Boundaries of the firm and intellectual production - wrestling with Transaction Costs Economics

Jan Vang-Lauridsen

Department of Industrial Economics and Strategy
Copenhagen Business School
Howitzvej 60
2000 Frederiksberg
Denmark
jbvl.ivs@cbs.dk

December 2000

Abstract

In this paper I explore the relevancy of the transaction cost economics (TCE) framework, represented by Coase and Williamson, for explaining the boundaries of the firm in intellectual production. I argue that at least when it comes to intellectual production the way Williamson and Coase understands the working of the price mechanism and market is not sufficient, which in turn has consequences for when a firm should 'make or buy' a good or service. Moreover, I point to that the potentials for opportunism is not reduced by vertical integration when it comes to intellectual production since intellectual production is difficult to monitor. On the other hand central concepts as opportunism, bounded rationality and economising are useful analytical concepts for intellectual production.

This is a very preliminary version, hence no quoting without permission from the author, please.

1. Introduction

Transaction cost economics - that has largely been developed by Coase (1937, 1960), Williamson (1975, 1979, 1985 and 1996), and Demsetz (1972 1994a, 1994b) - tries to answer three questions, namely why do firm exist, how are they organised internally and what determines their boundaries? TCE has despite a minor set back in the 90'ies been setting the research agenda within this field, later in competition with the resource-based view of the firm (including the knowledge, capability and competence view of the firm) and evolutionary theory – that often overlaps with the resource-based view of the firm – but includes the natural selection process (see Winter and Nelson (1982), Foss 1993 and Dosi and Teece 1993). Unfortunately the latter two approaches has to a far to large extent ignored the insights or been hostile to TCE (Foss and Foss 2000). According to RBV focus should be on resources, capabilities, knowledge, learning and bounded rationality while TCE so far has focused on opportunism, bounded rationality and incentives. Moreover, one can say that TCE primarily has focused on the commitment aspects of economic organisation while RBV has focused on co-ordination of production, not the commitment (as will be clear below this doesn't hold for Coase).

In this paper I hope to show that TCE holds central insights that are relevant to understand organisation of intellectual production. And just because certain hard criticisms are raised against TCE it doesn't mean that equally hard criticism couldn't be raised against RBV. RBV suffers from a lack definitional clarity, say when it comes to what a capability is, when it is similar and dissimilar, and so on.

In the paper – that is mainly conceptual in character – I focus on the question of the boundaries of the firm and set out to explore how far TCE brings us in explaining the boundaries of the firm in intellectual production. This focus is motivated by the fact that most empirical studies within TCE (and RBV and evolutionary economy in general) has been concerned with physical production. And as I explain below physical production is distinctively different from intellectual production.

The intention with the paper is less to contribute to a theoretical development than it is to unpack the potentials and limitations in the TCE since the following reflections consists of my first meeting with the field. I do not intent to cover all aspects the boundary question but focuses instead on the assumptions for explaining the boundaries of the firm, the role of information and foresight, and the way markets are conceptualised. These

concepts are both central to explaining the boundaries of the firm in general and maybe - especially information asymmetry and foresight - even more important for intellectual production than for physical production. This is due to that the element of tacit knowledge usually is larger in intellectual production than in physical, that human resources are not just an add on to machines, that the need for speed of innovations are higher, that products are often not even developed when 'ordered', that it is more difficult to measure efficiency / quality of an intellectual product than an physical also ex post, and finally that the degree of asymmetry in intellectual production is often larger than in physical production. Let me just bring a short example. If a firm buys paper for producing paper-plates, then it is rather easy to judge the quality of the product. This is also the case in some kind of intellectual production such as measuring the impact of a new processor in a computer in terms of for example speed of information processing. But other kinds of intellectual products quality are more difficult to measure, say measuring the quality of a newspaper article submitted by a journalist. It can of course be measured in terms of spelling mistakes, amount of sources that is quoted, an so on. But the editor doesn't know the articles that weren't written or could have been written – the journalist is usually the specialist in the field – and hence has often limited possibilities for actually judging the quality.

To deal with the boundary-question in intellectual production and TCE I address three questions. First and briefly, what is the difference between physical and intellectual production? Second, how does TCE explain boundaries of the firm? Third, are these explanations useful for intellectual production? Throughout the text, I will illustrate central points and arguments with examples from the newspaper industry. The newspaper industry is a sub-branch of intellectual production belonging to the more routine based part. The illustrations are illustrations only, not full prove empirical documentation.

