HUMAN CAPITAL AND TRANSACTION COST ECONOMICS

Nicolai J. Foss

SMG WP 2/2008

January 11, 2008
HUMAN CAPITAL AND TRANSACTION COST ECONOMICS

Nicolai J. Foss
Center for Strategic Management and Globalization
Copenhagen Business School
Porcelænshaven 24, 1st fl.; 2000 Frederiksberg; Denmark
njf.smg@cbs.dk

and

Department of Strategy and Management
Norwegian School of Economics and Business Administration
Breiviksveien 40; N-5045; Bergen; Norway

January 11, 2008; Word count (main body): 7,344

Prepared for Alan Burton-Jones and J.C. Spender, eds.

Acknowledgments: I am grateful to Alan Burton-Jones and JC Spender for comments
on an earlier version of this chapter.

JEL coding: D23, J24, M5

Key words: Human capital, transaction cost economics, property rights theory.
HUMAN CAPITAL AND TRANSACTION COST ECONOMICS

Abstract
Human capital has often been discussed in transaction cost economics, particularly in connection with understanding the employment relation. This chapter reviews Williamson’s thinking on the issue, explains how it differs from Coase and Simon’s, and briefly discusses the relevant empirical literature. The chapter also covers property rights theory, and discusses various critiques of the treatment of human capital in transaction cost economics.
INTRODUCTION

Human capital may be defined as the stock of valued skills, knowledge, insights, etc. controlled by an individual, in other words, the attributes of this individual that are valuable in an economic context. This stock may yield services such as labor services, management services or entrepreneurship. As slavery is almost universally prohibited, the issue is not one of allocation of ownership to human capital, but of whether these services are supplied through market transactions, transactions in an employment relation, in club-like organization (Buchanan, 1965) or within households (Becker, 1964, 1991). A key issue in the body of micro-economic thought commonly referred as “organizational economics” or the “theory of the firm” concerns the efficient organization of such transactions, specifically across contracts and governance structures. By supplying instruments of coordination and providing incentives, contracts and governance coordinate human capital services, but also provide incentives for investments in human capital. Under a norm of efficiency (value maximization) which specific contracts or governance structures that human capital is matched with depends on the characteristics of the latter.

While this overall perspective cuts across transaction costs economics, property rights/incomplete contracts theory, and agency theory, the specific terminology (“transactions”, “governance structures” etc.) applied here is particular to transaction cost economics (henceforth, “TCE”). Arguably, this is warranted as TCE has provided the first and still most comprehensive treatment of the organizational ramifications of human capital in economics. The present chapter surveys TCE as it pertains to human capital, concentrating on the work of TCE’s most prominent flagbearer, Oliver Williamson, the precursors of his work (i.e., Coase [1937] and Simon [1951]) as well as the various operationalizations and extensions of his approach. Also briefly discussed is the property rights approach (Hart, 1995), which like TCE rests on the notion of incomplete contracting and the need to safeguard transactions under such conditions.

While human capital plays a key role in these approaches as an important component in the explanation of important aspects of economic organization, notably the boundaries of the firm, it is important to stress that neither TCE nor PPR hold very specific implications
concerning the management of human capital. As Loasby (1995) observes neither carry any particular implications regarding what to “tell Mrs. Jones what to do on Monday morning.” TCE or PPR are not HRM or general management approaches, and they were never designed as such. They are rather somewhat static approaches concerned with the efficient matching of transactions and governance structures/allocations of property rights at a point of time, they are not taken up with process analysis, managerial decision premises, and so on, and like most of economics, they make heavily simplified assumptions concerning the motivation and cognition of agents. Nevertheless, TCE in particular has been subject to heavy criticism from management scholars who have indicted the theory for it purportedly being “bad for practice.” Other critiques on the part of management scholars concern the neglect of differential firm-level capabilities, and the basis of such capabilities in human capital. Such critiques are briefly discussed at the end of this chapter.

THE ROLE OF HUMAN CAPITAL IN TRANSACTION COST ECONOMICS

Transaction Cost Economics: Overall

The foundations of contemporary transaction cost economics were laid by Ronald Coase (1937, 1960). The logical starting for a transaction cost approach to governance and organizational issues is Coase’s (1960) insight that if it weren’t for transaction costs, all gains to trade would be exhausted and this could take place under any organizational arrangement. This connects to Coase’s earlier paper (Coase, 1937), for the argument in that paper is that the assessment of the net benefits of organizational and governance alternatives must proceed in terms of a comparative analysis of the costs of transacting under the relevant alternatives (Barzel and Kochin, 1992).

In a string of influential contributions, Williamson (e.g., 1971, 1973, 1975, 1985, 1996) has built a theory that while built on Coasian foundations also incorporates ideas from psychology and contract law. The behavioral starting points in Williamson’s theorizing are bounded rationality and opportunism. Simon’s notion of bounded rationality implies the

---

1 But for attempts to take TCE and PPR in a management direction, see Rubin (1987) and Milgrom and Roberts (1992).
presence of contractual incompleteness and, consequently, a need for adaptive, sequential decision-making. Opportunism is defined as “self-interest seeking with guile,” and its implication is that contracts will often need various types of safeguards, such as “hostages” (e.g., the posting of a bond with the other party). The unit of analysis in Williamson’s work is the multi-dimensional transaction. In addition to uncertainty (which is “frozen”), the dimensions of transactions that are primarily determinative of the costs of those transactions are frequency and asset specificity. In an early contribution, Klein, Crawford and Alchian (1978: 297) linked asset specificity to the concept of approducible quasi-rent:

Assume an asset is owned by one individual and rented to another individual. The quasi-rent value of the asset is the excess of its value over its salvage value, that is, its value in its next best use to another renter. The potentially approduciable specialized portion of the quasi rent is that portion, if any, in excess of its value to the second highest-valuing user.

