Esa Viitamo

SERVICE PRODUCTIVITY – ECONOMIZING ON CONTRACTUAL COSTS

LAPPEENRANNAN TEKNILLINEN YLIOPISTO
LAPPEENRANTA UNIVERSITY OF TECHNOLOGY

TEKNISTALOUDELLINEN TIEDEKUNTA
TUOTANTOTALOUDEN OSASTO

FACULTY OF TECHNOLOGY MANAGEMENT
DEPARTMENT OF INDUSTRIAL MANAGEMENT

TUTKIMUSRAPORTTI 207
RESEARCH REPORT

TUTKIMUSRAPORTTI – RESEARCH REPORT 207

Esa Viitamo

Service Productivity – Economizing on Contractual Costs

LAPPEENRANTA UNIVERSITY OF TECHNOLOGY
Faculty of Technology Management
Department of Industrial Management

P.O. Box 20
FI-53851 LAPPEENRANTA
FINLAND

ISBN 978-952-214-710-3 (Paperpack)
ISBN 978-952-214-711-0 (PDF)
ISSN 1459-3173

Lappeenranta 2008

Contents

Preface

The need of transacting features all organized and unorganized social interaction. The transactions between individuals and adjacent stages of value chains facilitate the productive performance of economic organizations, such as firms and spatial economic systems. Like the productive inputs of standard production processes, transactions provide intermediary services, which contribute to the value added of the consumed products and services. Parallel to service activities, transactions are internal or external to organizations. The magnitude of external transactions can be approximated with the size of trade and distribution services to be about 15 % share of the GDP, which eventually has to be transacted as well.

Service activities and businesses are pervasive in modern economies. Statistics bear witness of an expanding service sector, which accounts for three quarters of the GDP in advanced economies. The abundant evidence on differing productive performance across service industries indicates divergent technological and institutional trajectories that outline the evolutionary progress of the tertiary sector. While conceptual knowledge on services and their performance has accumulated substantially, theoretical advances in the research on service productivity and competitiveness are still limited. The shortage is striking, bearing in mind the pervasiveness of the subject matter itself.

Based on previous contributions and new insights, *Productivity of Business Services – Towards a New Taxonomy* (Viitamo, 2007*)* develops the analytical framework of service productivity further. The approach in this report builds on the notion that definitions and classifications of services and performance measurement are strongly interdependent. Given the ongoing restructuring of business activities with higher information content, the report argues that the dichotomy between manufacturing and services should not be taken too far. Industrial evolution also suggests that the official industry classifications are increasingly outdated, and new taxonomies for empirical research are needed.

The present report *On Service Productivity – Economizing on Contractual Costs*, is a logical extension to the ideas presented in Viitamo (2007). The principal driver for this inquiry is the enduring need to widen the analytical perspective on service performance into new terrains of disciplines. The diversification into the field of economics of organization and in particular transaction cost economics looks promising in two respects. The organizational setting of a firm is an important determinant of the productive performance of internal activities, through which the performance of external services to the customers is markedly influenced. Owing to the intensive interaction between the producer and the user, the production and transaction costs of services are intertwined. In total, organizational costs are highly relevant for service production and delivery.

The report demonstrates that new insights into the fundamental economic issues of the established approach of transaction cost economics can be gained. The main interest is in the practical implications of the examined theories on service productivity, and how these contributions are related to the analysis conducted in Viitamo (2007), in particular. The value added in this regard is derived from the fact that the organizational approach is not intrinsically, or explicitly, focused on a firm's productivity, least of all service productivity. On balance, the organizational approach provides useful ingredients that can be utilized for further development of the theoretical framework on service productivity. In effect, the organizational approach confirms some of the fundamental premises on which the service management and marketing literature implicitly rests on.

1 Introduction

Organizations are ubiquitous, which is a dominant characteristic of modern societies. Organizations are not only instances influencing the activities of individuals, but also actors in their own right. "Just as firms in a specific context, organizations make decisions on behalf of the individuals they represent. In this way organizations are collective actors that take actions, use resources, enter into contracts and own property" (Scott and Davis, 2003, p. 6). In contrast to neoclassical economics, where organizations play a negligible role, the various fields of economics of organization are interested in the focal issues of how organizations are capable of performing the value creation task they are assigned to, and which of the characteristics of organizations are thereby decisive.

Economics of organization utilizes the analytical tools of the standard neoclassical microeconomics and the premises of various organizational theories[1]. While sharing the interlinked foci of profit seeking enterprises and policy implications with the "orthodox" industrial economics, economics of organization discriminates abstract modelling and focuses more on the observed behaviour of economic agents and institutions. Lack of methodological formality and unity notwithstanding, the organizational approach enables robust theorizing with a solid linkage to real world phenomena. For economic theorizing this means balancing between the indivisible hand of impersonal market forces (Smith, 1776) and the visible hand of deliberated action of corporations and entrepreneurs (Chandler, 1990) is strongly influenced by social and behavioural norms (Kay, 2000).

1.1 Organizations and economics

For bringing the realism of human behaviour into economic action, the organizational theory can be regarded as a generalized theory of strategic management. This becomes more straightforward when "the firm" in the structuralist and resource-based framework of strategic management (Viitamo, 2008; Porter, 1985; Grant, 1991) is substituted for a "complex organization" with the capability of making deliberate decisions and taking independent actions of their own (Scott and Davis, 2003). Organizations are commonly conceived as social structures created by individuals to support a collaborative pursuit of specific goals. Organizations must define and refine their objectives, and they must induce the participants to contribute services, which have to be controlled. As with business enterprises, resources have to be garnered from the environment and the products and services dispensed (Scott and Davis, 2003)[2].

Accordingly, in consonance with strategic management and the organizational theory, economics of organization regards firms as hierarchically structured, effective decision-making units, which pursue wealth creation through a diversity of competitive strategies. In addition to these similarities and the adherence to realism of human behaviour, strategic management and economics of organization show fundamental differences as well. Within the Porterian framework, for instance, the purpose of the choice of organizational form is to serve high profitability, and follows directly from the choice of an industry and the associated strategy (Porter, 1980). Organizational analysts, on the other hand, are more

[1] Economics of organization is often used synonymously with new institutional economics (Williamson, 1985).
[2] In some cases, a substantial proportion of the capacity of the resources has to be expended to maintain the organization itself, in which case the means become the ends (Scott and Davis, 2003).

interested in why firms choose the particular strategies and organizations they do[3], rather than the performance consequences of these processes (Scott and Davis, 2003).

For the organizational theorist, the key focus is on the organization. The chain of logic followed by organizational theories is best illustrated by the often cited "contingence theory", originally introduced by Lawrence and Lorsch (1967) and Thompson (1967). Equivalent to the structure-conduct-performance paradigm (SCP) (Scherer, 1980), which maintains that the characteristics of the industry structure determine the strategies and actions of firms, and thereby influence their performance, Donaldson (1995) summarizes the various strains of contingency theories with the analogous structural-adaptation-to-regain-fit (SARFIT) model (see Figure 1).

In an open systems world environments with inherent uncertainties create requirements for organizations, influencing the strategic choices of the management. Strategies in turn create contingencies - scale, technology, degree of diversification etc – for which some organizations are better suited than others. In the case of a mismatch between the organizational structure and the contingencies caused e.g. by technological change, the performance suffers, which triggers a search for restoring the fit. In short, organizational innovation conducive to an improved organizational fit should improve the performance, e.g. the overall resource productivity (Scott and Davis, 2003).

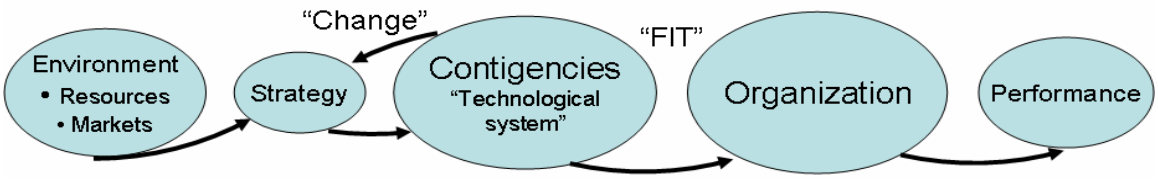

Figure1. Schematic presentation of the contingency model.

Variants of the contingency model are embedded – explicitly or implicitly – also in the theories of economics of organization. As noted by Dosi et al. (1998), organizational systems mediate the impact of technology on competitiveness. "Absent robust and adaptable organizational systems in firms, among firms and between firms and external institutions, the fruits of technology will become dissipated… conversely, well-designed organization structures and effective management are the handmaidens of competitive advantage, economic development, and growth" (op. cit., preface).

It should be noted that the organizational form is principally an endogenous variable within the context of strategic management, as well. Yet, organizational theorists and aligned economists, in particular, acknowledge the pervasiveness of organizations as governance institutions, ranging from multinational corporations to arms-length contractual relationships on the markets. For instance, markets and hierarchies (firms) may be regarded as viable alternatives for conducting transactions. Moreover, the diversity of the existing organizational forms reflects predominantly the limited cognitive capacities of the economic actors, in particular entrepreneurs and business managers (Williamson, 1985; Nelson Winter, 1982).