Before engaging in the more theoretical discussion, it is needed to clarify what actually is meant by the concept 'boundary of the firm'. Originally, the distinction was made between inside (direction as the dominant mode of co-ordination) and outside (co-ordination through the use of the price mechanism). But today the firm within TCE is view as a nexus of contacts 'involving' many different hybrid forms, making the distinction between inside and outside rather blurred. Moreover, the same activity may be within the boundaries of one firm if one focuses on the co-ordination mechanism but within he boundaries of another firm if one looks at ownership (Grandori 2001). Following Alchian and Woodward's statement "The firm is dead, long live the firm" one can say that "The boundaries are dead, long live the boundaries", meaning that the question is equally relevant but even more complex to answer than before, theoretically speaking that is. In practical terms, however, the boundary of the firm deals with what kind of governance structure by which a certain activity is governed.

Intellectual versus physical production. The activities I zoom in on is the so-called intellectual production but what is that and how is it different from physical production? For at detailed discussion, see Mahnke (1999), here I will just deal with it in a very brief and stylised manner. Intellectual production, sometimes also labelled knowledge production is different from physical production in several ways, namely concerning input, processing and output. The main input is knowledge and/or information in one or several of its forms, being codified or tacit, being know(ing) that or know(ing) how, and so on. Moreover, knowledge is not a scarce recourse in the sense that once used it is gone.

Essentially the knowledge is embodied in humans either inside the firm or outside but symbolically speaking it can also be embedded in organisations (firms), in the sense that no one in the organisation possesses al the organisational knowledge. The layouter knows how to layout the article, the journalists how to research and write articles and the editor how the submitted articles are to supplement each other.

Knowledge is relatively less important in physical production where capital, labour or land are the most important inputs. This is of course differences that are expressed in an idealised way; in the dirty reality things are far more blurred.

The production process is also different since 'it relies strongly on interactive production and is characterised by reciprocal and team-interdependencies rather than purely sequential ones" (Mahnke 1999), as it is often the case for physical production.

Finally, the product consists of information and/or knowledge where the physical attributes plays only a minor role, the important content is the information or knowledge parts. The physical content of a newspaper – a few hundred gram of paper with print – can hardly be worth the price, what one pays for is of course the intellectual content, namely the articles.

2. TCE and the boundary-question

The boundary-question has been one of the predominant questions within TCE- the other two questions being the existence and internal organisation of the firm- hence there is an overwhelming large literature on the topic, both empirical and theoretical. I do not attempt in here to deal with this extensive literature. Instead, I focus on Coase (1937) and Williamson (1985, 1996) since they represent different end of 'scales' of the TCE-literature both in approach and in time.

Transparent markets

Coase (1937 (1991)) founded the approach with his seminal article 'The nature of the firm' where he focused on co-ordination (Williamson later focused

more on opportunism). But Coase and Williamson share the idea that firms exists because of market failures, hence the boundaries of the firm are related to the way the price mechanism and/or market works though. The natural place to start then is to ask if their way of conceptualising the price mechanism is adequate for intellectual production. When addressing the cost of using the price mechanism – the shared idea – I will focus on Coase only (1937 (1991)).

Coase's text is dense and hence open to different conflicting interpretations. He takes his point of departure in a critic of how the market or price mechanism works according to what he labels as mainstream economists. He himself quotes Salter as documentation for his points about how they conceive the working of the price mechanism.

"The normal economic system works itself. For it is under no central control, it needs no central survey. Over the whole range of human activity and human need, supply is adjusted to demand, and production to consumption, by a process that is automatic, elastic and responsive" (quoted from Coase 1937(1991), p 18).

Obviously, the content of this quote can be unpacked and criticised from many perspectives. It resembles what people within development studies in a slightly misguiding way refer to a flee market instead of free markets since all most any modern adequate definition of a free market would involve some kind of institutional setting, either in terms of legal regulation or regulation by conventions. Nevertheless, it is probably not wrong to say that it is quite close to the way a free market is stylised in neo-classical economy today where it still is seen as a self-working system that through the price mechanism assures the most efficient allocation of the scarce resources.