Following Klein, Crawford and Alchian, asset specificity has increasingly become the central character in Williamson’s analysis. Williamson (1996) now identifies six different reasons why assets may be costly to redeploy, namely because of attachment to a brand name, a need to take quick actions (“temporal specificity”), market size (“dedicated assets”), localization, physical characteristics, and specialized knowledge (that is, human capital specificity).

Specific assets open the door to opportunism. If contracts are incomplete due to bounded rationality, they must be renegotiated as uncertainty unfolds, and if a party to the contract (say, a supplier firm) has incurred sunk costs in developing specific assets (including human capital), that other party can opportunistically appropriate an undue part of the investment’s pay-off (“quasi-rents”) by threatening to withdraw from the relationship. This situation leads to a Pareto-inferior outcome, for example, a no-trade outcome. Efficiency dictates the internalization within a firm of transactions that involve highly specific assets. More generally, Williamson (1985: 68) argues that variety in contracts and governance structures “… is mainly explained by underlying differences in the attributes of transactions.” The general design principle of discriminating alignment dictates aligning transactions that differ in the dimensions of
uncertainty, frequency and asset specificity with governance structures which differ in the capacities to handle different transactions (cf. the earlier discussion of governance structures and governance mechanisms) in transaction cost discriminating way. Thus, specific constellations of (values for) the uncertainty, frequency and asset specificity variables map directly into specific governance structures. This is the main predictive content of Williamsonian TCE.

So far, there is nothing particular about human capital; it is just a capital asset like any other which may be more or less specialized to specific uses and/or users. It may explain variation in the organization of transactions across markets, hybrids and hierarchies, exactly as, for example, physical assets may. However, it is important to bear in mind that Williamson not only emphasizes that transactions differ in the three mentioned dimensions; he is also adamant that alternative governance structures have differential capacities for governing different transactions. And it is here that (specific) human capital partakes of particular importance in Williamson’s thought, for it relates directly to the employment relation and the understanding of the costs and benefits of hierarchical governance. To see this it may be instructive to look at the precursors of Williamson’s analysis of the employment relation, namely Coase (1937) and Simon (1951).

**Coase and Simon on the Employment Relation**

In his classic paper, Coase (1937) essentially defines the firm as the employment relation, In the presence of uncertainty, Coase argues, contingencies are costly to anticipate and describe in advance, and rather than negotiating on a spot market basis over each contingency as they arise, an employment contract is concluded. The latter is defined as “… one whereby the factor, for a certain remuneration (which may be fixed or fluctuating) agrees to obey the directions of an entrepreneur within certain limits. The essence of the contract is that it should only state the limits to the powers of the entrepreneur. Within these limits, he can therefore direct the other factors of production” (idem. 242).

Simon (1951) defines authority as obtaining when a “boss” is permitted by a “worker” to select actions, $A^0$ ( $A$, where $A$ is the set of the worker’s possible behaviors. More or less authority is then defined as making the set $A^0$ larger or smaller. The boss then picks the action that he prefers and directs the worker to that action which ( for the latter to accept the
assignment (must lie within his or her “zone of acceptance.” A worker’s zone of acceptance is defined in Simon as that set of actions where the worker’s expected costs of carrying out these actions do not exceed the agreed-upon on wage. An important feature of authority is that the authority of a superior is constrained by the acceptance of the subordinate of the authority. “A subordinate may be said to accept authority,” Simon (1951: 22) explains, “… whenever he permits his behavior to be guided by a decision reached by another, irrespective of his own judgment as to the merits of that decision.” In contrast, in a market contract, the parties negotiate ex ante about the actions that the agent can take in response to various contingencies so as to fulfill the contract. Thus, the principal’s flexibility under market contracting is limited compared to what it would be under authority.

Clearly, Simon’s explanation of authority and the employment relation is quite akin to Coase’s (1937). To both of them authority is a decision right that an employer acquires, because he expects to obtain only ex post contracting the relevant information that will make it possible for him to pick his preferred actions within a specified subset of actions, which he will then direct the employee to carry out. Knowledge and knowledge asymmetries is not an essential part of either story (as they are in, e.g., Demsetz, 1991). Thus, in Simon symmetric knowledge is consistent with the authority relation; it is sufficient for the existence of that relation that one contracting party for whatever reason stands to gain more than the other from picking the actions once contingencies materialize. The thrust of Coase’s discussion, however, indicates that the employer is clearly the informed party who picks well-defined actions from a set of discrete actions on the basis of knowledge about contingencies that is superior to that of the employee, but this is not a central point and at any rate not the fundamental reason why the employment contract, and hence the firm, emerges. The absence of considerations of knowledge is also indicated by the fact that in both treatments the employer grants no discretion with respect to the choice of actions to the employee; such delegation would be warranted (ceteris paribus) if the employee possesses superior knowledge about contingencies, actions or the matching of contingencies and actions. In other words, characteristics of human capital are not included in Coase and Simon’s theories of the employment relation, where employees are implicitly taken to be homogenous.

Williamson on Human Capital and the Employment Relation
While Williamson accepts Coase’s emphasis on authority as a defining characteristic of the firm, he goes very significantly beyond the Coase-Simon analysis. He does so by taking explicit account of human capital, by examining the transaction cost ramifications of human capital specificity, and by emphasizing the distinct informational advantages of the hierarchy in dealing with specific human capital. While Coase had stressed the savings of transaction costs attendant upon changing a spot labor transaction to an employment contract, Williamson locates the relevant transaction costs in the opportunism problem that may arise under small-numbers bargaining.