In strategic management literature, business managers are implicitly assumed to possess unconstrained capacities with respect to rationality and action towards the pursued level of profitability[4]. In general,

[3] This involves an assessment of why industries are structured as they are.

business managers are rational profit maximizers constrained only by the imperfections of the internal and external settings of the firm. Organizational economics, instead, rests predominantly on the behavioural assumption first introduced by Herbert Simon. "Human behaviour is intendedly rational but only limitedly so" (1961, p. 24). "The capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behaviour in the real world – or even for a reasonable approximation to such objective rationality" (Simon, 1957, p. 198).

Problems of bounded rationality arise because of the limited human capacity to process information on the alternative courses of action that are available to the actor, and the consequences of these actions. An immediate consequence of bounded rationality is that managers and entrepreneurs cannot be rational profit maximizers, but they are "profit seekers" and pursue "satisfying" performance. This is a distinct deviation from the profit maximization assumption, which the standard neoclassical theory builds on[5]. As will be demonstrated in the following sections, organizations are capable of dealing with (mitigate) the problems of bounded rationality and uncertainty in various ways. The characteristics of the contingencies guide the managerial choice among the organizational alternatives conducive to the "most satisfying", anticipated productivity.

1.2 Linking to service productivity

"There is no such thing as good organization in any absolute sense. Always it is relative; and an organization that is good in one context or under one criterion may be bad under another" (Ashby, 1968). The snapshot on the contingency theory and the references made to the organizational strategy above (Scott and Davis, 2003) demonstrate that the organizational context and changes in it are important drivers of competitiveness of the economic activity the organization is intended to perform. More generally, it can be maintained that the mainstream of organization theories are directly or indirectly concerned with the competitiveness and productivity implications of organizations[6]. As with the theories of firms, the study of organizational strategy seeks to answer the question "why do some organizations perform better than others?" (Scott and Davis, 2003, p. 310).

When it comes to the issue of organizational productivity, the distinction between organizational theories and economics of organization becomes artificial. The early manifestations of academic thinking on organizations already put forward the productivity of organized action. For instance, analysts that became later to be known as the founders of the rational systems school (Thompson, 1967) were primarily interested in what the proper form should be in the interest of maximizing efficiency and effectiveness, rather than in examining and explaining organizational arrangements per se. They also focused the primary attention on managerial activities and functions rather than the wider subjects of organizations and organizing (Guillén, 1994). Consequently, the performance assessment of organizations has assumed a prominent role in various sub-fields of organizational research.

---

[4] Strictly speaking this is not the case. If the assumption of perfect rationality holds, strategic management as a normative discipline is redundant.

[5] Conceptually, bounded rationality is a complex and controversial issue. More detailed accounts are provided e.g. by Fransman (1998) and Radner (2005).

[6] Logically, once the theoretical interest is geared to existing organizations that pursue private interests, one is implicitly dealing with the capabilities of their survival and prosperity.

Scott and Davis (2003) identify three general types of indicators and dimensions of performance measurement, respectively. First, there is a set of indicators related to the *outcomes*, where the focus is on the specific characteristics of materials and objects, which the organization has produced[7]. As noted by the authors, such indicators are regarded as quintessential indicators of effectiveness. Second, there are indicators that measure the *process*, and thus focus on the quantity or quality of activities carried out by the organization[8]. The advocates of process measures emphasize the assessment of inputs or energy regardless of outcome, and thereby focus on efficiency. As pointed out elsewhere (see e.g. Viitamo, 2007) effectiveness and efficiency are the main components of service productivity.

Third, there are indicators that measure the *structure* – or approximate the quality - of the organization. The purpose of structural indicators, such as the skills level of workers, or the proportion of the faculty with doctoral degrees, is to assess the capacity of an organization for effective performance. As suggested by Scott and Davis (2003), these three types of performance indicators can be ranked ordinarily with regard to their remoteness of the core issue of interest. In comparison with effectiveness, i.e. the purpose of the activity itself, the scope of the processes (efficiency) is removed from the outcome and assesses the economic use of inputs. In this regard the structures (capabilities) are even more distant to the eventual outcome. Accordingly, structure measures the organizational capacity to perform the processes needed to attain the intended outcome.

Structural indicators are viable for performance measurement, however, as they account for the value creation potential of service organizations (Løwendahl, 2005; Viitamo, 2007). More specifically, structural considerations are embedded in the value-based productivity analysis (Viitamo 2007; Van Ark and de Jong, 2004), which intertwines the premises of the competing theoretical schools of service productivity, demarcation and assimilation (Metcalfe and Miles, 2006; Salter and Tether, 2006; Viitamo, 2007). Moreover, the value-based approach presupposes that the productivity of any service activity consists of two components, efficiency and effectiveness. Together these components account for technological opportunities and constraints to attain simultaneously low unit costs and high quality for a unit of service, produced and delivered to the customer (Viitamo, 2007).

The technological constraint stems from the *stylized fact* that service productivity is characterized by an inherent trade-off between effectiveness and efficiency. This means that, around the optimum, a higher quality of the service for the customer leads unavoidably to higher unit costs of producing the same service[9], and vice versa. While manufacturing processes are principally subject to an equivalent constraint, there are essential differences entailed in the high degree of the intangibility of the service processes and the outcomes. This results in higher uncertainty, and more limited preplanning and control of the service production relative to manufacturing processes with tangible products. Furthermore, owing to the higher flexibility of the resources and service technologies, the trade-off for services should indicate a higher degree of continuity approximated by a constant productivity frontier (Viitamo, 2008).

---

[7] Such outcomes are e.g. the reliability of product functioning, sales, changes in the health status of patients, etc.

[8] In this context quality refers to a pre-determined standard of the service or product. The performance of a process can be measured e.g. by the number of cars produced per day, the accuracy and completeness of the medical history taken etc. (Scott and Davis, 2003).

[9] More specifically, the similarity of services implies that services with similar or different cost structures are used for the same purpose.

The purpose here is to address how the complexity of service productivity can benefit from the organizational approach to a firm's competitiveness and productivity. In particular, the analysis implements a contractual perspective of transaction cost economics to the focal question of how organizational choices affect productivity, and are affected by productivity considerations. Economic exchange, the basic unit of analysis, "is fundamentally about service provision" (Vargo and Lusch, 2004, p. 326). Of specific interest are the ramifications of the transaction cost theories, which, through the contingency propositions, bear on the effectiveness and efficiency of service productivity. In general, the chosen organizational context, contingent of internal and external determinants, stands for a generic resource that provides services to the productive activity of the organization. Moreover, organizations are composed of complementary activities and assets, which provide services internally to the organization and externally to the clients it is supposed to serve. This creates a dual approach of "service organization" and "organizational services", as conveyed throughout the analysis below.

2 Theoretical Foundations

The exploratory orientation of organizational research suggests that even at the most generic level, the focus is to a high extent geared to the productivity ramifications of organized activity. Undoubtedly, the emphasis on the productive performance of business organizations, such as firms, becomes pronounced in the fields of economics of organization, transaction cost economics and evolutionary economics in particular. Of the numerous sub-fields of economics of organization, transactions cost economics is perhaps the most affiliated with the general organizational theory, as the argumentation of both disciplines rests heavily on the logic of the contingency theory and the related motives for internalizing and externalizing "service" activities and services provided by specific assets (Williamson, 1985; Thompson, 1967; Penrose, 1959).

2.1 Strategizing and economizing

In contrast with the "power perspectives" of strategic management and to a lesser degree with the organizational theory as well, transaction cost economics – in its original setting - puts exclusively forward the motives of economizing on the costs of organizational design in the context of private profit seeking. Theoretically that implies, at the intentional level, improved cost-efficiency and productivity, which through the externalities should benefit the economic system as a whole. Yet, several authors advance the view that the policy implications based on transaction cost analysis are of high relevance for strategic management as well (Rumelt et al., 1991; Jones and Hill, 1988). Reflecting such a view, economy is seen as the "best strategy" (Williamson, 1991) delimiting the options for strategic and opportunistic behaviour.

Owing to the generic focus on contractual relations, economic literature affiliated with transaction cost economics abounds. This means that the premises and implications of the discipline have also faced a wealth of resistance. While an exhaustive critique of the theory is overlooked here, the essence of the common pitfalls, such as problems related to sub-optimization is kept in mind throughout the discussion[10]. In particular, it is an undeniable fact that economizing and strategizing are most often inseparable motives of the organizational design, which, moreover, should involve the assessment of transaction and production costs simultaneously. On the other hand, the benefits of the limited foci on economizing and costs of transacting lie in the normative implications derivable from the simplified world of reality. For this purpose the theoretical treatment here will be selective and focused on the "niches" most relevant to the issue of service productivity.