Coase concentrates his criticism on one points when he says that if the market was working in this way firms wouldn't exists since the market would be the most efficient co-ordination mechanism, and since he isn't blind he cannot ignore that they actually exist out in the real world (the contra-factual method). The unavoidable question is then, why do firms exist? And Coase explains this in terms of market failure or less ideologically stated the cost of using the price mechanism.

"The main reason why it is profitable to establish a firm would seem to be that there is a cost of using the price mechanism. The most obvious cost of "organizing" production through the price mechanism is that discovering what the relevant prices are … The costs of negotiation and concluding a separate contract for each transaction which takes place on a market must also be taken into account" (Coase 1937 (1991). P. 21)

Behind Coases conceptualisation of the cost of using the price mechanism lies two assumptions, namely that the market exists – this question I wont deal with - and that the goods exists and are comparable. The latter is often not the case for intellectual production. Since many 'intellectual products' are customised to one customer only, this puts some restrictions on how

standardised the product can be, hence how easily it is substituted by another product, which in turn reduces the competition and then the possibility of identifying a relevant (market) price. In this vein one can say market prices rests on assumptions on transparency due to product quality and quantity, this is often not the case when it comes to intellectual products. In intellectual production it is difficult to judge the value of a product due to the nature of the product; it is difficult to measure/evaluate the effect of a product, since intellectual product often functions in what Sayer (1984) terms open systems. That is systems that is exposed to casual mechanism from several structures, so to isolate the effect of one is difficult and sometimes impossible. There are of course techniques to evaluate the effects of different products but usually they are quit costly both in time and, often, the result is very imprecise. Moreover, it is difficult to compare the product ex ante since the product is often not even made or the buyer often doesn't really know exactly what he needs/wants. This is often the reason why firms contact other experts such as specialised journalists or a consultancy bureau. The relation between producer and user in such cases involves a high degree of information asymmetry in both directions. The consultant is the specialist in the field, for example developing organisational competencies, but usually only has a limited knowledge of the organisation he is now going to work with. This involves a degree of opposite information asymmetry, as well as problems of so-called foresight both in respect of what is to be the outcome of the transaction and the division of the residual knowledge (this will be dealt with below). Not all kinds of intellectual production exists in 'open systems', some exist in situations resembling closed systems, and hence the effect, quality etc of a given product can easier be judged. It is for example rather easy to estimate how much the costs are reduced with introduction of new layout software-program, though not with out problems.

The differences between physical production and intellectual production is acknowledged by Williamson (1985) when he - quoting Arrow - says that information exchange are more open to moral hazard than conventional goods since the user/buyer doesn't know the value of the good before he has purchased it.

The mentioned characteristics are not unknown within physical production where in certain cases only certain producers can produce products with the right technological interface, hence hold some kind of monopoly position, certain products are not produced before ordered etc, but in comparison to intellectual production, one might say, that they belong to different end of the scale in relation to these problems - though the impact of a new physical product usually is testable under conditions resembling the closed system.

Polemically, one can say that the market is essentially a black box in TCE (Mahnke 1999), and it is even darker when it comes to markets dealing with knowledge products.

When to 'make or buy'

In TCE the boundary-question is essential a question of whether the firm should make a given good or service, that is suspend the market (or co-operate with other parties about making it) or buying it on the market. The explanations centres around markets since in Williamsons words 'in the beginning there were markets'. This is not to understood as a historical statement but a logical statement, relating to what and how they address this question, namely through explaining the firm as suspension of markets; firms will only exist when markets are not the most efficient mode of co-ordination (Coase) or governance structure (Williamson).

Following Coase in-house production is to be preferred when the costs of internal production is lower than the transactions costs. However, he is only superficially addressing why internal or in-house production should be cheaper. One main reason is that the search and writing of contracts becomes cheaper due to the reduced number of contracts that has to be written. Why this automatically should result in reducing the costs of monitoring and enforcing them is unclear but must be assumed a part of Coases argumentation. The consequences of Coases line of reasoning is then that in-house production is to be more commonly found in markets with high transaction costs than with low.

Central in Coases argumentation for in-house production is also that it benefits from entrepreneurial direction. Coase himself describes it this way:

"If a workman moves from department Y to department X, he does not go because of changes in the relative prices, but because he is ordered to do so" (Coase 1937 (1991)., p. 19).

The notion of direction was challenged in a misleading way by Demsetz and Alchian (1972) where they claimed that in general there was no difference between directing the local shop-owner and internal labour. It was misleading since the shop-owner has several customers and the exiting of one transactions with him has fewer consequences for him, than when an worker gets fired since he – usually – depends strongly on the income from this particular job. Alchian has later acknowledged the problems in their 1972-text.