Williamson’s first reflections on these issues from a distinct TCE perspective appear in his 1975 paper with Wachter and Harris (Williamson, Wachter and Harris, 1975; reprinted with modifications as chapter 4 in Williamson, 1975). Williamson begins by declaring that he

… is concerned with the implications of an extreme form of nonhomogeneity – namely, job idiosyncracy – for understanding the employment relation …The purpose is to better assess the employment relation in circumstances where workers acquire, during the course of their employment, significant job-specific skills and related task-specific knowledge” (1975: 57; emph. in original).

In a footnote, he observes that “the employment relation is not an isolated case of idiosyncratic exchange conditions. The vertical integration problem turns in no small degree on these same considerations” (p.57n). In other words, in 1975 Williamson’s reasoning on the employment contract proceeds by extension of the theoretical mechanism that he had developed in his 1971 and 1973 papers on vertical integration (Williamson, 1971, 1973).

Explicitly drawing on Becker’s work on human capital, as well as on labor law and the internal labor market literature (Doeringer and Piore, 1971), Williamson proceeds by identifying four different modes for contracting labor services in a comparative manner, namely by 1) sequential spot contracts (contract now for specific performance); 2) contingent claims contracts (i.e., contracting now for the specific performance of \( x_i \) in the future depending contingent on \( e_i \) obtaining); 3) long-term contracting (wait until the future unfolds and contract for the appropriate specific action); and 4) the authority relation (as in Coase [1937] and Simon [1951]). This identification of the relevant contractual alternatives allows Williamson to take
issue with Alchian and Demsetz (1972) and with Simon (1951). The former famously argued that talk of “authority” is a “deceptive way of noting that the employer continually is involved in renegotiation of contracts on terms that must be acceptable to both parties” (Alchian and Demsetz, 1972), an argument, Williamson (1975: 67) observes, that seems to involve zero or negligible transition costs associated with employee turnover, and which therefore corresponds to contracting alternative 3) above (rather than to the authority relation). Simon (1951) is criticized for only confronting alternatives 1) and 4) which means that “the terms are rigged from the outset” (Williamson, 1975: 71): As contractual alternative 1) lacks adaptability, transactions requiring such adaptability will naturally lead to the choice of the employment contract in Simon’s model. However, there are numerous contractual alternatives that do allow for adaptability, so adaptability cannot be the (main) explanation for the choice of the authority relation. That choice rather depends on “task idiosyncracies.”

Drawing on Doering and Piore (1971), Williamson outlines various sources of such idiosyncracies, which are to be understood as different manifestations of human capital specificity. Thus, such specificity may arise through workers acquiring knowledge about non-standard equipment (Penrose, 1959), specialized processes (Nelson and Winter, 1982), the capabilities of other employees (Alchian and Demsetz, 1972), and firm-specific communication channels and values (Arrow, 1974). The problem with task idiosyncracy/specific human capital is that “workers are strategically situated to bargain opportunistically” (Williamson, 1975: 74). Such opportunistic bargain may cause alternative 4) to be the preferred mode of organizing the supply of labor services. The employment relation, Williamson (1975) argues, means making use of a collective bargaining system where “wages rates are attached mainly to jobs rather than to workers. Not only is individual wage bargaining thereby discouraged, but may even be legally foreclosed … The incentives to behave opportunistically … are correspondingly attenuated” (idem.).

A later discussion (Williamson, 1985) expands on labor organization and makes more out of opportunism and asset specificity, compared to the treatment in the 1975 book. This reflects the increasing “hardening” of Williamson’s research program around the problem of hold-up in the presence of specific assets as the apparently universal mechanism that explains choices
between markets, hybrids and hierarchies. Thus, Williamson explains that what is at issue in explaining the organization of work is not the deepening of human capital, but its specificity; while both have implications for productivity, only the latter have direct implications for organization.

However, to nuance his analysis, Williamson adds one more dimension, namely that of “separability” of “work relations.” The latter is an attempt to capture the issues of inputs and output measurability highlighted in agency theory (Williamson specifically refers to Alchian and Demsetz, 1972). Accordingly, the separability of work relations depends on how costly it is to measure individual outputs. Armed with the dimensions of specificity and separability, Williamson constructs a 2x2 matrix that maps the organization of human capital services (1985: 247). Non-specific, separable human assets, such as migrant farm workers, certain kinds of engineers, IT specialists, etc., can move between employers without loss of productivity; there is therefore no efficiency interest in maintaining a specific relation. The “primitive team” obtains when human assets are non-specific but their outputs are difficult to measure on an individual basis. As analyzed by Alchian and Demsetz (1972) this kind of work organization motivates monitoring of inputs. The “obligational market,” in turn obtains when human assets are specific but separable. Under these conditions both firms and employees have an interest in maintaining the relation, and various devices, such as severance pay and employee benefits such as nonvested retirement, arise to discourage arbitrary dismissal or unwanted quits. Under the final conditions, high specificity and non-separability, the “relational” team organization arises, characterized by strong implicit contracts that can sustain cooperation. Williamson refers to Ouchi’s (1980) notion of “clan organization” and to the Japanese corporation (anno 1985) to exemplify the relational team.