The essence of transaction cost economics is best conveyed by the features that distance it from "orthodox", neoclassical economics[11]. Logically, a world with negligible costs on transacting is characterized by an atomistic production and competition, where organizations and ownership do not matter. Technology and the size of the markets, as well as characteristics of demand would determine the optimal scale and scope of production, and the extent to which productive units are technologically integrated, horizontally or vertically. In these circumstances a firm may be plausibly defined as a "black boxed" production function, and its behaviour is predictable by gravity towards a partial or a

[10] For critical assessments of transaction cost economics see e.g. Williamson (1985), Winter (1993) and Dosi et al., (1998).
[11] As with the majority of the disciplines under the economics of organizations, the origins of the transaction cost approach is rooted in dissatisfaction with the unrealistic assumptions of the neoclassical theory of the firm.

general equilibrium (Varian, 1984). Yet, the real world and the economic systems are far from such an ideal. Leaving strategic considerations aside, the observed diversity of organizational forms, including multi-divisionalized firms and inter-firm contractual relations, reflect the pervasiveness of transaction costs, which burden all economic activity in a discriminating manner.

It is misleading to think, however, that economic analysis of the firm and its size is confronted by an unresolved contradiction with abstraction and realism. As concluded e.g. by Viitamo (1996), the neoclassical theory deserves its merits for identifying the circumstances where technological integration is feasible, be the outcome conducive to a higher economic efficiency or not. The characteristics of technology and markets alone are not sufficient rationale for a common (integrated) ownership, however. That is, the degree and mode of unified control of inter-dependent activities are determined by the transaction costs associated with the alternative modes of coordinating them (Viitamo, 1996; Teece, 1986a). Analytically, it is convenient to think that the alternative organizational (contractual) modes, while resulting from discrete choices, locate along the continuum between markets and hierarchies.

Most of the significant advances in transaction cost economics took place in the 1970s and 1980s. Since the era of the "intellectual boom", path-breaking contributions have been few, and the scientific progress has been manifested mainly in applied research within the market-hierarchy setting. Of the latest contributions, the most influential has been the work of David Teece and his holistic approach in *Profiting from Innovation: Implications for integration, collaboration, licensing and public policy* (Teece, 1986a). Referred to by a number of authors (Pisano, 2006; Winter, 2006; Nelson, 2006), this article can also be regarded as a "synthesis" of the theoretical and practical implications of transaction cost economics. As suggested here, the "synthesis" becomes comprehensible through inquiries into the key domains of transaction cost economics.

2.2 The nature of the firm

In retrospect, the main catalyst for the emerging strands within the economics of organization was the seminal paper by Ronald Coase, *The Nature of the Firm*, published in 1937. As an antecedent to the ideas of Penrose (1959), Coase regarded a firm and markets as two alternative ways of coordinating production and transactions[12]. "Outside the firm, price movements direct production, which is coordinated through a series of exchange transactions on the market. Within a firm, these market transactions are eliminated and in place of the complicated market structure with exchange transactions is substituted the entrepreneur-coordinator who directs production…it is clear that these (modes) are the alternative methods of coordinating production" (Coase, 1937, p. 388).

Given the alternative modes of organizing it is straightforward to conclude that the principal reason, why a firm is a more profitable alternative is that the price mechanism of the markets entails costs. In particular, efforts have to be expended to discover what the relevant prices are. "Also costs of negotiating and conducting a separate contract for each exchange transaction which takes place on the market must be taken into account…It is true that contracts are not eliminated when there is a firm but they are reduced" (Coase, 1937, p. 391). Hence, the advantage of a firm, in the form of cost savings relative to market arrangement, draws upon its higher organizational productivity, or rationality

[12] Coase (1937) asked rhetorically, why firms exist at all. Why are all exchanges not among the different parts of an organization left to the market?

(Thompson, 1967). In effect, when the direction of resources is dependent on an entrepreneur, long term adaptive contracts are substituted for short term rigid contracts. Moreover, the firm enables utilization of economies of scale in management and specific activities, such as marketing.

Within the transaction cost framework as outlined by Coase, the growth of a firm is defined in terms of an increasing number of internalized transactions and associated service activities. Given the inherent cost advantage of the firm, it is justified to ask, why all production is not carried out by one big firm. The fact delimiting endless expansion is that there are decreasing returns to the entrepreneurial (managerial) function, that is, the costs of organizing additional transactions within the firm may rise. Ultimately, a point must be reached where the costs of organizing an additional transaction within the firm are equal to the costs involved in carrying out the transaction in the open market, or to the costs of organizing by another entrepreneur (manager) (Coase, 1937).

To conclude, the marginal productivity of the firm with respect to the number of internalized transactions is assumed to rise at a decreasing rate, *ceteris paribus*[13]. Hence, in consonance with the Penrosean argument on the congestion, availability and quality of managerial services, which inevitably limit the growth of the firm (Penrose, 1959)[14], Coase puts forward the congestion of managerial services provided by an entrepreneur, which draw on the logic of marginalism and bounded rationality. Another factor causing a diminishing marginal productivity of management is that an entrepreneur fails to place factors of production in uses, where their value is greatest. Put differently, congestion of the managerial services weakens allocative efficiency (Coase, 1937). In effect, comparison of transaction costs of market organization and management costs of internal organization, "has become the focusing conceptualization of the transaction cost theory in most of its subsequent applications" (Demsetz, 1993, p. 162).

2.3 Team production

Another research paper contributing to the emergence and development of the transaction cost approach addresses the relative merits of market and non-market allocation of resources in the case of team production. The point of departure in the widely referred article *Production, Information Costs and Economic Organization* (Alchian and Demsetz, 1972) is the notion that economic organizations, through which input owners cooperate, "will make better use of their comparative advantage to the extent that it facilitates the payment of rewards in accord with productivity" (Alchian and Demsetz, 1972, p. 778). Hence, to induce productive effort and attain aligned incentives, the key demands placed on any organization are metering input productivity and reward.

In contrast with the make-or-buy situation and the vertical interdependence addressed by Coase, Alchian and Demsetz (1972) discuss the horizontal interdependence in team-like production. As for the technology in team production, the individual owners of productive inputs (labour) have an incentive to collaborate since their marginal productivities are enhanced by the efforts of the other team members. In this regard assets are technologically interdependent and specific to the team itself. With the terminology of Thompson (1967), team production is characterized by reciprocal interdependence which yields higher productivity than pooled interdependence with negligible cross-input effects on the

[13] The decreasing marginal productivity of a firm's productive resources is a standard assumption, which the neoclassical theory of firm rests on (Kreps, 1990).
[14] In general this is called the "Penrose effect" (see e.g. Teece, 1982).

overall productivity. Yet, the comparative advantage of team production is partly outweighed by higher coordination costs relative to the technology of pooled interdependence.

Consequently, Alchian and Demsetz (1972) conclude that the monitoring of inputs within the technological regime of pooling interdependence is cost-efficient, as the productivity and rewards of each input can be aligned instantaneously by market coordination. In team production, however, technological interdependence disguises the contribution of each input on the whole, which necessitates monitoring the inputs. To follow the efficiency logic of the contingency theory, the monitoring and metering costs call for organizational arrangements for which the team departs from a closed system (Scott and Davis, 2003). Owing to the monitoring costs and asymmetrical information among the team members on the individual contributions, there is a common incentive to shirk, i.e. to behave opportunistically (Williamson, 1985). This is the built-in source of inefficiency of uncoordinated (decentralized) team production.

In these circumstances, Alchian and Demsetz (1972) suggest that the team members have an incentive to appoint a controller, who specializes in monitoring the services of input. The incentive arises from the greater productivity, "the fruits of which will be enjoyed by all team members that will follow from the reduction of shirking as a result of advent of monitoring and controlling" (Fransman, 1998, p. 153)[15]. In effect, to maximize the productivity of the team, the monitor should be made a residual claimant for the profit net of the monitoring costs. Necessary conditions for the existence of what Alchian and Demsetz (1972) call a *classical firm* are then 1) synergistic team production exposed to opportunistic shirking and 2) the possibility to estimate marginal productivities by observing or specifying input behaviour. These preconditions should lead to a contractual arrangement, where the residual claimant utilizes economies of scale in monitoring to reach the maximum team productivity.

In reference to the models of a firm by Coase (1937) and Alchian and Demsetz (1972), a complementary note on service technologies can also be made. Namely, Teece (2003) posits that the Coasean firm is less eligible for professional services than the "classical firm" outlined by Alchian and Demsetz. In the Coasean firm "the boss must know as much as the expert, if the boss is to provide direction inside the employee's zone of discretion. This is clearly difficult if not impossible…Moreover, the expert's zone of discretion (in case of Coasean firm) is likely to be quire narrow" (Teece, 2003, p. 995), which contradicts with the managerial requirements for professional service organizations (Løwendahl, 2005). Consequently, the model of team production, with looser hierarchical relations, is more plausible in the case of professional services. In a similar vein, Alchian and Demsetz are suspicious on the appropriateness of a tight control of capitalist firm for the productivity of specific professional services. Namely, "dock workers can be directed in detail without the monitor himself loading the truck, and assembly workers can be monitored by varying the speed of the assembly line, but detailed direction in the preparation of a law case would require in much greater degree that the monitor prepare the case himself…As a result, artistic or professional inputs, such as lawyers, advertising specialists and doctors, will be given relatively freer reign with regard to individual behaviour…If the management of inputs is relatively costly, or ineffective, as it would seem to be in these cases, but, if team effort is more productive than separable production with exchange across markets, then there will develop a tendency to use profit-sharing schemes to provide incentives

[15] In particular, if there is a net increase in productivity available by team production, the net of metering costs associated with disciplining the team, then team production will be relied upon rather than a multiple of bilateral exchange of separable individual outputs (Alchian and Demsetz, 1972).

to avoid shirking" (Alchian and Demsetz, 1972, p. 786)[16]. Hence, in reference to the necessary conditions of a classical firm, it is evident that if the members of the team are assigned the residual claimant status, the productivity losses resulting from shrinking are partly mitigated, and the classical firm is not a necessary condition for an effective organization of team production.

