if and only if

$$p_L < p < p_H \quad (11)$$

It is worth noting that p_H and p_L vary with the state of the world, they are random variables, but p is not, since it is chosen ex ante.

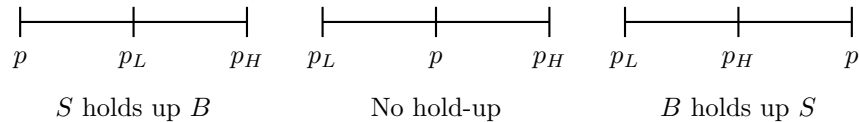


Figure 14.1. Hold-up.

To make progress Hart places more structure on the random variables, r_b and r_s . In particular he assumes

$$r_b = \alpha_b + \beta_b v + \varphi + \gamma_b \varepsilon \quad (12)$$

$$r_s = \alpha_s - \beta_s c + \gamma_s \eta \quad (13)$$

$$\text{where } 1 - \lambda > \beta_b > 0, 1 - \lambda > \beta_s > 0, \gamma_b > 0, \gamma_s > 0. \quad (14)$$

Here $\alpha_b, \beta_b, \gamma_b, \alpha_s, \beta_s,$ and γ_s are constants (later they will depend on the assets that each party owns) and φ, ε and η are independent random variables with zero mean. (12), (13) and (14) capture the idea the outside options for both B and S co-vary with v and c , respectively, but not too strongly, and are also subject to exogenous noise (ε, η). The noise term φ is a smoothing device: its rationale will become clear in the following section.

Using (12) and (13), it is possible to represent the state of the world as a 5-tuple $\omega = (v, c, \varphi, \varepsilon, \eta)$. Both B and S observe ω at date 1-. Note that a simple contract refers to a contract consisting of a single price p , where p is chosen before the state of the world ω is known.

It is useful to rewrite p_L, p_H as functions of ω , a process that gives

$$p_L(\omega) = \frac{1}{2}[\alpha_s + \gamma_s \eta - \alpha_b - \varphi - \gamma_b \varepsilon + ((1 - \lambda) - \beta_b)v + ((1 + \lambda) - \beta_s)c] \quad (15)$$

$$p_H(\omega) = \frac{1}{2}[\alpha_s + \gamma_s \eta - \alpha_b - \varphi - \gamma_b \varepsilon + ((1 + \lambda) - \beta_b)v + ((1 - \lambda) - \beta_s)c] \quad (16)$$

Clearly both $p_L(\omega)$ and $p_H(\omega)$ are increasing in v, c . Remember that hold-up occurs when $p_L(\omega) > p$ or $p_H(\omega) < p$, or, in other words, when either $p_L(\omega)$ is high or $p_H(\omega)$ is low. Thus it is clear that hold-up occurs, ceteris paribus, if v is exceptionally high or low or c is exceptionally high or low. This is intuitive: if v is high, S does relatively well in the renegotiation process because there is a lot of surplus available, even accounting for the λ loss, and so S has an incentive to hold up B . Similarly, should the value of c be low, B does relatively well in the renegotiation process and thus will have an incentive to hold-up S . From (15) and (16) we see that the effects of extreme values of v and c are less pronounced if β_b and β_s are large, since p_L and p_H will be less sensitive to v and c if they are.

Hart now looks at the optimal simple contract. Since date 0 lump-sum transfers can be

applied to allocate surplus, an optimal contract is one which maximises expected net surplus.

$$\begin{aligned} \underset{p}{Max} &= \int (v - c)dF(\omega) + \int (1 - \lambda)(v - c)dF(\omega) \\ & \quad p_L(\omega) \leq p \leq p_H(\omega) \quad \begin{array}{l} p < p_L(\omega) \\ \text{or } p > p_H(\omega) \end{array} \end{aligned} \quad (17)$$

where F is the distribution function of ω .

It is straightforward to see that under certainty the first-best can be achieved: just pick any price p in the interval $[p_L(\omega_0), p_H(\omega_0)]$, where ω_0 is the state of the world. However, the first-best typically cannot be achieved under uncertainty, because it is not generally possible to find a single price that lies in the intersection of a number of different $[p_L(\omega), p_H(\omega)]$ intervals.

Hart notes a shortcoming with the analysis so far:

“[s]uppose that the parties write a simple contract at date 0. Then, as we have observed, with uncertainty it is very likely that p will lie outside the $[p_L(\omega), p_H(\omega)]$ range for some ω , and so one party will hold up the other to get a better price. Why don’t the parties anticipate this and build the renegotiated price into the original contract? For example, the initial contract could state that the price will normally be 10 but can rise to 20 in unusual circumstances. Or the contract might give one party the right to choose the price from a menu of prices.” (Hart 2009: 277).

In the Appendix to his paper Hart deals with (some of) these possibilities. Here Hart allows for the parties to specify a range of possible trading prices, $[\underline{p}, \bar{p}]$, in their contract at date 0. For each state of the world the parties negotiate over which price in the range to choose. As long as they stay within the agreed range hold-up is avoided. The avoidance of hold-up is the major advantage of a large price range. The cost of such a range is that at date 1— each party may feel entitled to a different price within the price range. As in Hart and Moore (2008) not receiving what you think you are entitled to results in “aggrievement” and “shading”, that is, partial withholding of cooperation. If it were not for shading a price range of $\underline{p} = -\infty$, $\bar{p} = \infty$ would result in the first-best being able to be achieved. Given that this amounts to placing no restrictions on price, this is equivalent to no contract being written at date 0.

If, however, aggrievement and shading exists, a very large price range is not optimal since it will lead to shading in all states of the world. Thus there is a trade-off between the risk of hold-up and the cost of shading. The parties may accept a risk of hold-up by limiting the price range, to avoid shading. In the extreme case a fixed price contract will be signed.

Next Hart introduces assets into the model. Asset ownership matters because it determines which assets a given party can walk away with if trade does not take place. This affects each

party's outside options and their incentives to engage in hold-up.

Hart denotes by A the set of all assets at B and S 's command. A is assumed to be fixed and finite. Let A_b be those assets which B owns and A_s be the assets S owns. It is assumed that,

$$A_b \cap A_s = \emptyset, \quad A_b \cup A_s \subseteq A. \quad (18)$$

The first part of (18) tells us that both B and S can't walk away with the same asset while the second part states that if an asset is jointly owned then neither party can walk away with it.¹² This means that joint ownership gives each of the parties veto rights on the use of the asset.