When it comes to intellectual production, however, the concept of direction starts to become rather blurred. In formal terms, the manager still holds the right to command the labour as long as the direction stays within what's stated in the contract and this stays with the labour market law. But he's real possibilities for direction are limited due to the information asymmetry between managers and knowledge-workers where the knowledge-worker usually knows more about 'his' specialised activity than the managers. Say, how do one direct the creativity behind a designing a new web-newspaper that has to be different from existing ones: is the web-designer creative in his offices only or when he walks in the park, and how do

the managers know how to define his goals for him when the web-designer is the expert – apart from in very general terms. The problem also exists in physical production as Hayek (1942) made clear when explaining how division of knowledge was the flip side of the coin of division of labour. But one doesn't need the same kind of knowledge to direct a manual employed as an intellectual employed (differences in relation to shirking and free riding will be dealt with below).

Co-ordination of economic activities under uncertainty by the entrepreneur is actually the situation under which Coase focuses on direction. The degree of uncertainty, and the problems of co-ordination are central themes in intellectual production but Coase doesn't contribute which much of concrete relevancy since he holds he's reflections in very general terms. One central problem is that he's framework for co-ordination addresses sequential co-ordination, not co-ordination that is not linear as is the case for most kinds of intellectual production where several interdependent processes are being 'produced' at the same time, though intellectual production also has strong linear element. In newspaper production, this might be from submitting an article over lay outing it to printing it.

One illuminating aspect of Coases reflections that should be stressed is that he shows that uncertainty creates large co-ordination difficulties. Here he is much in line with Richardson (1972) who is the founder of the so-called capability view of the firm. They agree upon that getting the incentives right is not enough to avoid co-ordination problems under characterised by uncertainty.

Williamson and 'make or buy'

Williamson also subscribes to direction as one of the reason d'être for the firm but he sees it only as a partial explanation, he adds opportunism and bounded rationality which leads him away from the complete contract perspective that underlines much OE (organisational economics), since complete contracts are usually not relevant for intellectual production due to the differences in information asymmetries and foresight problems –it is questionable to which extent they are useful for physical production.

As said Williamson approaches the problem of 'make or buy' from another viewpoint, he sees governance structure (market, hybrid or hierarchy) as way of reducing the 'costs' of opportunism in a context of bounded rationality.

Opportunism with guile. Broadly speaking opportunism with guile means that humans by nature are assumed to lie, steal cheat and other more sophisticated forms of deceit. Williamson is not assuming that all transactions are driven by people who cheat etc. but instead that they are disposed to do so and that there is no possibility to know ex ante of whether the will behave in an opportunistic way, hence firms have to take opportunism as a point of

departure when writing contracts for the transaction or deciding to make the product them self. Here it is important to stress that Williamson is not just concerned with spot market transactions (exchange).

Opportunism can be divided into holdup (sometimes referred to as precontractual) and moral hazard (sometimes referred to as postcontractual). The transition from preconctracs to postcontracs is the so-called fundamental transformation because:

"Options available in the former stage are lost in the latter, the value of some resources becomes dependent on particular unique other parties because of loss of significant substitutability be equivalent resources" (Alchian 1988, p. 67).

The contracts terms making this possible includes both promises of performance (quantity, quality and 'time-schedule / duration of contract) and agreement on price (Williamson 1985, Alchian 1988).

Potentials for holdup (or reneging) occurs when there is uncertainty due to information asymmetry that allows one part to take advantage of this asymmetry or when there is uncertainty about price and / or compensation because the values of the resources / investment depends on the continued association between the resources, and so on. This can be a question of difficulties of foresight as well as the degree of information asymmetry. The potential for holdup essentially depends on whether the resource is redeployable; if it is redeployable almost immediately and with almost no costs, there are only limited possibilities for holdup while a low degree of deployability allows for opportunistic behaviour.

Moral hazard in relation to the postcontractual phase or stage, on the other hand, arises as a potential problem when one or more persons relies on behaviour of another and information about this is difficult to obtain, this is essentially a question of the problems or costs of the Principals monitoring of the agent, either to get all the information the agent gains in the process (hidden information) or reduction of shirking, free riding (hidden action). Moral hazard is, one can say, a measure of the cost of having to monitor the agent plus what is not found in the monitoring (Alchian 1988). The potential for moral hazard depends on the specificity or plasticity of the resources, meaning that resources with a high degree of plasticity – being easily useful for other activities – can and / or will be used for other activities. For example a journalist working for a newspaper who is working on a book in the working hours while submitting the article he is supposed to though with less effort put into them.