2.4 Microanalytics

The theories of the firm outlined by Coase (1937) and Alchian and Demsetz (1972) were breakaways from the dominating neoclassical doctrine. As such they represent also two antecedent approaches for the analysis of a firm's productivity in the economics of organization. While the analysis of incentive alignment in team production initiated by Alchian and Demsetz paved a way for the evolving principal-agent approach (Jensen and Meckling, 1976) and the theories of incomplete contracts (Grossman and Hart, 1986), the development of transaction cost economics, as known at present, draws more explicitly on the intellectual heritage of Ronald Coase, and thereby the contractual characteristics of a vertical relationship between the buyer and seller. Perhaps more than a plausible theory of the existence of the firm, the merit of the Coasean theory appears to be the explanation of vertical integration under unified control and ownership.

Coase's treatise of a firm as a contractual response to transaction costs of the markets was further advanced and conceptualized by Oliver Williamson, who laid theoretical foundations for a discipline that became to be known as transaction cost economics (Williamson, 1975; 1985). To quote, "transaction cost economics is a comparative institutional approach to the study of economic organization in which the transaction is made the basic unit of analysis…It is interdisciplinary involving aspects of economics, law, and organization theory…As indicated, transaction cost economics maintains the presumption that organizational variety arises primarily in the service of transaction cost economizing" (Williamson, 1985, p. 387). Institutions become the mechanisms through which economic agents attempt to regulate non-cooperative behaviour. "By moving a transaction from one institutional setting to another, certain strategies may be precluded and thus specific costs avoided" (Masten, 1982, p. 3)[17].

By definition (Williamson, 1981), transaction occurs when a good or service is transferred across a technologically separable interface. Transactions are analyzed in terms of contracts, which reflect voluntary agreements between economic actors for some kind of agreed performance (Macneil, 1978), and associated transfer of property rights (Chandler, 1990). Broadly taken, transaction costs are the costs of contracting equivalent to friction in physical systems (Williamson, 1985). Cost incurred prior to the assignment of the contract, called *ex ante* costs, include the expenses (in time and funds) of drafting, negotiating, and safeguarding contracts. The costs posterior to the contract, called *ex post* contracting costs, include costs of mal-adaptation and haggling, setup and running costs associated with the governance structures and bonding costs (Williamson, 1985). In general, post contractual costs arise

[16] In practise this implies partnership, which is a weaker, or more decentralized, form of a "traditional" capitalist firm. To illustrate the applicability of the team production model for professional, services Teece (2003) notes that it is the experts that hire bosses rather than the other way round.

[17] The institutions of specific interest here are relational contracts between economic units that govern the ways in which these units can cooperate and compete. They are distinguished from the institutional environment, which is a set of fundamental political, social and legal ground rules that establishes the basis for production, exchange and distribution. These rules govern *inter alia* elections, property rights, and the right of contract (Williamson, 1998).

because their causes cannot be foreseen ex ante, or the contract does not induce credible commitments from the transacting parties.

The ultimate sources of the *ex ante* and *ex post* transactions costs can be traced to two critical characteristics of human behaviour assumed by transaction cost economics. In consonance with the organizational theory (Simon, 1957) the rationality of economic agents is bounded, which implies *inter alia* that all contracting is unavoidably incomplete[18]. More specifically, bounded rationality assumes that economic behaviour is intendedly rational but limitedly so (Simon, 1961, Radner, 2000), as cognitive capacities come short of the economic objectives[19]. Furthermore, economic agents are assumed to behave opportunistically, which implies self-interest seeking with quile[20]. Williamson (1985) argues that opportunism is responsible for real or contrived conditions of information asymmetry, which vastly complicates the problem of economic organization. To correct the Williamsonian argument, Fransman (1998) comments that the key problem is not the asymmetrical information per se, but the costs of opportunistic behaviour compounded by asymmetrical information and bounded rationality.

Transaction costs become effective through the key dimensions by which the transactions differ. These dimensions are 1) the frequency with which they occur, 2) the level and type of uncertainty which they are subject to, and 3) the degree to which transactions are supported by durable, transaction-specific investments (Williamson, 1981; Williamson, 1993). When the frequency of a specific transaction conducted on the market increases, higher productivity is attained through organizational investments that facilitate a cost-efficient conduct of recurrent transactions[21]. As discussed above, uncertainty results directly from the behavioural characteristics of bounded rationality and opportunism. Assuming homogeneity of the economic actors in this respect, the differences in uncertainty should reflect the complexity of the transaction and task environment (Thompson, 1967, Masten, 1982)[22]. In particular, the management of post-contractual opportunism and uncertainty dictates the organizational efficiency of transacting.

The key determinant of transaction costs, which influences the choices among organizational alternatives, is asset-specificity. Asset-specificity in contractual relations, usually expressed as a continuous variable k, is the degree to which a productive asset of a buyer or seller can be redeployed in alternative uses and by alternative users without sacrificing its productive value. Asset-specificity is manifested by sunk costs which deter redeployability (Williamson, 1993) and the creation of quasi-rents (Klein et al., 1978)[23]. In principle, asset-specificity can arise in any of three alternative ways. Site-specificity means that the successive production units are located geographically close to each other, whereas physical asset-specificity indicates that the productive units are technologically

[18] In this regard contracts are inherently open systems (Scott and Davis, 2003)
[19] This means that the maximizing behaviour of consumers and firms is only an abstraction.
[20] More specifically, opportunism refers to incomplete or distorted disclosure of information, especially to calculated efforts to mislead, distort, disguise, obfuscate or otherwise confuse.
[21] An example is the utilization of scale and scope in the distribution of specific products and services.
[22] Note that in the absence of opportunism, the choice of organizational form is contingent on bounded rationality as suggested by Thompson (1967) and Scott and Davis (2003).
[23] Logically, quasi-rents are the outcome of the sunk investment (costs) in the relation-specific assets. In general, appropriable quasi-rent constitutes the additional profit created, appropriable by the seller or the buyer without terminating the bilateral business relationship.

interdependent. Human asset-specificity (human capital) refers to the case where employees' skills are specialized in relation to the employing firm or a customer firm[24].

[24] In a similar vein, organizational capabilities, i.e. the routines of a firm can be specific in relation to the customer firm. This domain of asset-specificity is not dealt with in the transaction cost literature. A fourth type of asset-specificity is of a dedicated form, where the assets are completely specialized at the outset.

3 Productivity implications

As with strategic management literature, transaction cost economics is not explicitly concerned with the productivity of a firm, or the productive ways of employing the specific resources it possesses. In a way this is logical, as the ultimate causes of transaction costs, bounded rationality and opportunism, are related to the external transactions between the production units, which are often beyond the immediate control of the firm itself. Yet, the issue of productivity should be of high relevance to transaction cost economics as well, since transaction costs, seen as "economic friction", suggest a wasted use of resources. The resources wasted in this case are mostly human. With a straightforward interpretation, transaction cost economics is implicitly concerned with the productivity of "labour services" which are employed in the planning and execution of business transactions, as well as the productivity of the transaction-specific assets which may be human or non-human.

While the allocation of labour services refers to the organizational dimension of productivity, asset productivity, while contingent on the organizational dimension, deals explicitly with the technological productivity of the services and products transacted. In reference to the duality of costs and output analyzed in the neoclassical theory of the firm (Varian, 1984, Kreps, 1990), transaction cost economics takes a cost-oriented view on productivity with the emphasis on efficiency. That is, given the specifications of the outcome that the assets and the associated transactions are expected to generate, the objective is to minimize, within the limits of bounded rationality, the use the labour services and cost of investments in the specific assets and asset-specificity. As pointed out below, such a view contrasts with output-orientated productivity and assessment of how labour services and specific assets, with a given cost structure, can generate an enhanced outcome. Of particular interest is the quality (effectiveness) of the transactions and the products and services thereby transacted.

3.1 Transactional efficiency

In the spirit of organizational approach it is appropriate to assume that a "realistic" approach to productivity analysis incorporates all the costs and productive resources involved in the production of products and services. This means that transaction costs, whether internal or external to the organization, add to the production costs and thus constitute an essential cost component in assessing the overall efficiency and effectiveness of any organization. An illustrative case in this regard is provided by Alfred Chandler's *Scale and Scope* (1990), which is a historical analysis of the emergence and dynamics of industrial capitalism in the USA, Germany and the UK. Chandler's extensive analysis on the success and failure of evolving corporations suggest that strategizing and economizing are intertwined and complementary means of business policies.