Now assume that the coefficients $\alpha_b, \beta_b, \gamma_b, \alpha_s, \beta_s, \gamma_s$ depend on the ownership of assets. In particular $\alpha_b = \alpha_b(A_b), \beta_b = \beta_b(A_b), \gamma_b = \gamma_b(A_b), \alpha_s = \alpha_s(A_s), \beta_s = \beta_s(A_s)$ and $\gamma_s = \gamma_s(A_s)$. It is also assumed that asset ownership increases the *marginal* payoffs of r_b, r_s with respect to v and c . That is,

$$\beta_b \text{ is nondecreasing in } A_b \quad (19) \quad \beta_s \text{ is nondecreasing in } A_s \quad (20)$$

The idea being captured by (19) and (20) is that the assets are specific to the trading that B and S are engaged in. An alternative would be to suppose that each party's total payoff increases in the assets they own, that is, α_b, α_s are increasing in A_b, A_s . But this is not needed in what follows. (14) and A4 are assumed to hold for all ownership structures.

Hart also assumes that the assets can be traded at date 0. Therefore a contract is now a 3-tuple made up of assets and a price, that is, (A_s, A_b, p) which specifies an ownership allocation, (A_s, A_b) and a date 1 price, p , where A_s, A_b satisfy (18). Assume that date 0 lump-sum transfers are possible, and thus an optimal contract maximises expected net surplus:

$$\text{Max}_{(A_b, A_s, p)} \left\{ \begin{array}{l} \int (v - c) dF(\omega) \quad \quad \quad + \int (1 - \lambda)(v - c) dF(\omega) \\ p_L(\omega; A_b, A_s) \leq p \leq p_H(\omega; A_b, A_s) \quad p < p_L(\omega; A_b, A_s) \\ \text{or } p > p_H(\omega; A_b, A_s) \end{array} \right\} \quad (21)$$

where p_L and p_H are indexed by both the state of the world, ω , as before, and the asset ownership allocation, (A_b, A_s) .

Consider what happens if, *ceteris paribus*, assets are taken, at date 0, from S and given to B . Then from (19) and (20), we know that β_b rises and β_s falls. This in turn means that p_L and p_H

¹²If $A_b \cup A_s \subset A$ then there must be some asset not owned solely by one of B or S . Such an asset is assumed to be under joint ownership.

will be less sensitive to v than they were before, note that from (15) and (16) we get,

$$\frac{\partial p_L}{\partial v} = \frac{1}{2}((1 - \lambda) - \beta_b) \quad (22)$$

$$\frac{\partial p_H}{\partial v} = \frac{1}{2}((1 + \lambda) - \beta_b) \quad (23)$$

and both of these decrease. On the other hand, it is also clear that p_L and p_H become more sensitive to c , since

$$\frac{\partial p_L}{\partial c} = \frac{1}{2}((1 + \lambda) - \beta_s) \quad (24)$$

$$\frac{\partial p_H}{\partial c} = \frac{1}{2}((1 - \lambda) - \beta_s) \quad (25)$$

and these two increase.

It seems sensible to consider a reduction in sensitivity in p_L and p_H as good since if the interval $[p_L, p_H]$ does not vary to any great degree it will be easier to find a price p which lies within $[p_L, p_H]$ for many ω . This reduces the opportunity for hold-up. This suggest that it is optimal for B to own all of the assets if only v varies, because such an ownership structure minimises the sensitivity of p_L and p_H with respect to the state of the world. It also suggests that it is optimal to have S owning all the assets if only c varies. Proposition 1 notes this.

Proposition 1 (1) *Suppose that $\varphi = \epsilon = \eta \equiv 0$ and $c \equiv c_0$ where c_0 is a constant. Then there exists an optimal contract in which B owns all the assets, that is, $A_b = A$, $A_s = \emptyset$.*

(2) *Suppose that $\varphi = \epsilon = \eta \equiv 0$ and $v \equiv v_0$ where v_0 is a constant. Then there exists an optimal contract in which S owns all the assets, i.e., $A_s = A$, $A_b = \emptyset$. (Hart 2009: 282).*

Hart proves a more general versions of the propositions of this section, in the Appendix to his paper. The more general versions allow for a range of prices to be specified in the date 0 contract.

Proposition 1 has an obvious limitation in that it only tells us when one party should own everything. In general both v and c will vary. Hart makes progress by introducing the idea of an idiosyncratic asset. Define an asset to be idiosyncratic to B if B 's ownership of it increases the sensitivity of r_b to v and S not owning it has no effect on the sensitivity of r_s to c . An analogous definition applies to any asset which is idiosyncratic to S .

DEFINITION (i) Asset a is idiosyncratic to B if $\beta_b(A_b \cup \{a\}) > \beta_b(A_b)$ for all $A_b \subseteq A$, $A_b \cap \{a\} = \emptyset$ and $\beta_s(A_s \cup \{a\}) = \beta_s(A_s)$ for all $A_s \subseteq A$.

(ii) Asset a is idiosyncratic to S if $\beta_s(A_s \cup \{a\}) > \beta_s(A_s)$ for all $A_s \subseteq A$, $A_s \cap \{a\} = \emptyset$ and $\beta_b(A_b \cup \{a\}) = \beta_b(A_b)$ for all $A_b \subseteq A$. (Hart 2009: 283).

Thus an asset is idiosyncratic to one firm if it is specific to that firm but not to the other. If this is so then taking the asset away from the other firm is unlikely to affect the sensitivity of that party's outside option to the state of the world. While it would seem that allocating an asset to the party for whom it is idiosyncratic is desirable, since it reduces variability in the $[p_L, p_H]$ range, actually proving such a result requires making strong assumptions about the stochastic structure. In Proposition 2 Hart assumes that with high probability v, c take on "normal" values $v = v_0, c = c_0$, whereas with a low probability, v, c , can each take on "exceptional" values. Hart further assumes, that because exceptional values are very unlikely, the possibility that v and c will take on exceptional values simultaneously can be ignored. It is also assumed that there is a small amount of exogenous noise via the random variable φ , which affects only r_b , but $\varepsilon = \eta = 0$.

Proposition 2 *Assume that $\varepsilon = \eta = 0$ and that φ is uniformly distributed on $[-k, k]$. Suppose that with probability $0 < \pi < 1$, event 1 occurs: $v = v_0, c = c_0$; with probability $(1 - \pi)\alpha_v$ event 2 occurs: $c = c_0, v$ has support $[v_L, v_H]$, where $v_L \leq v_0$ and*

$$v_H \geq \frac{-\beta_b(A)v_0 + (1 + \lambda)v_0 - 2\lambda c_0}{1 - \lambda - \beta_b(A)};$$

with probability $(1 - \pi)\alpha_c$ event 3 occurs: $v = v_0, c$ has support $[c_L, c_H]$, where $c_H \geq c_0$ and

$$c_L \leq \frac{-\beta_s(A)c_0 - 2\lambda v_0 + (1 + \lambda)c_0}{1 - \lambda - \beta_s(A)}.$$

Here $\alpha_v > 0, \alpha_c > 0, \alpha_v + \alpha_c = 1, k > 0$, and φ is independent of v and c in events 2 and 3, respectively. Then, for small enough k , the following is true: if π is close to 1, it is uniquely optimal for B to own asset a if a is idiosyncratic to B and for S to own asset a if a is idiosyncratic to S . (Hart 2009: 283-4).