Forbearance and the Informational Advantages of Hierarchy

In his later work Williamson has placed much less emphasis on collective bargaining than he did in his 1975 book, and instead emphasized the informational advantages of the authority relation and the hierarchical organization that it is embedded in. However, the theme of the informational advantages of hierarchy has been present in Williamson’s work at least since his 1971 paper. As he argues there, the hierarchy possesses certain inherent advantages over market
contracting: “... when conflicts develop, the firm possesses a comparatively efficient resolution machinery. To illustrate, fiat is frequently a more efficient way to settle minor conflicts (say differences of interpretation) than is haggling or litigation (Williamson, 1971: 114).” This advantage has come to play an increasing role in Williamson’s work. Thus, he has placed increasing emphasis on the argument organizational governance is characterized by its own implicit contract law, what he calls “forbearance”. To illustrate, whereas divisions will not normally be granted standing for a court, corporate headquarters and headquarters function as the firm’s “ultimate court of appeal”. For example, Williamson (1991) points out that disputes which arise within the firm, for example, between different divisions, may be easier to resolve than disputes arising between firms which sometimes require the use of the court-system. Thus, managerial authority partakes of an important role as arbitrator in the face of conflicts and disputes over unforeseen contingencies; in other words, Williamson’s extension of the Coasian view of authority is to analyze it as a “private ordering,” a private legal institution (Williamson, 2002). Part of that argument is Williamson’s assertion that there are qualitative and quantitative differences between the information structures that are available under market contracting and those that are available in the firm, an argument put to work in Williamson’s work on the M-form as an internal capital market (Williamson, 1975).²

**Employee and Employer Opportunism**

According to Coase (1937), the key reason why firms exist lies in the exercise of managerial authority in response to changes in the environment being an efficient mode of organization for certain transactions. In this view, echoed in countless other contributions to organization theory and the theory of the firm, managerial intervention is implicitly always seen as beneficial.³ However, in

² These are claims that have disputed by what is often referred to as the property rights theorists of the firm. In the words of Grossman and Hart (1986: 691): “... the transaction cost-based argument for integration does not explain how the scope for such behaviour changes when one of the self-interested owners becomes an equally self-interested employee of the other owner.”

³ However, as Williamson (1996: 150) points out this raises the “puzzle” of the “impossibility of selective intervention,” namely, “[w]hy can’t a large firm do everything that a collection of small firms can and more?” Thus, a large firm could replicate the market and only selectively intervene when there would be expected net gains from this, so that “... the firm will do at least as well as, and will sometimes do better than, the market.” Consequently, firms can grow without limits. As this is absurd, the answer to the question of what are the efficient boundaries of the firm must involve an understanding of why selective intervention of this kind is impossible, that is, a theory of managerial and hierarchical failure.
actuality, managerial intervention will typically override existing instructions of employees. Moreover, in firms where employees are given considerable discretion, managerial intervention may amount to overruling decisions that employees have made on the basis of decision rights that have been delegated to them (Baker, Gibbons, and Murphy, 1999). This suggests that employee utility may be harmed by managerial intervention, particularly when employees perceive opportunistic intentions behind such intervention (Foss, 2003). Negative firm-level effects may result when employees put in less effort and/or undertake less investment in specific human capital.

Contrary to the common claim that Williamson puts all the explanatory emphasis on employee opportunism (e.g., Dow, 1987), he has in fact never been blind to the possibility of employer opportunism. Thus, Williamson (1996: 150-151) points out that the option to intervene “… can be exercised both for good cause (to support expected net gains) and for bad (to support the subgoals of the intervenor).” This raises a series of fundamental incentive problems that are rooted in the fact that it may be difficult to verify the nature of the cause and that promises to only intervene for good cause are hard to make credible as they are not enforceable in a court of law. A problem of “… credibly [promising] to respect autonomy save for those cases where expected net gains to intervention can be projected” (Williamson 1993: 104) remains. An important problem therefore concerns how managerial intervention for good cause (i.e., value-increasing intervention) may be promoted, while intervention for bad cause (i.e., value-destroying intervention) is avoided. The practical relevance of this issue is underscored by a series of empirical studies that suggest that management breaking the psychological contract with employees is a phenomenon of frequent occurrence (Rousseau and Parks, 1992; Coyle-Shapiro and Kessler, 2000). While Williamson himself has not dealt with the issue in any detail, subsequent work has pointed to various mechanisms, such as implicit contracts (Kreps 1990; Baker et al. 1999), explicit credible commitments, or organizational structure (Aghion and Tirole 1997; Foss, Foss, and Vazquez, 2006) that may reduce the incidence and severity of harmful managerial interventions.

**Empirical Application**

There is a now a huge body of empirical literature examining various issues in transaction cost economics. Many empirical studies have examined the make-or-buy decision, focusing
mainly on the transaction cost approach. The empirical TCE literature is surveyed and summarized in Shelanski and Klein (1995) and Macher and Richman (2006), and the bulk of the evidence is interpreted as consistent with the predictions of TCE (see David and Han [2003] for a contrary view). Despite challenges associated with the measurement and definition of key variables, the role of asset specificity, comparison with rival theories, and causality, the transaction cost model seems to have straightforward empirical implications, such that observed forms of organizational governance can be explained in terms of asset specificity, uncertainty, frequency, and so on.

Indeed, much of the success of Williamson’s formulation of TCE lies exactly in his identification of the key attributes that determine governance costs (Masten, 1996: 45). In Coase (1937), the decision rule/prediction is that GS* = GS1, if GC1 < GC2 and GS* = GS2, if GC1 ( GC2 (“GS” is governance structure, “GC” is governance cost). From the empirical point of view, Williamson’s contribution is to specify the cost relations ( GC1 = 1X+ e1 and GC2 = 2X+e2, X being the vector of observable attributes that impact governance costs (, yielding testable reduced form relations by means of qualitative choice models (e.g., probit models).

Early empirical work in TCE all took such an approach (Monteverde and Teece, 1982; Anderson and Schmittlein, 1984; Masten, 1984; John and Weitz, 1988). Monteverde and Teece draw on Teece’s (1980) earlier work on how firm-specific human capital may help shaping the horizontal boundaries of the firm (i.e., diversification) to examine how firm-specific “industrial know-how,” represented by the applications engineering effort (rated on a 1-10 scale) associated with the development of a given automobile component, shapes the incentives to engage in vertical integration. Anderson and Schmittlein (1984) examine whether the function of personal selling (as opposed to mass communication) is integrated or not, in the sense of whether a manufacturer’s representative is used or rather an employed salesperson. The human capital specificity of the sales person is measured through variables that measure, for example, the difficulty of the salesperson learning the ins and outs of a company, including learning about accounts, the importance of customer loyalty, etc. John and Weitz (1988) measured the

---

4 Comparatively little empirical work looks at agency theory (excepting corporate finance and governance applications), the resource-based view, and the property-rights approach.
specificity of human capital investments by the time required for newly hired salesperson with industry-experience to familiarize himself with a given firm’s products and customer-base.