In the 19th and 20th centuries, competitive corporations pursued cost reduction and efficient use of resources through the utilization of economies of scale and scope in production and distribution, as well as a reduction of the costs of transactions involved. According to Chandler (1990), the costs of transactions are reduced by more efficient exchange of goods and services *between* units, whereas the economies of scale and scope are closely tied to a more efficient use of (production) facilities and skills *within* such units. To quote, "transaction cost economies are, of course, closely related to those of scale and scope…Just as changes in the processes of production and distribution within units have powerful impact on the nature of transactions between units (as they are defined through the contractual

relations), so do changes in contractual relations affect the operations carried on within units" (Chandler, 1990, p. 18).

Chandler's observations support the argument put forward here that transaction costs are indeed pervasive. They exert profound influence on the overall cost efficiency of any productive technologies based on sequential or reciprocal interdependencies between productive sub-units. Transforming the perspective from costs to performance, Chandler asserts that efficient coordination of throughput does not occur automatically. It demands constant attention of the managerial team or hierarchy. The potential economies of scale and scope, as measured by rated capacity, are the physical characteristics of production technology. The actual economies of scale and scope as determined by throughput are organizational, however. "Such economies depend on knowledge, skill, experience, and teamwork – on the organized human capabilities essential to exploit the potential of technological processes" (Chandler, 1990, p. 24). Put differently, organizational efficiency and capabilities are a function of the human capabilities and assets (bounded rationality), as well as the alignment of incentives to release the services of human assets (management of opportunism) at the individual level. As noted by Demsetzt (1993), productivity derives in part from transaction and monitoring cost considerations, but it also depends on other conditions that underlie the acquisition and use of knowledge.

Another reason why transactional efficiency[25] matters for productivity analysis, is that transaction cost analysis is capable of linking intra- and inter-organizational efficiencies to a broader analytical framework, which examines the efficiency outcomes of networks, clusters and industrial sectors from the organizational contingency perspective. While the theoretical advances have in so far been in partial analysis of individual transactions, the transaction cost approach on industry structures could well contribute to the equilibrium analysis of competitive markets (Kreps, 1990). A persistent challenge in this regard is to distinguish between the actual production costs, transaction costs of markets, and the managerial costs of internal governance (Demsetz, 1993). Accordingly, the methodological progress of transaction cost economics has not been achieved by the development of the techniques for measuring transaction costs directly but by the development of operationalizing hypotheses to suggest where transaction difficulties are likely to be severe (Winter, 1993).

Despite the methodological differences between the empirical and strategic analysis of Alfred Chandler, and the microanalytic perspective of Oliver Williamson, both analysts are, according to Granovetter (1998), intrinsically contingency theorists (Thompson, 1967; Scott and Davis, 2003). Chandler and Williamson "predict the balance between federations of firms and single amalgamated units to derive from the need to adapt to variations in technology, consumer demand and market structure" (Granovetter, 1998, p. 77). This is manifested by the unqualified pursuit of transactional efficiency. In general transaction cost economics urges managers and firms to organize economic activity so as to "economize on bounded rationality while simultaneously safeguarding the transaction against hazards of opportunism" (Williamson, 1993, p. 93). As a difference to the generic organization theories though, opportunism represents an added contingency, which influences the choice of the governance mode for a transaction.

[25] The term transactional efficiency, instead of transactional productivity, is used here to indicate the performance of individual transactions. This is because a) productivity is conceptually associated with production, b) cost management is more linked to efficiency, and c) effectiveness is not a plausible concept for the transactional performance within the setting here.

More specifically, the *behavioural principle* of transaction cost economics implies the objective of balancing between the transaction costs entailed in the external uncertainty of environment (managing bounded rationality), and the transaction costs entailed in the internal uncertainty of the contractual relationship (managing opportunism)[26]. For the market-mediated transactions governed by explicit contracts this implies a trade-off between effectiveness and flexibility, which determines the horizontal coverage (scope) of the contract. Another trade-off exists between the opportunity costs of being tied in an inflexible long-term contract, and the cost of negotiating a series of short term contracts in the condition of bilateral dependency. The latter trade-off determines the vertical (temporal) span of the contract (Masten 1982).

Strategically more significant is the *principle of organizational contingency,* which urges business managers to "align transactions (which differ in their attributes) with governance structures (the costs and competencies of which differ) in a discriminating (mainly transaction cost economizing) way" (Williamson, 1993, p. 95). This involves first evaluating the frequency, uncertainty and asset-specificity associated with the specific transactions, and second, it is necessary to identify and outline the alternative governance structures, i.e. a hierarchical firm, markets and hybrid modes, which the transactions might feasibly be assigned to. "The discriminating (efficiency-driven) match between transactions and governance structures plays a prominent role in both conceptual and empirical parts of the transaction cost economics research agenda" (Williamson, 1993, p. 96). For the principle of organizational contingency, three domains of strategic choices relevant to a firm's productivity are taken under a closer scrutiny here. These domains are vertical integration, diversification and corporate structuring.

3.2 Asset-specificity

The discretionary choice of the organizational mode for the governance of a transaction is reduced to the problem of make-or-buy, or vertical integration. To assess their relative efficiencies, it is essential to identify the trade-offs associated with the markets and hierarchies. Decentralized forms of organizations (e.g. markets) support high-powered incentives[27], which work efficiently when autonomous adaptation to disturbances is needed. Conversely, high-powered incentives are inefficient and costly in case of non-market governance, where cooperative adaptation is needed. In a similar vein, low-powered incentives characteristic of hierarchy are inefficient in autonomous adaptations but efficient in cooperative adaptation (Williamson, 1998). Accordingly, given the duality of incentives and holding the technology constant, three things occur when a transaction is moved out from the market and placed under unified ownership: ownership changes, the incentives change and the governance structures change (Williamson, 1985). Along with technology and the prices of products and services transacted, these variables are in practise jointly determined.

Within a strong, i.e. well-protected, property rights regime the relative efficiency of the governance structures is determined by the costs of producing and transacting, which are technologically interdependent. With a simplified assumption of constant technology and production costs, the choice of the organizational mode, with regard to the three attributes of transactions, is ultimately contingent

[26] As Williamson (1981) notes, this is not inconsistent with the imperative to "maximize profits" but it focuses the attention somewhat differently.

[27] High-powered incentives guide entrepreneurial (private) profit-seeking behaviour, while low-powered incentives guide the behaviour of hired managers and employees.

on the degree of asset-specificity. Put differently, the growth of frequency and uncertainty necessitates non-market coordination only when some degree of asset-specificity exists. Hence, in the absence of asset-specificity (k = 0), the transaction is most economically coordinated by markets, regardless of frequency and uncertainty. From the transaction cost perspective the advantage of markets rests on the high-powered incentives of buyers and sellers, which "elicit autonomous adaptation to any unanticipated contingency" (Williamson, 1991, p. 82).

The transaction cost advantage of the markets tends to diminish, however, when the assets of the buyer or seller become increasingly specific to the transaction. This may occur through a gradual process of learning-by-doing or through once-for-all investments in transaction-specific assets and proximate locations (Masten, 1986). In both cases competitive bidding among the potential sellers, prior to the initial contract transforms[28] into bilateral relationship *ex post*, where the identity of the seller and the buyer matters (Williamson, 1981). With a higher degree of asset specificity, which locks the seller and buyer into a bilateral exchange, the ex ante and the ex post transaction costs should rise. The seller and the buyer have to employ resources in writing detailed contracts, safeguard against potential opportunism, and mitigate the problems of actual opportunism effective in executing the contract. Hence, "absent opportunism the rationale for coordinating an exchange within a hierarchy would be substantially reduced" (Williamson, 1985, p. 31).
When asset-specificity continues to grow, the increasingly inefficient market procurement will be replaced by more complex forms of contracting, and in an extreme case by a hierarchical firm. The relative advantage of hierarchy over the market procurement is fostered, if the transaction is associated with higher frequency and uncertainty. The organizational choice, contingent on the degree of asset specificity is depicted in Figure 2, where the costs on the positive part of the vertical axis measure the sum of transaction costs of markets and the management costs of the hierarchy (Demsetz, 1993)[29]. In general, the market displays the highest cost advantage when asset-specificity remains between zero and point k^{hy}, whereas hierarchy (an integrated firm) is superior when asset-specificity exceeds point k^{hi}. For this region of asset-specificity the bilateral interdependence can benefit from cooperative adaptation of a firm (Masten, 1982)[30]. Hybrid modes of contracting are feasible in the middle ($c_1 - c_2$) of the *locus* of the cost-efficient, i.e. the most productive modes of governance, $\beta^m - c_1 - c_2 - c_3$ (see Figure 2).

[28] Williamson (1985; 1993) calls this a *fundamental transformation*.
[29] By assumption the hybrid forms are subject to both types of organizational costs.
[30] A specific advantage of the firm over the market is the ability to settle disputes efficiently, as well as the access to, and efficient management of, information.