Hart explains the use of such strong assumptions on the probability distribution of v and c as follows,

"[i]t is useful to understand why Proposition [2] requires such strong assumptions about the probability distribution of v and c . The reason is the following. Let p be the optimal price for the general case where v, c are uncertain. Suppose that we transfer an asset that is idiosyncratic to B from S to B . (In what follows, we suppress

assets in the notation.) We know that this will reduce the variability of $p_L(\omega), p_H(\omega)$ with respect to v . But this might reduce the probability that p lies in $[p_L(\omega), p_H(\omega)]$ if, for example, $p \in [p_L(\omega_1), p_H(\omega_1)], p \notin [p_L(\omega_2), p_H(\omega_2)]$, and $[p_L(\omega_1), p_H(\omega_1)]$ moves closer to $[p_L(\omega_2), p_H(\omega_2)]$. The stochastic structure in Proposition [2] avoids this kind of situation.” (Hart 2009: 284)

Hart uses, as an example of an application of Proposition 2, the case of strictly complementary assets. Assume that assets a_1 and a_2 are strictly complementary, in the sense that they are valuable only when they are used together. So a_2 by itself is of no use to S , but a_1 and a_2 combined are of use to B . Suppose that B owns a_1 . Now we can define a new economy such that a_1 is always owned by B , i.e. a_1 is inalienable, and thus the set of alienable assets is $A \setminus \{a_1\}$. For this economy a_2 is idiosyncratic by Definition (i). Thus by Proposition 2 B should own a_1 and a_2 . The same basic argument will also show that if S owns a_1 it is optimal for S to own a_2 as well. The conclusion of this reasoning is that complementary assets should be held together. A similar line of reasoning results in the conclusion that joint ownership is suboptimal under the requirements of Proposition 2.

Thus far, Hart has emphasised the notion that ownership of an asset is good for one of the parties since it reduces the variability of that party’s payoff, via changes to v or c , relative to its outside option. It is, however, possible that asset ownership increases the variability of outside options relative to inside values. In such a situation, it may be better to take assets away from people. To consider this case, Hart holds the values of v and c constant and lets the values of φ and ε or η vary. Here he shows that joint ownership may be optimal.

Proposition 3 *Assume γ_b, γ_s are strictly increasing in A_b, A_s , respectively, and φ is uniformly distributed on $[-k, k]$. Suppose that with probability $0 < \pi < 1$, event 1 occurs: $v = v_0, c = c_0, \varepsilon = 0, \eta = 0$; with probability $(1 - \pi)\alpha_\varepsilon$, event 2 occurs: $v = v_0, c = c_0, \eta = 0, \varepsilon$ has support $[\varepsilon_L, \varepsilon_H]$, where*

$$\varepsilon_L \leq \frac{2\lambda(c_0 - v_0)}{\gamma_b(\emptyset)}$$

and $\varepsilon_H > 0$; with probability $(1 - \pi)\alpha_\eta$, event 3 occurs: $v = v_0, c = c_0, \varepsilon = 0, \eta$ has support $[\eta_L, \eta_H]$, where

$$\eta_L \leq \frac{2\lambda(c_0 - v_0)}{\gamma_s(\emptyset)}$$

and $\eta_H > 0$. Here $\alpha_\varepsilon \geq 0, \alpha_\eta \geq 0, \alpha_\varepsilon + \alpha_\eta = 1, k > 0$, and φ is independent of ε and η in events 2 and 3, respectively. Then, for small enough k , the following is true: if π is close to 1, it is

uniquely optimal for all assets to be jointly owned by B and S. (Hart 2009: 285).

4 Hart and Holmström (2010)

Hart and Holmström, (HH), claim that they present a new model of the boundaries of the firm, which takes into account strategic decisions that are taken in the absence of *ex post* bargaining. The absence of bargaining is important because the assumption that *ex post* conflict is resolved via bargaining is one feature that has limited the applicability of the standard property rights approach to the firm. To justify the use of authority rather than bargaining HH follow Hart and Moore (2008) and take a “contracts as reference points” approach.

The HH model assumes there are two ‘units’ involved in a lateral relationship. A unit is thought of as an irreducible set of activities for which it would be pointless to break up any further. Each of the units is operated by a manager who has to take a decision which effects not only his unit but also the other unit, that is, there are externalities associated with the decision. What HH are considering is a decision so significant that it warrants careful consideration of the organisational form which best supports it. An example would be where the units are deciding on the adoption of a common standard or platform for technology that is used jointly.

The two managers of the units have a binary, “Yes” or “No”, decision to make. HH assume that there are only two aggregate outcomes, which they denote as “coordination” and “non-coordination”. If, and only if, both units choose Yes do we get coordination. That is, either party can veto coordination by selecting No.

HH also assume that the decision by each unit is *ex ante* non-contractible but it is *ex post* contractible. Each unit has a boss. Importantly the boss has the right to make the decision in his unit *ex post*.¹³ HH consider two forms of organisation: non-integration and integration. Under non-integration the units are separate firms and the managers of the units are the bosses which mean they make the Yes/No decisions. Integration means that the two units are part of the same firm and a third party manager is the boss and thus takes the Yes/No decision. In this case the boss instructs the unit managers to choose either Yes or No and the manger follows these orders, that is, such instructions are contractible. The two managers can, however, shade on performance.

An important factor in the HH model is the assumption that each unit creates two forms of benefits: monetary profit, which is transferable with ownership, and private benefits which

¹³In other words, the boss has residual controls rights.

are not transferable. These private benefits can be thought of as a measure of job satisfaction, broadly defined.

The importance of the two types of benefits in the model can be illustrated in the following way. Let the pair (v_A, w_A) $[(v_B, w_B)]$ denote the profits and private benefits accruing to unit A [B]. HH assume that the manager of each unit is the only worker for that unit and thus the private benefits are a measure of his job satisfaction. They also assume that the boss of the unit can divert all profit from that unit to themselves.

If the two units are non-integrated and the manager of unit A [B] is also the boss of unit A [B] then the payoff for the manager of A [B] is $v_A + w_A$ [$v_B + w_B$] since he diverts all profits to himself and cares about his own personal benefits. If, on the other hand, the two units are integrated, then, assuming an outside manager is the boss, her payoff will be $v_A + v_B$ since she diverts all profits to herself and does not care about any private benefits. HH use as a benchmark social surplus equal to $v_A + v_B + w_A + w_B$.

HH note that the key point is that under integration less weight is placed on private benefits than under non-integration. Under non-integration one of w_A, w_B appears in each of the boss's objective function while under integration neither w appears in the overall objective function. Offsetting this under appreciation of the importance of the private benefits is the fact that under integration, total profits, and not individual unit profits are maximised.

So far the effects of deadweight losses, due to shading, have been ignored. Taking these into account complicates the analysis somewhat. Shading does bring about some internalisation of externalities: a boss will place some weight on the payoff of the other parties given that they can shade.