Holdup and moral hazard problems would only have a limited significance for the boundaries of the firm if it weren't connected to bounded rationality, since all contracts would be able to deal with the problems of holdup and moral hazard, hence I turn to bounded rationality.

Bounded rationality. Bounded rationality, a concept he borrows from Simon , is very rather unclear. Sometimes it is used as simply meaning less than full information, which could be interpreted simply as an empirical modification to full information. This is what Foss (2000) labels thin bounded rationality. Full information then signals rationality. Thick bounded rationality – again a concept developed by Foss (2000) – then is on the ontological level where it can be seen as referring to cognitive capacity, structure, scripts, information channels etc. The concept of rationality itself also seems rather vaguely developed. Williamson just comes with references to maximising behaviour. Mostly bounded rationality serves the purpose of opening the door to incomplete contracts, apart from this it holds very few stated consequences for the theory despite that it should have consequences for conceptualising firms decision making, learning, believe in the idea of best practise, etc.

Bounded rationality is linked to opportunism in explaining governance structures, see below, but opportunism could also to be linked the other way around, meaning that bounded rationality could have as a consequence that parties involved in a transaction don't know that it is in their best interest to behave opportunistically. Williamson, however, has a strong believe in the opportunistic gene so this reflection is not warranted.

Boundaries according to Williamson

This leads to the question: what kinds of governance structure a firm should choose in a context of opportunism and bounded rationality?

To deal with opportunistic behaviour in a context of bounded rationality a firm can chose governance structures ranging from using the price mechanism over so-called hybrid forms to integrated or unified governance structure. The most efficient one is the one that reduces opportunistic behaviour with the least problematic trade off between costs and incentives.

The boundaries of the firms depends on the bounded rationality in relation to the degree of asset specificity, frequency and uncertainty of the transaction and possibilities for scale economies – keeping everything else equal. By asset specificity – here I only focus on asset specificity[1] - he means:

"A specialized investment that cannot be redeployed to alternative uses or by alternative users except at a loss of productive value. Asset specificity can take several forms, of which human, physical, site and dedicated assets are the most common" (Williamson 1996, p. 377).

That simply means that the investment only at (high) costs can – if at all - be used for other purposes, hence in a supplier-buyer relation it potentially opens for holdup and moral hazard. The degree depends on the degree of bounded rationality, and possibilities for dealing ex ante with the question in

[1] Among the three factors he claim that empirical studies backs him up in saying that asset specificity is the most important (this empirical documentation has been disputed several times).

terms of safeguards. This means that in cases where investments involves a high degree of asset specificity the transaction should be internalised, unless the possibilities for opportunism can be reduced ex ante in contractual agreements in forms of safeguards etc. Else their exists risk of opportunism where the strongest part takes advantage of its position. Two assumptions ground these reflections. First, 'in the beginning there were markets' (Williamson 1985, p 87) as he says, which means that markets-efficiency is taken as the point of departure, and that it is deviations from this that have to explained. Willamson is quite open about the way he echoes Hayek when he talks about the market. Second, we talk about one-shot contracts, in the sense that future options for renewal, new contracts etc, are not taken into account.[2]

It would be unfair to criticise Williamson for dealing with one-shot contracts only but it might be fair to open the field a bit by asking, how it would look if the future held possibilities for new contracts between the parties.

This question has been dealt with in the repeated game perspective (by among other Baker, Gibbon and Murphy (2000)). They conclude that repeated games affects the involved parties temptations to renege, create holdup. Williamson has not yet made an explicit reference to this perspective. Consistent with Williamsons understanding of opportunistic behaviour one cannot ex ante know if other more attractive offers some up for one of the parties, then firms always have to treat possibilities for holdup; this means following the line of argument behind one-shot contracts. The exception to this, as Williamson himself points to, is within fields where reputation plays an important role. But reputation is a difficult concept to deal with, and open questions are how much or for how long time a firm can behave opportunistic before it looses its reputation. Reputation might be rather relevant for intellectual production where due to the character of the knowledge product, the problems related to information asymmetries and foresight, firms will often choose a 'supplier' that holds a good reputation for not cheating, abusing the differences in information asymmetries.