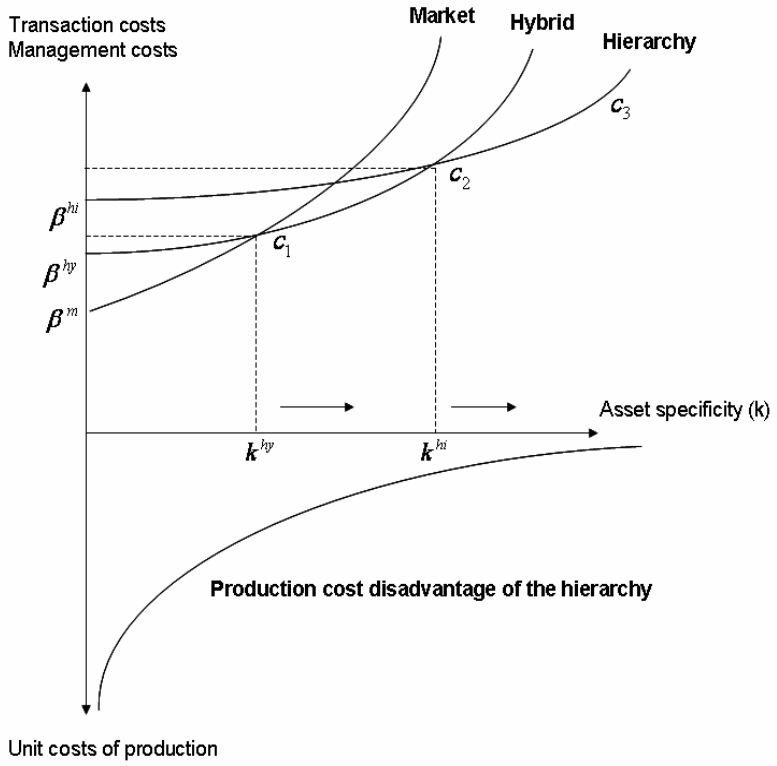

Figure 2. Organizational choice as a function of asset-specificity (modified from Williamson, 1991).

The organizational productivity of hybrid forms, such as long-term contracts, and the range of the comparative advantage, $k^{hy} - k^{hi}$ are contingent e.g. on the fixed costs of contracting, β^{hy} relative to respective costs of markets β^{m}, and hierarchy (firm) β^{hi}. They represent the bureaucratic costs of governance which, by assumption, are lowest for the markets, and highest for the hierarchy[31] (Williamson, 1985; 1991). In his more formal presentation Masten (1982; 1884)[32] shows that the feasible range of a long-term contract with regard to k depends *inter alia* on the durability of the transaction-specific investment, the size of the appropriable profits[33] generated by the joint production and external uncertainty. All these factors tend to make contracting more costly, which moves k^{hi} to the left in Figure 2. Uncertainty and the complexity of the product or service transacted may raise the opportunity costs of being locked into an inflexible contract, tending to increase the range of k, where markets and hierarchy are more productive forms of governance (Masten, 1982).

To conclude, asset-specificity is the key determinant of the degree of vertically unified control of adjacent stages or production. Interestingly, there is an analogy with the organizational technologies identified by the generic organizational theory, which distinguishes between pooling, sequential and

[31] The differing slopes of the cost curves for the management costs and transaction costs reflect the comparative disadvantage of the markets in cooperative adaptability with respect to the hybrids and with the hierarchy, as well as the comparative disadvantage of the hybrids relative to the hierarchy.
[32] The focus of Masten (1982; 1986) is the trade-off between transaction-specific investments and the costs of contracting, the trade-off between bargaining and vertical integration, and ultimately the factors which influence the choice among spot markets, extended contracting and internal organization.
[33] Bilateral profit is called quasi rent (Klein et al., 1978).

reciprocal interdependencies (Scott and Davis, 2003; Thompson, 1967). The contractual relationship examined here corresponds to what Thompson calls sequential interdependence between successive stages. Sequential interdependence is best coordinated by a *plan*. In case of market procurement with a low degree of asset-specificity, there is pooled interdependency between the seller and the buyer. Market-based transactions rely on autonomous adaptation, the coordination of which draws on the *rules* of markets. When asset-specificity grows, pooled interdependence gives way for reciprocal interdependence, which is more complex to coordinate and thereby assumes *mutual adjustment* (Thompson, 1967).

Thus far, the costs of production have been excluded from the analysis. Given the fixed output of the contractual relationship, the model implies that the hierarchy (firm) cannot outperform market governance in production efficiency. In effect, for all values of k, market is the preferred mode of governance[34]. This is mainly because a decentralized market – i.e. a specialized supplier – can aggregate diverse demands and is thus less limited in the utilization of economies of scale and scope than the integrated firm (hierarchy), which produces for its own needs only. The scale advantage shrinks, however, when the asset specificity and reciprocal interdependence increases. Highly specific assets cannot be redeployed in other uses without sacrificing their productivity value (Williamson, 1981). In Figure 1 this means that the production cost advantage enjoyed by the markets diminishes asymptotically to zero as the degree of asset-specificity grows. In Figure 2 the inclusion of production costs shifts k^{hy} and k^{hi} to the right.

The heuristic model discussed here illustrates the generic efficiency argumentation put forward by transaction cost economics. Namely, as long as there are no serious hazards in using market exchange, transaction should be left under the decentralized governance of markets, which is superior in production cost efficiency. Even when transaction is associated with moderate asset-specificity, uncertainty and frequency, a specialized supplier guided by high-powered incentives attains a higher productivity than a vertically integrated producer in particular, when the integrated firm uses the input internally only. Whereas the strategizing approach represented e.g. by Porter (1980; 1985) regards vertical integration as "means of power", transaction cost economics accepts vertical integration only as a "necessary evil" to mitigate the hazards on efficiency caused by bounded rationality and opportunism.

3.3 Scale, scope and effectiveness

The analysis above suggests a close affiliation of the ransaction cost approach to the contingency models of organizational choice by Thompson (1967) and Lawrence and Lorch (1967). Technology, as manifested by asset-specificity, together with uncertainty, plays a decisive role in selecting the most productive governance mode. A related and central issue, unexplained in Figure 1 is, however, what explains the origins of asset-specificity and its increase over time. It is rational to assume that asset-specificity does not occur by chance, exogenously, but because it is profitable and contributes somehow to the productivity of the vertically interdependent joint production.

[34] This is a strong and often criticized assumption. It implies that supplying and buying firms possess undifferentiated technological opportunities allowing no firm-specific advantages to exist (Demsetz, 1993). This contrasts sharply with the assumptions of the resource-based view on strategic management and evolutionary economics (Nelson and Winter, 1982), which stresses the uniqueness of the firm.

Comparative analysis of markets and hierarchies, where asset specificity enters as a costly but productivity-enhancing input would bring added realism to heuristic modelling. Utilizing the neoclassical profit maximizing framework, Riordan and Williamson (1985) show that when increased asset-specificity enhances productivity in vertical joint production, either through lowered unit costs of production or increased revenues generated by higher quality of the final product or a service, unified ownership (hierarchy) will produce more output with a higher degree of asset specificity than market governance (Riordan and Williamson, 1985)[35].

In particular, if the productivity growth that accrues to asset specificity is extensive, (Williamson and Riordan, 1985) demonstrate that vertical integration becomes increasingly profitable, and is progressively favoured over markets. In other words, if transaction and management costs approximate the functional forms depicted in Figure 2, unified ownership can be regarded as an effective means of safeguarding the productivity-enhancing impact of highly specific assets. Consequently, it is the incidence of transaction costs of markets that impede the optimal resource allocation. This is the central outcome of Masten (1984) as well. The reasoning of Riordan and Williamson (1985) holds even if the vertically integrated firm is for specific ranges of production volumes, subject to production cost disadvantage in comparison to market governance. With reference to Figure 2, such a "penalty" can be measured by the economies of scale lost under the unified ownership (hierarchy).

The implications of asset-specificity on economic efficiency has an interesting linkage to the productivity trade-offs between efficiency and effectiveness discussed by Porter (1998) in the context of strategic differentiation and cost leadership, as well as to quality and efficiency in the context of service productivity examined by service management research (see e.g. Grönroos and Ojasalo, 2004). To make the inference of Riordan and Williamson (1985) more generic, the central issue, beyond the focus of the authors, is to define the real cost of asset-specificity. As conveyed by the premises of transaction cost economics, the opportunity cost of asset-specificity is the reduced redeployability of the asset, which equals with potential economies of scale and scope that remain unutilized in the alternative first-best allocation of the asset.

Hence, irrespective of the organizational choice (make-or-buy) the owner of the asset is faced by a choice between two sources of productivity, scale and scope, which are based on general non-specific assets, and effectiveness generated by specific assets. Transaction-specific assets should enable extraction of customized services from the assets, which are provided with a few clients, and in an extreme case, one client[36]. As Riordan and Williamson (1985), note this can be achieved either by tailored cost reduction or tailored increase in the quality of the product and service produced by the client. Contingent upon the type of asset specificity (locational, physical, human), which by assumption is a continuous variable, there exits a continuous trade-off between the first-best efficiency (scale and scope) and effectiveness, the combination of which (ess^{k*}, eff^{k*}) can be varied to generate a fixed level of maximum productivity. This is illustrated in Figure 3, where the level of productivity is assumed to grow as the productivity frontier shifts outward.