HH assume that the ability to shade under non-integration depends on the relationship that exists between the two parties. If the relationship is one of “non-integration without cooperation” then shading is not possible. This is because the relationship between the units is limited and it will be terminated should non-coordination occur. In the case of “non-integration with cooperation” the relationship continues and shading can take place even under non-cooperation. Under integration HH suppose that shading is always possible.

In summary HH write that “[...] under nonintegration, bosses have the right balance between private benefits and profits, but are parochial (they do not take into account their effect on the other unit), whereas, under integration, they have the right balance between units, but ignore private benefits. In our model, where the only issue is whether the units coordinate, we show that nonintegration and integration make opposite kinds of mistakes. Nonintegration can lead to

too little coordination when the benefits from coordination are unevenly divided across the units. One unit may then veto coordination even though it is collectively beneficial. In contrast, under a weak assumption—specifically, that coordination represents a reduction in “independence” and therefore causes a fall in private benefits—integration leads to too much coordination.” (Hart and Holmström 2010: 487-8).

The HH model considers two units A and B that have a lateral relationship. The units operate in the same input or output markets. Each unit has no workers but does have a manager. A decision has to be made by each unit that affects the other unit. The strategic coordination decision is modelled as a binary “Yes” (Y), “No” (N) choice. HH assume there are two aggregate outcomes which they refer to as “coordination” or “non-coordination”. Coordination will occur if and only if both units choose Y.

The time line is given in Figure 23.1. At the start an organisation form is chosen, then each unit gets to choose Y or N and lastly the payoffs are realised. The choice of organisation form is between having the two units as separate firms (non-integration, i.e. there are two bosses) or having just one firm with each unit being a division of that firm (integration, i.e. there is one boss).

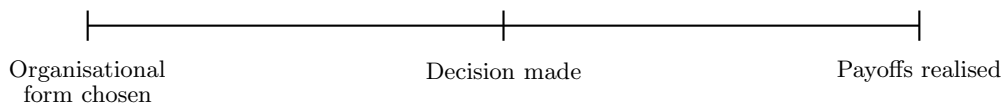


Figure 23.1. Time line.

There are two types of payoffs generated within each unit: a monetary profit, v , and private (and non-transferable) benefits, w . w are a (monetary) measure of the job satisfaction of the manager of each unit. In Table 23.1 Unit A is the row player and Unit B is the column player and v represents profit and w the private benefits. Subscripts denote the unit concerned.

		Unit B	
		Y	N
Unit	Y	A: $\Delta v_A, \Delta w_A$	A: 0,0
		B: $\Delta v_B, \Delta w_B$	B: 0,0
A	N	A: 0,0	A: 0,0
		B: 0,0	B: 0,0

Table 23.1. Payoff matrix.

Importantly HH assume that the boss of a unit can divert all profits generate by that unit to

herself. Private benefits, on the other hand, must always remain with the manager. HH use Table 23.1 to represent the payoffs from the different outcomes. HH state that they assume that payoffs are non-transferable and perfectly certain. They normalise so that the monetary profits and private benefits under non-coordination are zero for both units.

HH now introduce some additional notation:

$$\Delta z_A \equiv \Delta v_A + \Delta w_A, \quad \Delta z_B \equiv \Delta v_B + \Delta w_B \quad (26)$$

Here Δz_A [Δz_B] represents the change in total surplus, profits plus private surplus, in unit A [unit B] from a move to coordination and $\Delta z_A + \Delta z_B$ equals the change in aggregate total surplus. Note that equation (26) does not take into account the costs due to aggrivement, which depends on the *ex ante* contract in addition to the *ex post* decision.

As noted above the private benefits can be seen as a measure of the job satisfaction or on-the-job consumption. HH suppose that part of job satisfaction is the ability to take an independent course of action and thus assume that coordination results in a reduction in private benefits. That is,

$$\Delta w_A \leq 0, \quad \Delta w_B \leq 0. \quad (27)$$

They place no restrictions on whether profits are greater or lower under coordination and, further more, they make no assumption as to whether even if profits are increased by coordination that increase is greater than the fall in private benefits.

HH's focus is on two main organisational forms:

1. Non-integration: Manager A is the boss of unit A and manager B is the boss of unit B with the payoff to manager A being $v_A + w_A$ and that of manager B being $v_B + w_B$. Each manager diverts all his units profits to himself and receives the private benefits from his unit.
2. Integration: An outsider manager is the boss of both of both units and managers A and B are subordinates. The boss receives $v_A + v_B$. The unit managers' remuneration amounts to a fixed wage and the private benefits of his unit.

Organisational form and contracts are determined *ex ante*. At this stage the coordination decision are assumed to be so complicated that they can not be contracted upon. However authority to make these decisions can be allocated. It is assumed that the boss of each unit has the residual rights of control and thus has the authority to take the Y/N decision for her unit.

These *Y/N* decisions can be contracted upon *ex post*. Each of the unit managers chooses *Y* or *N* for his unit under non-integration while the overall boss instructs the two unit managers as to which of the two options they should choose under integration. HH suppose that the unit managers must follow the boss's commands as they are contractible but they also assume that the managers can shade. Shading can also take place under non-integration.

Specifically HH assume that each party feels entitled to their most preferred outcome or the best decision allowed by the contract and any party which receives k_i less than that outcome will feel aggrieved by k_i . This aggrievement will result in shading of θk_i . Here θ is an exogenous shading parameter which is assumed to be the same for all parties and $0 < \theta < 1$. The total deadweight loss from shading is therefore $\theta \sum_i k_i$.

Again there is a trade-off between a rigid contract with little shading and a flexible contract which can adjust to circumstances but with the possibility of greater shading. In the HH model there is no payoff uncertainty, however, their assumption that decisions are contractible only *ex post* means that changes in circumstances makes the *ex ante* choice of organisational form relevant for the deadweight losses that result from aggrievement.

The ability to shade can depend upon the nature of the transaction that the party is involved in. Under “non-integration without cooperation” the relationship between the parties ends if cooperation does not take place and thus the shading possibilities are greatly reduced. Under “non-integration with cooperation”, the broader relationship between the parties continues even if coordination fails and thus shading is possible. Under integration HH suppose that shading is always possible.

Given this framework *ex post* renegotiation is not costless. This follows because under the shading assumption each party will feel entitled to the best outcome in the renegotiation and all parties can not achieve that and thus some will shade. In addition, insofar as renegotiation revisits the terms and entitlements underlying existing contracts, renegotiation could make all the parties worse off. HH rules out *ex post* renegotiation on these grounds.

HH assume that bargaining at the *ex ante* stage results in the organisational that is chosen being the one which maximises expected future surplus net of *ex post* shading costs.¹⁴ At the *ex ante* stage it is assumed that at least one side of the market is competitive so that each side achieves the best outcome available in the negotiations. This means that there is no shading at the *ex ante* stage. At the *ex post* stage, the parties are locked in and thus there is the possibility of shading.