Since these classic studies, numerous studies in industrial organization and regulation (e.g., Crocker and Masten, 1996; Saussier, 2000), alliance organization (Dyer, 1997), international business (and marketing (e.g., Anderson, 1985; Heide and John, 1988) have examined the impact of human capital specificity on governance and contractual choice. Most confirm the TCE prediction(s). However, in a provocative paper Masten, Sheehan and Snyder (1991) suggest that standard TCE may have misconstrued the reason for this consistent finding. The standard story (Williamson, 1975, 1985) has specific human capital being associated with employment contracts because the costs of organizing the sourcing of the relevant services through market contracting is are too high. However, it may be, Masten et al. argue, that it is simply less costly to manage workers with more specific skills; thus, firm organization is preferred for specific human capital, not because the costs of using the market for organizing such assets are too high, but because internal costs of organization are low. This clearly smacks of the knowledge-based approach to the firm (particularly Kogut and Zander, 1992) rather than of TCE.

However, as Poppo and Zenger (1999) point out the standard empirical research designs do not discriminate between TCE and competing approaches such as the measurement approach (Barzel, 1997) or knowledge-based approaches (e.g., Arrow, 1974; Demsetz, 1991; Kogut and Zander, 1992). All three approaches will, roughly, predict that specific human capital and firm organization are highly correlated. Poppo and Zenger develop and test a model of comparative institutional performance rather than institutional choice, and thereby examine the degree of support for the three competing explanations of boundary choice. For their sample of firms, they disconfirm the Masten et al. hypothesis. However, they also conclude that the evidence points to the need to eclectically combine the three theories for a satisfactory understanding of the boundaries of the firm, at least in the context of the sourcing of IT-services.

5 Comparatively little empirical work looks at agency theory (excepting corporate finance and governance applications), the resource-based view, and the property-rights approach.
PROPERTY RIGHTS APPROACHES AND HUMAN CAPITAL

Property Rights and Incomplete Contracts

The property rights approach was initiated in Grossman and Hart (1986) and explicitly motivated as an attempt to model Klein, Crawford and Alchian (1978) and Williamson’s ideas on asset specificity, using game theoretical modeling conventions (see Hart [1995] for an accessible introduction). However, the outcome of that attempt was essentially a new theory. The basic logic of the approach is as follows. As in Williamson’s work, a central assumption is that real-world contracts must necessarily be incomplete in the sense that the allocation of control rights cannot specified for all future states of the world (Grossman and Hart, 1986). Ownership is defined as the possession of residual rights of control, that is, rights to control the uses of assets under contingencies that are not specified in the contract. Residual rights over human capital cannot be transferred. Control implies the ability to exclude other agents from deciding on the use of certain assets, for example, the owner has the right to pull out the assets he owns from a relation. Such ownership rights can confer bargaining power, and play an important role in the determination of the efficient boundaries of the firm. Thus, control rights determine the boundaries of the firm: A firm is defined as a collection of jointly owned assets. Obviously, the relevant assets are alienable assets plus, possibly, the owner’s inalienable human capital.

Asset ownership is important because the willingness of an agent to undertake a non-contractible investment (notably an investment in human capital) which is specific to the non-alienable asset(s) in the relation depends on who owns the asset(s). If the agent who undertakes

6 Indeed, Williamson (1996: 372-73) argued that TCE had progressed from “pre-formal” (i.e., Coase’s work) to “semi-formal” (i.e., Williamson’s own work) and “fully formal” (i.e., the new property rights view) stages. However, in a later discussion Williamson (2000) finds the new property rights finds it substantially different from his own framework in several key respects.

7 Klein, Crawford and Alchian as well as Williamson have unforeseen contingencies at the heart of their stories: It is the haggling that arises when contracts are pushed outside of their self-enforcing range by unforeseen contingencies that is the main problem. What matters is the ex post action. In contrast, most of the incomplete contracting approach assumes that ex post bargaining is efficient and that actions (e.g., with respect to investment) are taken immediately after the contract is signed. Thus, what drives these models are misaligned ex ante incentives, particularly with respect to investment in vertical buyer-supplier relationships.

8 The basic distinction between an independent contractor and an employee, that is to say, between an inter-firm and an intra-firm transaction, turns on who owns the physical assets which the agent utilizes in his work. An independent contractor owns his tools etc., while an employee does not.
the investment does not own the asset, she may, as in Williamson’s work, be subject to a hold-up by the owner. On the other hand, the ability to deprive an agent of the piece of capital with which she works (and to which she may be heavily specialized) provides room for authority. Efficiency then dictates that the agent who is to make the most important (non-contractible) asset-specific investment should own the asset. It is not that opportunism can be avoided by internal organization/integration \textit{per se}. Integration may shift incentives for opportunistic behavior, but it does not remove such incentives. Given this, one should choose the ownership arrangement that \textit{via} its impact on incentives minimizes the consequences of opportunism. In sum, the property rights approach is a theory of the efficient allocation of ownership to alienable assets, given investments, (typically in human capital (that may be important to value creation, but are non-contractible.