[35] As demonstrated by the reduced production and transaction costs, the model conforms to the economizing argument of transaction cost economics. That is, vertical integration, when implemented, should lead to an increased economic efficiency and welfare.

[36] This assumes that the choice of the degree of asset-specificity is technically possible and not too costly.

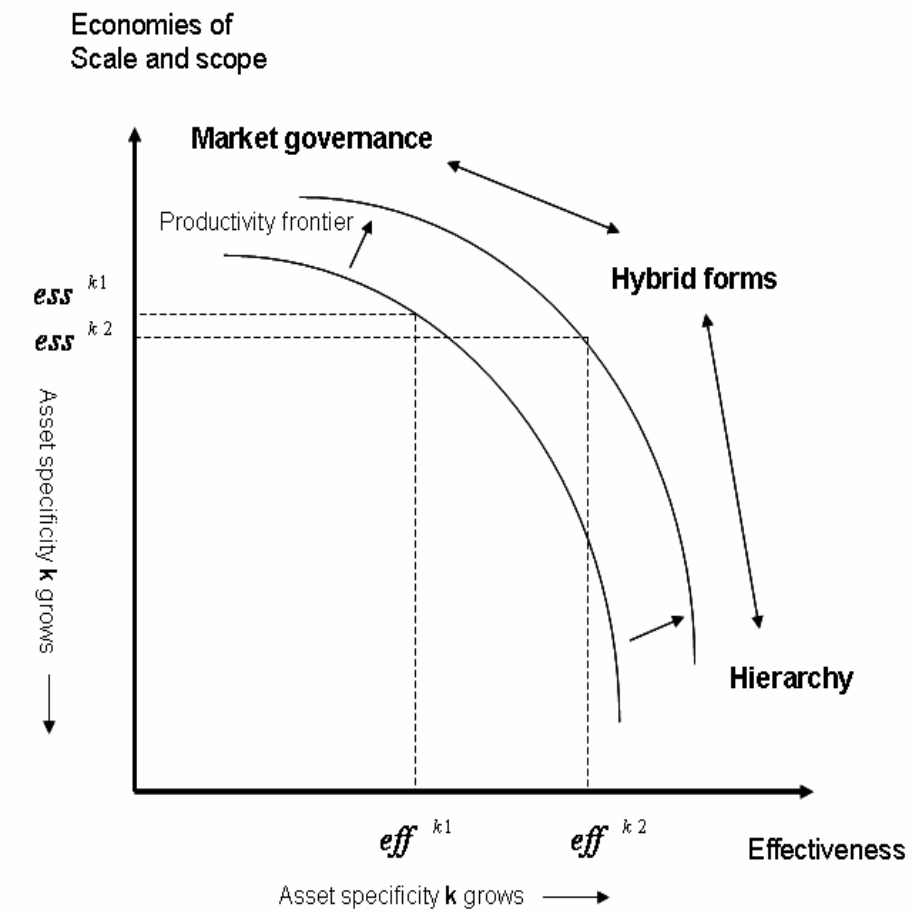

Figure 3. Linking asset-specificity to the analysis of service productivity.

Note that the trade-off is continuous to the extent that the degree of asset specificity k is a continuous variable as well. Consistent with the assumptions of transaction cost economics the point on the trade-off curve is determined by asset-specificity, which also indicates the number of customers served. The model approximates the theoretical setting which the literature of service management and productivity is implicitly focused on. Related to the make-or-buy issues, transaction cost economics predicts, moreover, that each combination of efficiency and effectiveness generated by the productive asset is associated with the most productive governance structure as well. As indicated by Figure 3, the governance is most often based on contractual hybrids, where the controlling mechanism varies between the markets and hierarchies. Yet, in comparison with asset-specificity and the combinations of efficiency and effectiveness, the choice of governance modes assumes a higher degree of discontinuity.

3.4 Quality and externalities

Related to the issues of effectiveness and quality raised by Riordan and Williamson (1985), transaction cost economics suggests another set of incentives to adopt a vertically integrated ownership structure. Such incentives rise to deter the loss of intellectual property rights under a weak property rights regime (Williamson, 1991; Teece, 1986a). While the transacting firms, locked in by specialized assets, may be capable of suppressing opportunism effectively, effective control over the relational innovations may still pass inadvertently into the hands of third parties (Williamson, 1991). In such a case, integration

into an adjacent stage by a firm may be to safeguard the source of competitive advantage irrespective of the degree of the transactional specificity of the assets.

Often created through internalization of an asset subject to a bilateral hold-up, the asset of concern here is specific to the firm and redeployable at related activities outside the core business and boundaries of the firm. In particular, this holds for human capital and knowledge, the acquisition of which is an irreversible investment and "usually less costly than retention and utilization of that knowledge" (Masten, 1982, p. 13). In market exchange such firm-specific knowledge is exposed to externalities "which may arise in conjunction with the intended or unintended debasement of quality for a branded good or service" (Williamson, 1981, p. 1549). The externality is thus a manifestation of measurement problems and costs associated with transactions (Barzell, 1982; Teece, 1982).

With regard to make-or-buy considerations, the problem of quality debasement appears typically at the interface between production and distribution. In particular, unintended quality debasement by an independent distributor poses a problem for the producer, if the actions taken by individual distributors affect one another[37], when one retailer's poor performance in customer service limits the sales of other retailers (Teece, 1984). While usually reflective of bounded rationality, the problem is compounded in the presence of opportunistic retailers. Namely, intended degradation of quality and reputation for a short-term profit increase may be a profitable strategy since the opportunistic retailer bears only part of the consequent costs.

Accordingly, Williamson concludes "if the quality enhancement (debasement) efforts of distributors give rise to positive (negative) externalities, the benefits (costs) of which can be incompletely appropriated by (assigned to) the originators, failure to extend quality controls over distribution will result in sub-optimization" (1981, p. 1549). Since all relevant information cannot be costlessly displayed and assessed, hierarchical governance (vertically unified ownership) should supplant market exchange. Product and service differentiation are typical situations, where externalities matter and may thus degrade the core assets of the producer's competitiveness (Williamson, 1981; Porter, 1980; 1985).

While franchising and forward integration become progressively desirable when externalities among retailers become significant, the argument is symmetric for backward integration as well. Teece (1984) suggests that producers of high-quality products and services ought to be vertically integrated backwards into the production of intermediate products when limited opportunity exists to develop an experience rating on suppliers, and when effective in-plant monitoring of suppliers' production activities involves significant costs.

3.5 Diversification and innovation

"Of all outstanding characteristics of business firms perhaps the most inadequately treated in economic analysis is the diversification of their activities" (Penrose, 1959, p. 104). Put forward by David Teece (1980; 1982), the argument acknowledges the inadequacy of the orthodox (neoclassical) economics to explain the existence of a multi-product firm. Whereas neoclassical theory of multi-product firm as outlined e.g. by Panzar and Willig (1981) and Baumol (1982), provides the necessary technological conditions, under which diversified production of related products and services is feasible, that is

[37] The reasoning here is analogous with team production with externalities and shrinking analyzed by Alchian and Demsetz (1972).

economies of scope, it does not provide a plausible explanation of why these activities should be coordinated under unified ownership. As with the market-hierarchy dichotomy in the case of vertical integration, diversification under strong and weak property right regimes is explained by economizing on transaction costs (Teece, 1982; Viitamo, 1996)[38].

The principal difference between vertical integration and diversification is related to the types of transactions being internalized (Teece, 1980). Whereas vertical integration involves mainly internalizing the supply of tangible inputs specific to a single production process, integration in the case of diversification involves internalization of the production of related products and services which share common know-how or other inputs supplied internally. To clarify the point further, horizontal or lateral integration is internal exploitation of the excess capacity exhibited by the firm's assets, which avoids the contractual hazards of market transactions between competing firms[39]. This is not to say that multi-product firms cannot emerge within an economy operating under neoclassical assumptions, but they do so only "by accident" (Teece, 1982, p. 41). Whether firms are organized along specialized or multi-product lines is economically irrelevant for the neoclassical theory, as markets and internal organization are perfect substitutes (Teece, 1982).

As indicated elsewhere, Teece advocates the efficiency doctrine in strategic management, which stresses the centrality of unique resources and dynamic capabilities for the productivity and competitiveness of the firm (Teece et al., 1997; Teece and Pisano, 1998). Combined with the transaction cost reasoning, such affiliations are reflective in Teece's efficiency-oriented theory on the multi-product firm and related diversification. The point made by Teece is that the decisive catalyst of diversified expansion of a firm is the firm-specific knowledge, which is difficult, if not impossible to trade. The difficulty follows from the notion of evolutionary economics (Nelson and Winter, 1982) that transfer of productive expertise across firm boundaries requires – to be effective – the transfer of complementary organizational as well as individual knowledge.