¹⁴Lump-sum transfers are assumed to be available to redistribute surplus.

Bargaining at the *ex ante* stage also determines the wages of the managers. In the special case where the *ex ante* market for managers is competitive the wage for a manager plus the expected private benefits will equal the reservation utility for the manager. An implication of this is that should an organisational change reduce private benefits, there will be a compensating increase in the wage.

HH compare three organisational forms: “non-integration without cooperation”, “non-integration with cooperation” and “integration”. There is no renegotiation.

HH denote the social surplus, net of shading costs, by S but refer to S simply as social surplus. More precisely, S is the relevant payoff from Table 23.1 less any shading costs. *First-best* refers to the cases where aggregate surplus is maximised and the costs of shading are zero. *First-best efficient* refers to situations where the decision maximises total surplus ignoring shading costs.

Non-integration without coordination: under non-integration the payoff for manager A [B] is $v_A + w_A$ [$v_B + w_B$] and either manager can veto coordination by selecting N . HH deal with three cases.

Case 1: $\Delta z_A \leq 0$, $\Delta z_B \leq 0$. Here the change in total surplus for both managers is negative (or zero) and thus the managers’ preferences are aligned, neither of them wants coordination. Given this, there is no disagreement and thus no aggrievement. Social surplus is given by

$$S = 0. \tag{28}$$

Case 2: $\Delta z_A \geq 0$, $\Delta z_B \geq 0$. Again the managers’ preferences are aligned, both parties want coordination and thus we get coordination without aggrievement. The social surplus is given by

$$S = \Delta z_A + \Delta z_B. \tag{29}$$

Case 3: $\Delta z_i < 0$, $\Delta z_j > 0$ ($i \neq j$). In this case the managers are in conflict, manager i doesn’t want coordination and can veto it via his choice of N . Remember that under “non-integration without cooperation” manager j can not shade since the parties do not have an on going relationship. Given this, manager i will use his veto and thus the outcome will be non-coordination.

$$S = 0. \tag{30}$$

We see that the first-best occurs in Case 1, non-coordination gives zero which is the best payoff the managers can get and there is no shading, and in Case 2. You get coordination if and

only if

$$\Delta z_A + \Delta z_B \geq 0. \quad (31)$$

that is, the change in aggregate total surplus is non-negative. Case 2 implies equation (31) and given equation (31) the only way you can get first-best is with $z_i \geq 0$, $i = A, B$. In Case 3 the first-best may not occur, it depends on who the winners and losers are. Even if aggregate surplus increases, the distributions of the gains may be such that one of the parties loses out and thus will veto coordination.

In short: there is too little coordination under “non-integration without cooperation”. Whenever coordination occurs it is first-best efficient, that is, Case 2 implies (31); but coordination may not occur when it is first-best efficient, that is, (31) does not imply Case 2. Last, note that there is no shading in equilibrium under “non-integration without cooperation”, whether the outcome is coordination or non-coordination.

Non-integration with coordination: The fact that cooperation will continue even if cooperation doesn’t means that shading can occur even if non-cooperation is the outcome chosen. This clearly will not effect Case 1 or Case 2 as no shading takes place under either of these outcomes. Both are still first-best. Things can change, however, for Case 3. In particular, manager i may not choose to veto coordination because manager j will be aggrieved and thus may shade. The level of grievement will be the difference between manager j ’s payoff under his most preferred outcome, that is coordination, and what he actually gets. Shading by j will be proportional to this. Thus j will be aggrieved by Δz_j and will therefore shade $\theta \Delta z_j$. Coordination will occur if manager i ’s loss in total surplus under coordination is less than the cost of shading imposed on i by manager j , that is, $\Delta z_i \geq -\theta \Delta z_j$ which implies

$$\Delta z_i + \theta \Delta z_j \geq 0. \quad (32)$$

If equation (32) does hold, manager i coordinates with j but only reluctantly, and will be aggrieved by amount $-\Delta z_i$ since his most preferred outcome is not to coordinate. Shading by i will therefore be $-\theta \Delta z_i$ which results in a deadweight loss of this amount. Note that equation (32) implies

$$\Delta z_j + \theta \Delta z_i > 0. \quad (33)$$

This follows since

$$\Delta z_j + \theta \Delta z_i > \Delta z_i + \theta \Delta z_j \geq 0$$

which holds because $\Delta z_j > \theta \Delta z_j$ and $\theta \Delta z_i$ is less negative than Δz_i , remembering that $\Delta z_i < 0$ and $1 > \theta > 0$. Given (33) manager j is willing to coordinate in spite of the shading by i . If, however, equation (32) does not hold coordination does not occur and j shades by an amount $\theta \Delta z_j$.

Here social surplus is either,

$$\begin{aligned} S &= \Delta z_A + \Delta z_B + \theta \Delta z_i && \text{if (32) holds (that is, coordination)} \\ \text{or } &= -\theta \Delta z_j && \text{if (32) does not hold (that is, non-coordination)} \end{aligned} \tag{34}$$

So we end up with first-best being achieved in Cases 1 and 2 but not Case 3. Note that (32) implies (31) since $\Delta z_A + \Delta z_B > \Delta z_A + \theta \Delta z_B = \Delta z_i + \theta \Delta z_j \geq 0$ and so if (32) holds the condition for first-best coordination also holds, but (32) doesn't always hold and therefore there is too little coordination compared to the first-best. Also the social surplus given by (34) always involves a strictly positive shading cost, so no matter the decision made one side or the other will be unhappy.

HH sum up by noting that “[i]t is evident that “non-integration with cooperation” is potentially desirable (to the extent that it is a choice) only if coordination is the outcome (i.e., [(32)] holds. When [(32)] does not hold, the parties are better off with “non-integration without cooperation”. In the case where there is uncertainty (to be discussed later) it is possible that parties attempt “non-integration with cooperation”, only to find that [(32)] fails.” (Hart and Holmström 2010: 496-7).

Integration: HH divide their analysis of integration into two cases. Remember equation (27): $\Delta w_A \leq 0, \Delta w_B \leq 0$.

Case 1: $\Delta v_A + \Delta v_B \leq 0$. Given equation (27) the managers' and the boss's preferences are aligned. In this case coordination does not occur since neither the boss nor the managers want it. Thus there is no aggrievement and no shading. Social surplus is

$$S = 0. \tag{35}$$

Case 2: $\Delta v_A + \Delta v_B > 0$. Here there is conflict since the boss wants coordination but the managers do not. As the managers do not want coordination they will be aggrieved by an amount $\Delta w_A + \Delta w_B$ and thus will shade by $\theta(\Delta w_A + \Delta w_B)$ if it occurs. The boss will therefore coordinate

if and only if her payoff net of shading is greater than zero (her payoff from non-coordination).

$$\Delta v_A + \Delta v_B + \theta(\Delta w_A + \Delta w_B) \geq 0. \quad (36)$$

Here we see that the boss will partially internalise the preferences of her managers. If (36) does not hold then the boss will not coordinate and the managers get their preferred outcome. However if this is the case then the boss is not getting her most preferred outcome and thus she will be aggrieved, by $\Delta v_A + \Delta v_B$, and will shade by $\theta(\Delta v_A + \Delta v_B)$ thereby reducing the managers' payoffs by this amount.