One more thing has to be mentioned – as Williamson himself does – this framework does not address the question of situations where firms have to be able to respond fast to changes. This is to a large extent also more of an explicit co-ordination problem, I assume, hence of more relevance to the Coasian framework, but one might ask whether asset specificity couldn't be opened to cover aspects of time responsiveness. Of course, this is needed if one is going to use the framework in studying the newspaper industry where changes and fast responses are close to the rule rather than close to the exception.

Disadvantages of in-housing

[2] Central to explaining boundaries of the firm for intellectual production is human assets but looking apart from behavioural opportunism Williamson has very little to say (Mahnke 1999).

So far we have focused on the advantages of in-house production and the critic that might be related to Williamsons line of arguing but in-house production also has negative consequences according to Williamson. The disadvantage of in-house production is that it reduces the incentive intensity by moving away from high powered incentives to low powered ones, and hence opens for shirking, and so on, but it also allows for direction (instead of following the signals from the price mechanism). Therefore establishing an internal incentive and monitoring structure that reduces shirking becomes essential. It is in that spirit Demsetz and Alchian (1972) and Zenger and Hesterly (1997), among others, have argued that outsourcing is the most efficient solution to align incentives, allocate residual income and reduces the possibilities for shirking / free riding since it results in smaller units, hence it is easier to monitor and / or estimate each individuals contribution to the final products. Whether smaller size automatically lead to less free riding / shirking in intellectual production is not unlikely due to the fact that if just one is shirking it is readable in the result but these advantages might be on behalf of not getting access to economies of scale. The implications of economies of scale and firm size for learning and intellectual production remain an open question.

Williamson himself has rather little focus on internal organisation and only provides a rather limited argumentation as to why internal organisation should be more efficient (on the contrary he deals with bureaucratic failures) since from the perspective of opportunism the internal organisation would be driven by opportunism too; that would be for all the workers as well as for the managers in relation to shareholders etc. And with no hard core monitoring possibilities for intellectual production it is likely to – following Williamson's own assumptions – either to be very costly – and still rather imperfect - to monitor the workers, hence the organisation would be to close to being in an internal war-like situation. Polemically one might say that the obsession with opportunism creates a snowball-effect of monitoring (and increased contract costs) in a way that not only will shareowners monitor the manager who monitors the labour and so on. But shareholders will also have to start monitor each other to see if one of the others try to persuade the manager to make arrangements that are in somebody's specific interest. This could be an agreement with a firm that one of shareholders holds shares in. All these monitoring-activities have a cost too, that has to be included in shareholders 'calculations' of the profitability of their investment. All this monitoring will increase the de facto 'size' of the firm hence increases the monitoring cost even more. And one might not be surprised if a less 'suspicious' firm runs with the price advantages, and they might even gain from not treating everybody with suspicion since this might positively influence the so-called atmosphere (see below) – to use a concept from Williamson.

Williamson does point to the 'fact' that it is easier to audit internally and this might very well be true to some extent but as in relation to the direction-question, this is quite difficult for intellectual production.

He also provides the reader with a few ad hoc explanations for why in-house production should reduce opportunism. For example, he refers to 'a sense of being in it together' for employed and employee which of course – empirically speaking – is a truth with modifications. Furthermore, by doing this he indirectly shifts assumptions away from one-shot contracts assuming that workers are there for longer time than the duration of one long term contract. Empirically this might of course often be the case but this is the same with repeated contract between firms, and not a logically necessity.

Moreover, there need not be the difference between workers and supplier that TCE claims since suppliers also often depends on the well being of the company – of course there might be differences in 'scale' in relation to this question but no logical difference. Nowadays one supplier might depend strongly on one buyer while an IT-consultant gets ten attractive job-offers a day in his email inbox.

In the same vein, Williamson also refers to something called atmosphere more or less a synonymous with corporate culture but he does not explain where it comes from. He explicitly states (1996) that these kinds of institutions are not a part of his micro-focus but a part of the institutional environment.

One more thing that should be raised is that even if the firm gets the incentives right, and the contracts deals in the most appropriate way with foresight-problem this doesn't solve all the problems related to the production process, one might say that one thing is in the place but the most important is still missing, namely to organise the actual production process and the co-operation between user and producers.