\textbf{Extensions: Access, Implicit Contracts, and Organization Ownership}

Employees are not, strictly speaking, part of the firm because employees cannot be owned. However as Rajan and Zingales (1998: 388) point out, “... there is a sense in which employees ‘belong’ to an organization ... This sense of belonging arises from the expectation ‘good citizens’ of an organization have that they will receive a share of future organizational rents”. This sort of belonging can be explained within the confines of the property rights approach by invoking the concept of “access.” This means that agents are allowed to work with critical resources, specialize themselves relative to these resources, and make themselves valuable in this way. Since a specialized employee can control her own specialized human capital, she now has additional power, although she doesn’t possess more residual rights of control. As Rajan and Zingales show, access may sometimes provide better incentives for making efficient effort choices or investments in human capital than direct ownership.

Incentives may also be provided by even softer means. When it is difficult to write complete state-contingent contracts, for example, when certain variables are either ex ante unspecifiable or ex post unverifiable, people often rely on “unwritten codes of conduct”, that is, on implicit contracts. These may be self-enforcing, in the sense that each party lives up to the other party’s (reasonable) expectations from a fear of retaliation and breakdown of cooperation. In an early extension of the property rights approach, Kreps (1990) argues that employers and
employees may be seen as playing a prisoners dilemma-game, that with repeated plays a cooperative norm (interpreted as corporate culture), may be established, and that this established norm tells employees (as well as outside contractors) that firm management will not opportunistically take advantage of them. The firm is seen in Kreps’ paper, not as a collection of physical assets as in Grossman and Hart (1986), but rather as a carrier of reputation capital. Baker, Gibbons and Murphy (2002) also merge ideas on self-enforcing contracts with property rights, showing how the allocation of ownership influences which relational contracts are feasible. Klein (1988) takes a different approach. He argues that rather than focusing on individual human or physical assets in the analysis of the boundaries of the firm, attention should be directed to what “organization ownership”. Thus, Klein argues that vertical integration may imply a certain degree of ownership of human capital after all. For an organization can obtain ownership of another organization’s organization capital, that is, the firm-specific knowledge embodied in the organization’s team of employees (i.e., its capabilities). This can alleviate the hold-up problem, for the reason that it is (post integration) hard for the now integrated team to hold up the acquiring organization. The costs of collective action may be prohibitive and/or the hold-up attempt may be illegal according to labor law.

TRANSACTION COST ECONOMICS AND OTHER PERSPECTIVES ON HUMAN CAPITAL

Since its takeoff in the mid-1970s, TCE has continuously been subject to strong criticism, in the beginning mainly from Marxists (Marglin, 1974) and sociologists (Perrow, 1986), but increasingly from management scholars (e.g., Kogut and Zander, 1992; Ghoshal, 1994; Ghoshal and Moran, 1996). The treatment in the theory of the employment relation and human capital has been criticized in particular. It is characteristic of TCE that human capital is treated like any other asset, and not even its inalienability is emphasized. (In contrast, the inalienable character of human capital is a key point in the property rights approach). While theoretical generality is to be strived for, there is, of course, the fundamental distinction between human capital and other assets that only human capital is touched by intention, rationality, motivation, and so on. Implicitly, this recognition lies underneath many of the critiques of TCE.
TCE and Bounded Rationality

As mentioned, bounded rationality is a behavioral starting point for TCE. “But for bounded rationality,” Williamson argues (1996: 36), “all issues of organization collapse in favor of comprehensive contracting of either Arrow-Debreu or mechanism design kinds.” More generally, Williamson (1998) argues that taking more account of the relevant psychological literature will improve the understanding of organization “… as an instrument for utilizing varying cognitive and behavioral propensities to best advantage” (Williamson 1998: 12).

In spite of his insistence on the necessity of assuming boundedly rational behaviour, Williamson is in actuality rather vague on bounded rationality. He notes that “[e]conomizing on bounded rationality takes two forms. One concerns decision processes and the other involves governance structures. The use of heuristic problem-solving … is a decision process response” (Williamson, 1985: 46). The latter “form” is not central, however, in transaction cost economics, which, Williamson argues, “… is principally concerned … with the economizing consequences of assigning transactions to governance structures in a discriminating way.” Thus, Williamson is interested in making use of bounded rationality for the purpose of explaining the existence and boundaries of firms and therefore the choice between alternative governance structures rather than for the purposes of explaining “administrative behavior,” as in Simon (1947).

However, it is open to some debate whether it makes much sense to separate bounded rationality as an important ingredient in the understanding of governance structures from bounded rationality as the starting point for the understanding of decision processes, as different governance structures likely exhibit different decision process properties (March and Simon, 1958). Clearly, from an organizational theory point of view, the lack of concern with decision processes means that the important possibility that bounds on rationality may be endogenous to organization is not inquired into. Moreover, even from a mainstream TCE position, interest in more far-reaching notions of bounded rationality seems warranted. Thus, while TCE seems to presuppose that organization is highly flexible and plastic, the existence of, for example, endowment effects (Camerer, Heath, and Knez, 1993) among employees may complicate organizational change and make such change sluggish (for more speculation on these issues, see Foss, 2001).
TCE and Employee Motivation

While the role of bounded rationality in the theory of the firm has given rise to a fair amount of debate (e.g., Furubotn and Richter, 1997; Foss, 2001), it is nothing compared to the enormous amount of critical writings on the motivational assumptions in the theory. In particular, opportunism has been a favorite \textit{bête-noire}. Some such criticisms (e.g., that the theory assumes all agents to be opportunists) can be dismissed as demonstrably false. However, according to a recent critique, which is not so easily dismissed, a primary problem with the treatment of motivation in TCE is not opportunism \textit{per se}, but rather that modern economic approaches assume that all motivation is of the “extrinsic” type (Ghoshal and Moran, 1996; Osterloh and Frey, 2000). In other words, all behavior is understood in terms of encouragement from an external force, such as the expectance of a monetary reward. (In contrast, when “intrinsically” motivated, individuals wish to undertake a task for its own sake). These arguments do not necessarily deny the reality of opportunism, moral hazard, and so on, but assert that there are other, more appropriate ways to handle these problems than providing monetary incentives, sanctions, and monitoring. The arguments are often based on social psychological research (Deci and Ryan, 1985) and on experimental economics (e.g., Fehr and Gächter, 2000).