The social surplus will be,

$$\begin{aligned} S &= \Delta z_A + \Delta z_B + \theta(\Delta w_A + \Delta w_B) \quad \text{if (36) holds (that is, coordination)} \\ \text{or} &= -\theta(\Delta v_A + \Delta v_B) \quad \text{if (36) does not hold (that is, non-coordination)} \end{aligned} \quad (37)$$

Case 1 results in the first-best outcome while Case 2 does not. For Case 2 there is too much coordination relative to the first-best since (31) implies (36) but not vice versa. We can see this from the fact that $\Delta v_A + \Delta v_B + \theta(\Delta w_A + \Delta w_B) > \Delta z_A + \Delta z_B = \Delta v_A + \Delta v_B + (\Delta w_A + \Delta w_B) \geq 0$ since $\Delta w_i \leq 0$, $i = A, B$ and $1 > \theta > 0$.

HH then state that they have established Proposition 4.

Proposition 4 *Nonintegration errs on the side of too little coordination (when coordination occurs it is first-best efficient, but it may be first-best efficient and not occur), whereas integration errs on the side of too much coordination (when coordination is first-best efficient it occurs, but it may occur even when it is not first-best efficient). If noncoordination is first-best efficient, “nonintegration without cooperation” achieves the first-best. If coordination is first-best efficient then (a) integration leads to coordination, but may not be optimal given the deadweight losses from shading; (b) integration is optimal if the changes in private benefits from coordination are sufficiently small; and (c) integration is uniquely optimal if in addition the distribution of profits is sufficiently uneven. (Hart and Holmström 2010: 497-8).*

4.1 a digression on profit maximisation

In Section 6 of their paper HH argue that the firm in their model may not maximise profits because the boss could care about the non-transferable private benefits or because they may be forced to internalise them given that employees can shade. If we think of the private benefits as

on-the-job consumption then the HH model can be contrasted with the model of Demsetz (1995: First commentary). In the Demsetz model the firm produces only for those outside the firm and not for internal (on-the-job) consumption and both profit maximisation and utility maximisation are achieved, with the former being implied by the latter. The HH model, on the other hand, allows for on-the-job consumption, but neither profit maximisation nor utility maximisation can be guaranteed.

But it should be noted that if money profits are less due to the boss caring directly about the non-transferable private benefits of employees and the boss is also the owner then it is not clear that this is not profit maximising behaviour.¹⁵ What this means is that part of the profits realised by the firm are being taken by the owner in the form of the transfer to employees via their private benefits. Assuming the owner is acting rationally, it must be that the value to the owner of the transfer is greater than the money profits sacrificed. That is, the sum of the value of the transfer plus the lower money profits is larger than the value of the sum of no transfer and a greater money profits. One way to view this is as the owner maximising the money profits of the firm and then spending some of this profit, which is now part of his personal wealth, to pay for the transfer to his employees. Viewed in this way, the transfer is just on-the-job consumption by the owner and this doesn't diminish the usefulness of assuming that owners seek to maximise profits. In effect on-the-job consumption is being secured by using earned (maximised) profits in-house. If the boss is not the owner then the boss caring directly about the private benefits of employees raises questions about agency problems between the owner and the boss.

5 reference points, property rights and transaction costs

Hart (2008: 406) argues that shading costs are akin to “haggling costs”. The modelling of haggling costs can be seen as a move towards the modelling (however imperfectly) of transaction costs. Hart and Moore (2008: 4-5) argue that “[...] the costs of flexibility that we focus on—shading costs—can be viewed as a shorthand for other kinds of transaction costs, such as rent-seeking, influence, and haggling costs.” Exactly how similar the reference point and transaction-cost explanations are is, however, open to debate. There is also the question of the relationship between these two approaches to the firm and the property rights approach.

In a discussion of the differences between the Grossman-Hart-Moore (GHM) theory of the firm and the transaction-cost approach, Williamson (2000: 605-6) argues that the most important

¹⁵Obviously no assumption that production is completely specialised for sale is being made.

difference between them is that GHM introduce inefficiencies at the ex ante investment stage while the transaction-cost approach emphasises that ex post haggling and maladaptation drive inefficiencies.¹⁶ There are no ex post inefficiencies in GHM due to their assumption of common knowledge and ex post costless bargaining. Gibbons (2010: 283) explains it this way:

“[t]he model in question is Grossman and Hart’s (1986), which explores an alternative to Williamson’s (2000, p. 605) emphasis that “maladaptation in the contract execution interval is the principal source of inefficiency”. Instead, in the Grossman-Hart model, there is zero maladaptation in the contract execution interval, and the sole inefficiency is in endogenous specific investments.

It is striking how different the logic of inefficient investment can be from the logic of inefficient haggling. In their pure forms envisioned here, the two can be seen as complements. For example, the lock-in necessary for Williamson’s focus on inefficient haggling could result from contractible specific investments chosen at efficient levels. But by assuming efficient bargaining and hence zero maladaptation in the contract execution interval, Grossman and Hart focused attention on non-contractible specific investments and hence discovered an important new determinant of the make-or-buy decision: in the Grossman-Hart model, an important benefit of non-integration is that both parties have incentives to invest; in Williamson’s argument, an important cost of non-integration is inefficient haggling. In short, the two theories are simply different.”

This emphasis on ex post haggling and maladaptation can be interpreted as reflecting a view that internal organisation is better at reconciling the conflicting interest of the parties to a transaction and facilitating adaptation to changing supply and demand conditions when such cost are high.

The reference point approach can be seen as a movement away from the ex ante GHM approach and back towards transaction cost thinking in so much as contracting is not perfectly contractible ex post. This fact, as Hart (2008: 294) points out “[...] is a significant departure from the standard contracting literature. The literature usually assumes that trade is perfectly enforceable ex post (e.g. by a court of law). Here we are assuming that only perfunctory performance can be enforced: consummate performance is always discretionary”, and thus inefficiencies can arise ex post. The development of a tractable model of contracts and organisational form that exhibits ex post inefficiency is one of motivations for advancing the reference point approach in the first place.

¹⁶Holmström and Roberts (1998: 75-9) also highlights the distinctions between the transaction cost and property rights theories. Whinston (2003) looks at the empirical differences between the two theories.

(Hart and Moore 2008: 4). Hart’s interpretation of the reference point theory is “[i]n a sense, this work can be viewed as a “merger” of the transaction cost and property rights literatures.” (Hart 2011b: 106).