Hybrids and markets – briefly

What about hybrids forms then? Hybrids exists when it is possible to make a trade off between differences in possibilities for taking advantage of a 'power' position, say if the asymmetry can be dealt with by a hostage-investment. That is an investment that reduces the asymmetries in 'power' positions.

This will open for closer co-operation between firms in form of for example joint ventures, being a partly suspension of the market, though still be rather exposed to a high degree of high-powered incentives. Williamson treats the hybrids as a stepson, and considers them to be of temporary character, meaning that they will only exist for a while before turning either into pure market or into pure hierarchy. Taken the earlier mentioned critique of both markets and hierarchies into account it is doubtful how far this explanation bring us, though of course it is a form that exists in the real world and has to be given a lot of attention.

Market-based transaction are chosen when there are few negative side-effects or risk of opportunistic behaviour since it allows for full exposure to high powered incentives, and where they provide sufficient statistics, as Williamson polemically writes with reference to Hayek – which they don't very often for intellectual production as became clear when discussing Coase and the market/price mechanism. This discussion there is no need to repeat but one might add that market transactions are transactions between producers, that is just mediated by the market, and the supplier might suffer from the same low incentives problems as the buyer would have.

Moreover, it is difficult to see why exposition to market gives high powered incentives, there is nothing in Williamsons assumptions about opportunism that 'dictates' this; and opportunism as self-realisation can be as high powered incentive as say a few more bucks in the pocket. Clearly, this indicates a need for incorporating different motivational structures.

However, when it comes to absolute survival of the firm it seems like a legitimate claim that markets are high powered with incentives but this is not the same as saying that this is most effective since in some branches, one can easily imagine that the firms that survive are those that can 'afford' a longer time horizon; say can afford to invest more in R and D.

Learning and knowledge creation

Finally, I will briefly touch upon the role of learning and knowledge creation in relation to boundaries of the firm. This will only be briefly since learning and knowledge creation doesn't play a central role in TCE, though Williamson acknowledges the importance several times. He mentions learning in relation to the tacit dimension of knowledge in the discussion of asset specificity; the '.. human asset specificity that arises in a direct learning-by-doing fashion …' (Williamson 1985, p. 95) and firms learning of handling transactions. Though nobody will be assaulted when I claim that learning doesn't hold a central position in TCE. The intriguing questions are then: Are there a elements in the nature of TCE that prevents it from focusing on or integrating learning? Is there something in TCE that put pressure on the approach to integrate learning more, and finally can it contribute to the learning/competitiveness debate?

Let me briefly dwell with these questions in a rather telegraphic style. In a dogmatic reading, short term economising can be seen as the pillar on which TCE rests but this isn't necessarily so since learning to manage the transactions in a more efficient way are a part of efficient managing of transactions, and this might cost extra in the short run. On the other hand, production is not central to TCE, hence it has only limited to offer in this respect, and it is not enough to improve ones learning and knowledge creation about transactions since perfect management of transactions doesn't ensure that there is a relevant product to sell - to state it in a rather popular way.

This doesn't mean, however, that there isn't a 'drive' in TCE to focus on or integrate learning more. This is among other things due to the central position that bounded rationality takes in the theoretical framework, and bounded rationality clearly indicates that there can be more to learn, better ways to do things etc, hence there is a pressure towards learning. Foss and Foss (2000) have given this a first try where the show that '… firms may have learning advantages because the use of incomplete contracts and the discretion it provided by authority are low-cost ways of conducting experiments …" (Foss and Foss 2000). Exactly the low-cost arguments points to why the resource-based approach could eventually benefit from looking more into TCE. RBV hasn't, generally speaking, developed very sophisticated stories about the links from knowledge creation and learning to firms competitiveness. TCE has a developed framework for dealing with this. In Foss and Foss's words:

"From an OE perspective, these issues should be addressed using a micro-analytical focus; an economizing approach which directs attention to efficiency issues, and comparative instutionalism which reequip arguments of why one institution may more efficiently organize learning processes".

An open question is, however, still in which areas it is that TCE has most to contribute. For example in areas with Schumpetarian rents the low-cost line of reasoning might not be the most convincing.