In one version of the argument, Ghoshal and Moran (1996) argue that the theory of the firm misinterprets the causal relation between motivation (e.g., the tendency to shirk) and the surrounding environment (the type of governance structure in place). They claim that individuals within an organization perform not according to the incentives and opportunities offered, but to their “feelings for the entity.” “Hierarchical” controls, they state, reduce organizational loyalty and thus increase shirking. Reliance on internal governance in the presence of relationship-specific investments, they hold, causes the very problems it is designed to alleviate: Williamson’s approach becomes a “self-fulfilling prophecy,” and is therefore “bad for practice.” Osterloh and Frey (2000) ask which organizational forms are conducive to knowledge creation and transfer. They note that elements of market control (e.g., high-powered incentives) are often introduced in firms to accomplish this. However, Osterloh and Frey argue
that this only works to the extent that there is no “motivation crowding-out effect,” in which extrinsic motivation does not crowd out intrinsic motivation (Deci and Ryan, 1985).

While these critiques are subtle and founded on important research from motivational psychology, it is somewhat questionable how much bite they really have as a critique of TCE. To be sure, intrinsic motivation etc. is not usually raised as an issue by TCE scholars. However, it has been a recurrent theme in TCE that the incentives that characterize internal organization are —and, indeed, generally should be — low-powered (Williamson, 1996). While the mechanism that supports this conclusion may be different (namely, multi-tasking agency theory, Holmström and Milgrom, 1991), the conclusion is the same as the one obtained by psychology-based arguments.

**TCE and the Capabilities View**

A recurring critique of TCE has concerned its neglect of productive capabilities (e.g., Kogut and Zander 1992). Differential capabilities give rise to different production costs, and that such cost differentials may crucially influence the make or buy decision: firms may internalize activities because they can carry out these activities in a more production (not transaction) cost-efficient way than other firms are capable of. The factors that make capabilities distinctive and costly to imitate, for example, complexity and tacitness, also mean that such differences in production costs may be long-lived. Knowledge-based writers also claim that the existence of the firm can be explained in knowledge-based terms and without making use of the assumption of opportunism (Hodgson 2004). Kogut and Zander (1992) argue that firms exist because they can create certain assets — such as learning capabilities or a “shared context”— that markets purportedly cannot create (Kogut and Zander 1992: 384). The problem with this argument is that it does not sufficiently characterize *firms*: Markets can cultivate learning capabilities and shared context (as in industrial districts) as well. If these assets are indeed largely internalized in firms, it is, from the TCE point of view, exactly because they are firm-specific assets that yield an appropriable quasi-rent. Teece’s (1982) early work on diversification, which merged Penrosian arguments (Penrose, 1959) with TCE, made exactly this point.
While TCE scholars may take issue with specific knowledge-based explanations the view does point to some weak points in the theory of the firm. For example, differential capabilities probably do play a role in determining the boundaries of the firm (Walker and Weber 1984; Monteverde 1995; Argyes 1996). However, there are two major problems in this area that may hinder progress. The first is that the nature of the central construct (i.e., capabilities) itself is highly unclear. It is not clear how capabilities are conceptualized, dimensionalized, and measured, and it is not clear how capabilities emerge and are changed by individual action (Abell, Felin and Foss, 2008). The second problem partly follows from the first: the mechanisms between capabilities and economic organization are unclear. This points to a need for understanding in a detailed manner how specific human capital assets combine inside firms to yield firm-level capabilities.

CONCLUSIONS

TCE represents the first systematic attempt to explore the organizational ramifications of human capital for economic organization, that is, the choice between alternative contractual and governance structures. While they are certainly taken up, with the rationale of the employment contract, Coase and Simon take no particular interest in human capital. In contrast, Williamson, building on earlier contributions by Gary Becker and labor market economists, argues that the services of specific human capital would tend to be organized through employment contracts. His treatment of such assets is not fundamentally different from his treatment of other specific assets; the “fundamental transformation” etc. may obtain for human capital as well as for any other kind of productive asset. However, Williamson clearly recognizes that the law regulating labor transactions are different from the law regulating other transactions. The property rights approach presents a complementary logic, based on the inalienability of human capital.

TCE is first and foremost intended by its creator(s) as an economics-based theory of the efficient choice of contractual and governance structures. It is not a management theory of, for example, human resource management, although the theory obviously does hold implications for important strategic management and organization issues, such as vertical integration, diversification, etc. However, in actuality, TCE has emerged as one of the important source
theories for much theorizing in management. This implies that the critiques that have been leveled against the theory over at least the last two decades for an inadequate treatment of, particularly, cognition and motivation need to be taken seriously. For example, a more adequate treatment of bounded rationality may lead to the identification and conceptualization of other sources of transaction costs than opportunism (Foss, 2001), transaction costs that are relevant to economic organization.
REFERENCES


2003-6: Marjorie Lyles, Torben Pedersen and Bent Petersen: Knowledge Gaps: The Case of Knowledge about Foreign Entry.


2003-9: Kate Hutchings and Sneijna Michailova: Facilitating Knowledge Sharing in Russian and Chinese Subsidiaries: The Importance of Groups and Personal Networks Accepted for publication in Journal of Knowledge Management.