Also the reference point approach differs from the property rights theory in that it does not require an assumption of the use of relationship specific investments as is standard in the property rights theory. Relationship specific investments can be introduced to the reference point theory, Hart (2011a) is an example where this is done, but, in general, the reference point theory does not rely on such investment.

The reference point approach also highlights the importance of Williamson’s notion of the “fundamental transformation”.¹⁷ Hart and Moore argue that the move from an ex ante competitive market to an ex post bilateral setting—what Williamson (1985: 61-3) terms the fundamental transformation—provides a rationale for the idea that contracts are reference points. “A competitive ex ante market adds objectivity to the terms of the contract because the market defines what each party brings to the relationship. HM assume that the parties perceive a competitive outcome as justified and accept it as a salient reference point.” (Fehr, Hart and Zehnder 2009: 562). This is an idea which finds experimental support: see Fehr, Hart and Zehnder (2009), Fehr, Hart and Zehnder (2011) and Hoppe and Schmitz (2011).

But we must also be aware that important features of the transaction-cost theory may still have been left out. How fully shading costs capture the costs of ex post maladaptation and haggling is an open question. When discussing some opportunities for the future of transaction-cost economics, Robert Gibbons (2010: 283) notes that “[...] it may be that Hart and Moore’s (2008) “reference points” approach is a productive path. Time will tell [...]”. Hart (2011b: 106) concludes “[w]hether this merger [resulting in the reference point theory] will be successful remains to be seen.”

6 conclusion

The papers discussed here have, on the positive side, developed a new framework within which to study the theory of the firm based on the ‘reference point’ approach to the study of (incom-

¹⁷Given the importance of fundamental transformation to the analysis of economic organisation Williamson (1985: 63) asks why this notion was ignored for so long. In footnote 23 he gives three reasons: “One explanation is that such transformations do not occur in the context of comprehensive, once-for-all contract—which is a convenient and sometimes productive contracting fiction but imposes inordinate demands on limited rationality. A second reason is that the transformation will not arise in the absence of opportunism—which is a condition that economists have been loath to concede. Third, even if bounded rationality and opportunism are conceded, the fundamental transformation appears only in conjunction with an asset specificity condition, which is a contracting feature that has only recently been explicated.”

plete) contracts. They have demonstrated that the trade-off between contractual rigidity and flexibility has important implications for the organisation of firms. The unifying feature shared by the papers summarised above is the application of the idea that contracts act as a reference point for feelings of aggrievement and thus acts of shading. The Hart and Moore (2008) theory and its extensions provide an explanation for the existence of long-term contracts in the absence of relationship specific investments, which are assumed in most of the incomplete contracts approaches to the firm. Also the reference point theory can shed new light on the roles of the employment relationship and authority. In work extending the theory, Hart (2009) reintroduces assets into the model and shows that previously hard to explain observations in the empirical literature on contracting and integration can be explained by the reference point approach. Hart and Holmström (2010) offers a theory of firm scope. They provide an analysis that moves the focus of the theory away from the role of non-human assets in determining a firm's boundaries towards a theory where the activities undertaken by the firm determine the firm's scope.

On the negative side the Hart and Moore theory, and thus the papers surveyed here, rests on strong (ad-hoc) behavioural assumptions which have only limited experimental support and, thus far, no direct empirical backing. There are concepts in the behavioural literature, such as reference dependent preferences, the self-serving bias or reciprocity, which are broadly consistent with the Hart-Moore model but, as already noted, the experimental/empirical support for the model is, at best, limited. An important topic for future work in the reference point literature on contract theory is to show that the ad-hoc nature of the behavioural assumptions used within it are consistent with utility-maximising behaviour.

Integration is a central topic in the theory of the firm. As outlined above in the reference point literature the papers Hart and Moore (2007) and Hart and Holmström (2010) both deal with the issue. For Hart and Moore integration is vertical and can be seen as a way of ensuring supply under conditions of uncertainty. For Hart and Holmström the relationship between the agents is lateral with the coordination of the activities of the parties being of prime importance and here integration can be optimal if, for example, the changes in private benefits from coordination are small, but can also result in too much coordination. So the motivations for integration differ depending on the nature of the relationship between the agents, i.e. is the relationship vertical or lateral, but in either case integration can be efficiency enhancing.

Hart and Moore (2007) argue that one failing of the property rights theory is its reliance on noncontractible, relationship-specific investments which are hard to measure empirically. But it is not clear that aggrievement and shading present any less in the way of measurement issues for

empirical work. So the reference point approach may not be an improvement in terms of the ease of empirical evaluation.

In addition, on the theory side, an assumption made with regard to shading is that the reason for aggrievement does not affect the amount of shading that takes place, clearly it could. If one party thinks another agent's behaviour is opportunistic they may react differently than if they feel the other agent's action is the result of external factors. The first case could result in more shading than the second. But such concerns over the reasons for aggrievement may have wider application. For example, the choice of price by B in the Hart and Moore (2007) model may be seen by S as self-interested behaviour, but not truly opportunistic, where as the deliberate decision to attempt "hold-up" in the Hart (2009) model could be more readily seen as opportunistic behaviour and thus punished by a greater amount of shading.

Also the theory has an inherent human capital bias. The reliance on aggrievement and shading could be seen as limiting the applicability of the theory to areas relatively dependent on human capital. In firms in which production is mainly dependent on non-human capital, which cannot be aggrieved and cannot shade, the theory may be of less value. Also for areas where it is possible to write contracts that cover most of, if not all of, the relevant actions - thereby reducing the likelihood of aggrievement and the ability to shade if aggrievement does occur - the theory would be of more limited usefulness. In other words the theory is less applicable to situations where the set of actions which defines consummate, as opposed to perfunctory, performance is small. Thus the theory seems to have greater potential when applied to firms who have a greater dependence on human capital and where monitoring is ineffective.

This human capital bias is important, for example, when Hart and Holmström (2010: 511) state that "giving private benefits a pivotal role in the analysis moves the focus of attention away from assets toward activities in the determination of firm boundaries." More properly they have moved the focus of attention away from asset ownership to human capital utilisation. They tie the 'unit' to the manager in such a way that an expansion of activities requires the addition of extra managers (human capital). What happens if, for example, the activities the firm undertakes can be expanded by simply expanding the range of physical capital the firm employs? There is no basis for an increase in the level of aggrievement and shading and therefore Hart and Holmström's reference point model would say the boundaries of the firm would not have changed but the firms activities have increased. What Hart and Holmström exploit is an implied positive correlation between the range of human capital a firm employs and the range of its activities. While theories such as the property rights approach to the firm may go too far in defining firms solely in terms

of asset ownership, the reference point theories may go too far in the opposite direction by over emphasising human capital.