3. Concluding remarks

Central in TCE's explanation of the boundaries of the firm is the functioning of the market, since 'in the beginning there were markets'. Markets are, however, still to a large extent a black box in TCE which Williamson (1985) indirectly acknowledges when he says that TCEs understanding of the market are open to much of the criticism that Nelson and Winter (1982) puts forward about orthodox economics. The working of the market or price mechanism becomes even more problematic when dealing with intellectual production as Williamson also acknowledges but he doesn't take any consequences of it for his understanding of the markets, hence the boundaries of the firm. This means that the conclusions about when to 'make or buy' are more difficult to deal with than what TCE proposes. Williamson himself points to the limitations of the framework for dealing with firms that depend on fast responsiveness. This shows that there is a need for at more elaborated understanding of the market in TCE, especially about how it works in intellectual production, and what the consequences are for the boundary-question.

This doesn't mean, however, that the discussions of assets specificity is not relevant to understand 'make or buy' decisions but this has to be

understood in relation to bounded rationality and opportunism as they manifest them self in intellectual production.

Williamsons focus on one shot long term contracts is maybe not the most relevant transaction to deal with for intellectual production since it to a large extent is a 'business' that depends on reputation, among other things due to the character of the product. This doesn't indicate that firms shouldn't be aware of the 'danger-signals' related to transactions as a supplementary to their focus on reputation. How one can address reputation within a TCE framework is unclear.

Why in-house production should reduce the consequences of opportunism is the best case rather unclear, and Williamsons explanations not fully consistent, and with the difficulties of monitoring intellectual production is becomes even more unclear why in-house production should reduce the cost of opportunism.

Co-ordination is another central aspects of intellectual production and TCE represented by Coase is mostly focused on sequential co-ordination not production with simultaneous engineering, and so on, that characterises intellectual production. When Williamson defines a transaction in relation to crossing a technological interface, he shows how he actually thinks in terms of the running assembly line, that is the incarnation of modern physical production.

Finally, there is no way around learning and knowledge creation, when dealing with intellectual production, and this issue holds only a marginal position in TCE. There are potentials for changing this and due to the centrality of the idea of bounded rationality, there is even an endogenous 'pressure' for integrating learning more. Not all the problems related explicitly to production are dealt with, and these come before having the 'perfect' transactions. On the other hand, the strength of TCE is the dealing with efficient organisation and from this RBV can certainly learn a lot.

References:

Alchian, A and S. Woodward (1988): The firm is dead, long live the firm. Journal of Economic Literature, 26, 1, 65-79.

Baker, G; R, Gibbon and K. Murphy (2000): Relational contracts and the theory of the firm. Unpublished.

Coase, R (1937): The nature of the firm, in: The nature of the firm (1991), ed. By Williamson and Winther. Oxford University Press.

Demsetz, H (1994a): Ownership, control and the firm. Vol 1.Blackwell.

Demsetz, H (1994b): Ownership, control and the firm. Vol 2.Blackwell.

Demsetz, H and Alchian, A (1972): Production, information cost, and economic organization. American Economic Review.

Ghoshal, S and Insead, P.M, (1996): Bad for practise: A critique of the transaction cost theory. Academy of Management, The Academy of

Management Review, downloaded, no paging.
Grandori, A (2001): Co-ordination, fortcomming.
Hayek, F. A (1945): The use of knowledge in society, in idem. 1948. Individualism and Economic Order. Chicago: University of Chicago Press.
Foss, N (1993): Theories of the firm: competence and contractual perspectives. Journal of Evolutionary Economics.
Foss, N (2000): Thin and thick bounded rationality … unpublished.
Foss, N and Foss, K (2000): Learning in firms: Knowledge-based and property rights perspectives. Unpublished.
Mahnke, V (1999): Intellectual production and the strategic theory of the firm. Handelshøjskolen i København.
Nelson, R and S. Winter (1982): An evolutionary theory of economic change, The Belknapp Press.
Richardson, G (1972): the organisation of industry. Economic Journal 82, 883-96.
Sayer, A (1984): Method in social science. Hutchinson.
Williamson, O (1996): The mechanism of governance. Oxford University Press.
Williamson, O (1985): The economic institutions of capitalism. Free press.
Williamson, O (1979): Transaction-Cost economics: The governance of contractual relations. Journal of Law and Economics, 22, 233-261.
Williamson, O (1975): Markets and hierarchies: analysis of antitrust implications. Free Press.
Zenger, .R and Herterly, W.s 81997): The disaggregation of corporation: selective intervention, high-powered incentives, and molecular units. Organizational Science, 8, 3, 209-222.