2004

2004-1: Nicolai J. Foss: Knowledge and Organization in the Theory of the Multinational Corporation: Some Foundational Issues

2004-2: Dana B. Minbaeva: HRM Practices and MNC Knowledge Transfer

2004-3: Bo Bernhard Nielsen and Snejina Michailova: Toward a Phase-Model of Global Knowledge Management Systems in Multinational Corporations


2004-5: Teppo Felin & Nicolai J. Foss: Methodological Individualism and the Organizational Capabilities Approach

2004-6: Jens Gammelgaard, Kenneth Husted, Snejina Michailova: Knowledge-sharing Behavior and Post-acquisition Integration Failure

2004-7: Jens Gammelgaard: Multinational Exploration of Acquired R&D Activities

2004-8: Christoph Dörrenbächer & Jens Gammelgaard: Subsidiary Upgrading? Strategic Inertia in the Development of German-owned Subsidiaries in Hungary


2004-10: Jens Gammelgaard & Thomas Ritter: The Knowledge Retrieval Matrix: Codification and Personification as Separate Strategies


2004-12: Akshey Gupta & Snejina Michailova: Knowledge Sharing in Knowledge-Intensive Firms: Opportunities and Limitations of Knowledge Codification

2004-13: Snejina Michailova & Kate Hutchings: Knowledge Sharing and National Culture: A Comparison Between China and Russia

2005


2005-4: Dana B. Minbaeva: Negative Impact of HRM Complementarity on Knowledge Transfer in MNCs

2005-5: Kirsten Foss, Nicolai J. Foss, Peter G. Klein & Sandra K. Klein: Austrian Capital
Theory and the Link Between Entrepreneurship and the Theory of the Firm

2005-1: Nicolai J. Foss: The Knowledge Governance Approach


2005-6: Marjorie A. Lyles, Torben Pedersen & Bent Petersen: Closing the Knowledge Gap in Foreign Markets - A Learning Perspective

2005-7: Christian Geisler Asmussen, Torben Pedersen & Bent Petersen: How do we Capture “Global Specialization” when Measuring Firms’ Degree of internationalization?

2005-8: Kirsten Foss & Nicolai J. Foss: Simon on Problem-Solving: Implications for New Organizational Forms


2005-12: Koen H. Heimeriks: Alliance Capability as a Mediator Between Experience and Alliance Performance: An Empirical Investigation Into the Alliance Capability Development Process


2005-14: JC Spender: Management, Rational or Creative? A Knowledge-Based Discussion

2006

2006-1: Nicolai J. Foss & Peter G. Klein: The Emergence of the Modern Theory of the Firm


2006-3: Volker Mahnke, Torben Pedersen & Markus Venzin: Does Knowledge Sharing
Pay? An MNC Subsidiary Perspective on Knowledge Outflows

2006-4: Torben Pedersen: Determining Factors of Subsidiary Development


2006-6 Nicolai J. Foss & Ibuki Ishikawa: Towards a Dynamic Resource-Based View: Insights from Austrian Capital and Entrepreneurship Theory

2006-7 Kirsten Foss & Nicolai J. Foss: Entrepreneurship, Transaction Costs, and Resource Attributes

2006-8 Kirsten Foss, Nicolai J. Foss & Peter G. Klein: Original and Derived Judgement: An Entrepreneurial Theory of Economic Organization

2006-9 Mia Reinholt: No More Polarization, Please! Towards a More Nuanced Perspective on Motivation in Organizations

2006-10 Angelika Lindstrand, Sara Melen & Emilia Rovira: Turning social capital into business? A study of Swedish biotech firms’ international expansion


2006-12 John Holt, William R. Purcell, Sidney J. Gray & Torben Pedersen: Decision Factors Influencing MNEs Regional Headquarters Location Selection Strategies

2006-13 Peter Maskell, Torben Pedersen, Bent Petersen & Jens Dick-Nielsen: Learning Paths to Offshore Outsourcing - From Cost Reduction to Knowledge Seeking

2006-14 Christian Geisler Asmussen: Local, Regional or Global? Quantifying MNC Geographic Scope


2006-16 Nicolai J. Foss & Giampaolo Garzarelli: Institutions as Knowledge Capital: Ludwig M. Lachmann’s Interpretative Institutionalism

2006-17 Koen H. Heimriks & Jeffrey J. Reuer: How to Build Alliance Capabilities

2006-18 Nicolai J. Foss, Peter G. Klein, Yasemin Y. Kor & Joseph T. Mahoney: Entrepreneurship, Subjectivism, and the Resource – Based View: Towards a New Synthesis

2006-19 Steven Globerman & Bo B. Nielsen: Equity Versus Non-Equity International Strategic Alliances: The Role of Host Country Governance

2007

2007-1 Peter Abell, Teppo Felin & Nicolai J. Foss: Building Micro-Foundations for the Routines, Capabilities, and Performance Links
2007-2 Michael W. Hansen, Torben Pedersen & Bent Petersen: MNC Strategies and Linkage Effects in Developing Countries

2007-3 Niron Hashai, Christian G. Asmussen, Gabriel R.G. Benito & Bent Petersen: Predicting the Diversity of Foreign Entry Modes

2007-4 Peter D. Ørberg Jensen & Torben Pedersen: Whether and What to Offshore?


2007-6 Nicolai J. Foss: Strategic Belief Management

2007-7 Nicolai J. Foss: Theory of Science Perspectives on Strategic Management Research: Debates and a Novel View

2007-8 Dana B. Minbaeva: HRM Practices and Knowledge Transfer in MNCs

2007-9 Nicolai J. Foss: Knowledge Governance in a Dynamic Global Context: The Center for Strategic Management and Globalization at the Copenhagen Business School


2007-11 Nicolai J. Foss & Peter G. Klein: Organizational Governance


2008

2008-1 Kirsten Foss & Nicolai J. Foss: Managerial Authority When Knowledge is Distributed: A Knowledge Governance Perspective