Avenues for future work on the reference point approach to the firm would include the re-introduction of non-contractible investments into models of the firm. A paper related to the reference point approach to the firm literature, which analyses non contractible investments in a model with shading, is Hart (2011a). This paper studies a model where a long-term contract is utilised to encourage a seller to make quality-enhancing investments in addition to trying to achieve ex post efficiency and avoid shading. Each party has an outside option which depends on the assets they own. Having the seller make a take-it-or-leave-it offer is optimal when there is no contract renegotiation and no shading but when shading can occur such a contract will result in deadweight losses. Here the optimal contract will place limits on the seller's offers and can result in ex post inefficiencies. Asset ownership can reduce inefficiencies even when revelation mechanisms are used. As already noted asset ownership also reduced inefficiencies in Hart (2009) but in this paper there is no investment and the allocation of assets was used to avoid ex post hold-up. Such hold-up was the cause of any aggrivement and thus shading. Asset ownership in this paper should be allocated to those whose payoffs are relatively uncertain and to those for whom the asset is idiosyncratic. In Hart (2011a) seller incentives replace uncertainty as the dominate reason for asset ownership as investment incentives are crucial and this drives us in the direction of seller ownership of assets. Idiosyncratic assets are again important but one major difference between Hart (2011a) and Hart (2009) is that although in Hart (2011a) the seller should own assets idiosyncratic to her, the buyer should not necessarily own assets idiosyncratic to him since this can interfere with the seller's investment incentives. An obvious extension to these ideas is an application directly to the theory of the firm. Such an extension would move the reference point approach back towards the property rights framework.

Like the property rights approach to the firm in the reference point approach, firms or divisions within firms, are fundamentally individuals and thus the same question can be asked of the reference point framework as of the property rights approach, Can it be applied to more sophisticated organisational firms? Questions also have to be asked about where does a reference point come from and how can we pin it down? Do reference points have to be the same at the contractual performance stage as they were at the contractual negotiation stage? If they can change, when, how and why? Answering such questions will provide the basis for a fruitful and exciting future for the reference point approach to the theory of the firm.

7 References

- Aghion, Philippe and Richard Holden (2011). ‘Incomplete Contracts and the Theory of the Firm: What Have We Learned Over the Past 25 Years?’, *Journal of Economic Perspectives*, 25(2) Spring: 181-97.
- Demsetz, Harold (1995). *The Economics of the Business Firm: Seven Critical Commentaries*, Cambridge: Cambridge University Press.
- Fehr, Ernst, Oliver D. Hart and Christian Zehnder (2009). ‘Contracts, Reference Points, and Competition – Behavioral Effects of The Fundamental Transformation’, *Journal of the European Economic Association*, 7(2-3) April-May: 561-72.
- Fehr, Ernst, Oliver D. Hart, and Christian Zehnder (2011). ‘Contracts as Reference Points – Experimental Evidence’, *American Economic Review*, 101(2) April: 493-525.
- Foss, Nicolai J., Henrik Lando and Steen Thomsen (2000). ‘The Theory of the Firm’. In Boudewijn Bouckaert and Gerrit De Geest (eds.), *Encyclopedia of Law and Economics*, Volume III, Cheltenham U.K.: Edward Elgar Publishing Ltd.
- Gibbons, Robert (2005). ‘Four formal(izable) theories of the firm?’, *Journal of Economic Behavior and Organization*, 58(2): 200-45.
- Gibbons, Robert (2010). ‘Transaction-Cost Economics: Past, Present, and Future?’, *Scandinavian Journal of Economics*, 112(2): 263-88.
- Grossman, Sanford J. and Oliver D. Hart (1986). ‘The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration’, *Journal of Political Economy*, 94(4): 691-719.
- Hart, Oliver D. (1989). ‘An Economist’s Perspective on the Theory of the Firm’, *Columbia Law Review*, 89(7) November: 1757-74.
- Hart, Oliver D. (1995). *Firms, Contracts, and Financial Structure*, Oxford: Oxford University Press.
- Hart, Oliver D. (2008). ‘Economica Coase Lecture: Reference Points and the Theory of the Firm’, *Economica*, 75(299) August: 404-11.
- Hart, Oliver D. (2009). ‘Hold-up, Asset Ownership, and Reference Points’, *Quarterly Journal of Economics*, 124(1) February: 267-300.
- Hart, Oliver D. (2011a). ‘Noncontractible Investments and Reference Points’, *National Bureau of Economic Research*, NBER Working Papers: 16929.
- Hart, Oliver D. (2011b). ‘Thinking about the Firm: A Review of Daniel Spulber’s The Theory of the Firm’, *Journal of Economic Literature*, 49(1) March: 101-13.
- Hart, Oliver D. and Bengt Holmström (2010). ‘Contracts as Reference Points’, *Quarterly Journal of Economics*, 125(2) May: 483-513.
- Hart, Oliver D. and John Moore (1990). ‘Property Rights and the Nature of the Firm’, *Journal of Political Economy*, 98(6): 1119-58.
- Hart, Oliver D. and John Moore (2007). ‘Incomplete Contracts and Ownership: Some New Thoughts’, *American Economic Review*, 97(2) May: 182-6.
- Hart, Oliver D. and John Moore (2008). ‘Contracts as Reference Points’, *Quarterly Journal of Economics*, 123(1) February: 1-48.
- Holmström, Bengt and John Roberts (1998). ‘Boundaries of the Firm Revisited’, *Journal of Economic Perspectives*, 12(4) Fall: 73-94.

- Holmström, Bengt and Jean Tirole (1989). 'The Theory of the Firm'. In Richard Schmalensee and Robert D. Willig (eds.), *Handbook of Industrial Organization* Vol. I, Amsterdam: North-Holland.
- Hoppe, Eva I. and Patrick W. Schmitz (2011). 'Can contracts solve the hold-up problem? Experimental evidence', *Games and Economic Behavior*, 73(1) September:186-99.
- Maskin, Eric (2002). 'On indescribable contingencies and incomplete contracts', *European Economic Review*, 46(4-5) May: 725-33.
- Maskin, Eric and Jean Tirole (1999). 'Unforeseen Contingencies and Incomplete Contracts', *Review of Economic Studies*, 66(1) January: 83-114.
- Tirole, Jean (1988). *The Theory of Industrial Organization*, Cambridge Mass.: MIT Press.
- Whinston, Michael D. (2003). 'On the Transaction Cost Determinants of Vertical Integration', *Journal of Law, Economics, and Organization*, 19(1)April: 1-23.
- Williamson, Oliver E. (1975). *Markets and Hierarchies: Analysis and Antitrust Implications*, New York: The Free Press.
- Williamson, Oliver E. (1985). *The Economic Institutions of Capitalism*, New York: The Free Press.
- Williamson, Oliver E. (1996). *The Mechanisms of Governance*, Oxford: Oxford University Press.
- Williamson, Oliver E. (2000). 'The New Institutional Economics: Taking Stock, Looking Ahead', *Journal of Economic Literature*, 38(3) September: 595-613.