Firm-specific knowledge is often fungible to a high degree. That is, the human capital and routines employed by the firm are not always entirely specialized in the production of the particular products and services the firm is currently producing. Hence, a firm's comparative advantage is not defined in terms of the products or services offered, but in terms of the capabilities the firm possesses. The firm is seen as "establishing a specialized know-how or asset base from which it extends its operations in response to competitive conditions" (Teece, 1980, p. 233). As a corollary, "a firm's capability lies upstream from the end product – it lies in the generalizable capability which might well find a variety of final product applications" (Teece, 1982, p. 45).

In particular, prerequisites for the emergence of economies of scope are created by dynamic competition, involving uncertainty, entrepreneurial struggle and disequilibrium in the spirit of

[38] This rests on the argument by Williamson (1991) that "economy is the best strategy". The issue of diversification is a central focus in the strategic management literature as well. According to Porter (1985), there are two fundamental issues in corporate strategy for the diversified firm. The first is the selection of industries in which the diversified firm should compete. The second issue is how the strategies of the firm's business units should be coordinated. Accordingly, these issues are addressed in a similar fashion as more fundamental questions, namely, what is the attractiveness of an industry and how to create a defendable position against the five competitive forces (Porter, 1980). While conceptually strategizing, the motives for diversification at the operative level identified by Porter are economizing on common costs and tangible and intangible interrelatedness between the strategic business units.

[39] This is the standard logic in explaining the horizontal expansion of multinational enterprises (Teece, 1986b).

Schumpeter (1950). Dynamic competitive processes are, according to Teece (1982), characterized by knowledge accumulation through R&D and constantly changing market conditions with new profit opportunities. Complementary insights are provided by Penrose (1959) who stresses the importance of indivisibilities and managerial learning. These factors generate excess capacity and inducements for diversified growth delimited ultimately by the "Penrose effect"[40]. The organizational question is then how economies of scope transform into the scope of an enterprise, which calls for market-hierarchy perspectives.

In the presence of a strong property rights regime, where indivisible but non-specialized physical capital can be used as a common input in the joint production of several products and services, diversification (hierarchy) on the grounds of transactional efficiency is ill-founded. This is because redeployability of the asset deters hold-up and hazards of opportunism. Consequently, market governance based on simple contracting appears to be superior in such a case (Teece, 1982). Conversely, if the services of the excess capital are specialized, the markets for the capital are thin and therefore exposed to hazards of opportunism. In effect, contractual costs favour joint production under unified ownership. More generally, increasing the scope of the enterprise follows the same line of argumentation as vertical integration in the presence of asset-specificity.

When the source of economies of scope is human capital or proprietary knowledge, asset-specificity is less effective and the contractual relationship is faced with contingencies that approximate the quality externalities discussed in the context of vertical integration. With the "technological impediments" associated with inter-organizational transfer of embedded knowledge and tacit information, the utilization of economies of scope that draws on proprietary information is subject to specific costs arising from market exchange. The owner of the asset incurs costs of identifying the potential customers and ex ante transaction costs of reaching an acceptable agreement[41]. Similarly, if the effective transfer of technology and human capital requires services of expert teams coordinated by the seller, contingencies entailed in high incompleteness of the contract and pursuit of sub-goals will abound (Williamson, 1981). Moreover, if the transfer of proprietary knowledge also assumes recurrent transactions (high frequency), the overall productivity of the human capital calls for a hierarchical governance structure, i.e. diversification[42].

Analytically, the most prominent and consequential is the case where a weak property rights regime rules and the disclosure of proprietary information and information impactedness turn out to be the major problem. The seller and buyer have to be wary of the potential opportunism of each other, which entails the fundamental paradox of information (Teece, 1982; Arrow, 1971). That is, the value of the information for the purchaser is not known until he has received the information, but then he has in effect acquired it without cost. Diversified integration (hierarchy) is an effective and organizationally the most productive alternative to eliminate information externalities in the utilization of competitive advantage based on fungible proprietary information. The comparative advantage of hierarchy over market is increased by the presence of asset-specificity.

---

[40] The Penrose effect implies a congestion of managerial assets. The congestion can be mitigated by corporate restructuring and adoption of multidivisional structure.
[41] In the absence of bounded rationality these costs would be avoidable.
[42] This means e.g. a change from a licensing contract to a joint venture and further to an integrated firm.

The duality of economizing on transaction costs of asset-specificity and knowledge transfer is the main argument of David Teece, and at the same time the major contribution to the internal coherence of transaction cost economics. More generally, Teece provides a generic framework of economic integration, which rests on the propositions of transaction cost economics. Along with vertical and diversified integration, the same principles are applicable to horizontal integration, and the theory of multinational enterprises, in particular (Teece, 1986b). The synthetic view was made operative in the path-breaking *Profiting from Innovation* (PFI) (Teece, 1986a), which brings the analysis of markets and hierarchies on the agenda of strategic management and innovation. The mission of PFI is to delineate the technological conditions under which markets and hierarchy are the preferred modes of governance for firms to appropriate the profits from the innovations they succeed to generate[43].

With specific re-interpretations, it is postulated here that PFI implicitly summarizes the three efficiency principles that guide any organizational choice between the markets and hierarchy. The first prerequisite for efficient integration (hierarchy) is that commercialization of an innovation by a firm requires access to complementary assets. Complementarity of asset refers to the capabilities needed to perform primary activities such as manufacturing, distributing and marketing (Porter, 1985). The nature of the complementary assets in turn determines the direction of integration (vertical, lateral, horisontal). For the analytical purpose here this can be called the *principle of complementarity* (Grossman and Hart, 1986; Alchian and Demsetz, 1972), which is the source of technological efficiency and effectiveness.

A second prerequisite for feasible integration is that the services and technology of the complementary assets are specialized to the innovation, which is equivalent to high asset-specificity entailed by relational dependency on the transaction of the complementary asset. This may be thus called the *asset-specificity principle* (Williamson, 1985; Riordan and Williamson, 1985). A third prerequisite is the prevalence of a weak property right regime, which implies that the innovation cannot be protected by patents or trade secrets. Consequently, the technological and managerial information is exposed to information externalities of human capital (Teece, 1986a). This is the *principle of externality* (Williamson, 1981; Teece, 1980; Teece; 1982). Consequently, even if the complementary asset were specific to entail a bilateral hold-up, market procurement would be preferred under the strong property rights regime, which prevents the externalities appropriable by the competitors and the holder of the complementary asset.

Showing a weaker linkage with the core propositions of transaction cost economics, Teece (1986a) introduces a set of efficiency criteria stemming from the specific needs of corporate management. Hence, irrespective of complementarity, asset-specificity and externalities, the internalized asset has to be critical[44] for the utilization of the innovation. While essential, the condition of relative importance is not accounted for by transaction cost analysis, which examines each transaction in isolation. Undoubtedly, this entails in most cases in sub-optimization, which transaction cost economics is repeatedly accused for (Dosi et al., 1998). Relatedly, if the innovating firm exhibits a disadvantage in finance or a weaker competitive position (Porter, 1980; 1985), co-operation with better-positioned rivals and the holder of the complementary asset is needed (Teece, 1986, Teece, 2006).

---

[43] Relatedly, the paper offers strategic advice to business managers of how to avoid loss of the competitive advantage generated by the firm's innovation activity. In retrospect, Teece (2006) notes, however, that PFI does not try to explain how to build and maintain durable competitive advantage.

[44] Put differently, the relative importance of the complementary asset has to be high compared with the other assets and activities co-employed with the innovation.

A logical corollary of PFI is that big integrated corporations should enjoy a comparative advantage over small and medium size enterprises in systemic innovations (Teece, 1984), which require extensive design and coordination of complementary assets and activities. More specifically, when technological interdependencies are important, it is likely that the commercialization of an innovation will require complementary investments in several activities of the industry. If the needed complementary assets are highly dedicated to the innovation and risks of externalities abound, smaller companies dependent on market governance and networking are disadvantageously positioned. Autonomous, or stand-alone innovations, which are less dependent on a complementary asset, and do not require complex transactions between organizational units, will proceed efficiently in small unintegrated enterprises (Teece, 1986a; 1984).

3.6 Corporate restructuring

Along with the comparative analysis of markets and hierarchy for the conduct of single transactions and service activities that the firm is dependent on, corporate (re)structuring, geared to the efficient organization of internal transactions and service activities, is the second major field of analysis in transaction cost economics (Teece, 1984). In contrast with the issue of efficient boundaries and make-or-buy in the microanalytic context, the broader agenda of the internal organization of a corporation is less subject to the problem of sub-optimization. Another distinctive feature of corporate restructuring is the implicit assumption of path-dependency in the growth of the enterprise. In this regard history matters (Penrose, 1959). As the internalization of transactions proceeds, the wealth of corporate activities becomes as given without exposing them on recurrent make-or-buy assessments. Detailed cost evaluations are further complicated by the increasing complexity and interdependence of corporate activities brought about by the growth.

The internal organization of a corporation and restructuring became a popular agenda through the path-breaking studies of Alfred Chandler on US business history (Chandler, 1962; 1977). He observed that the intensive expansion of US companies in the 1920s, not only by scale but to a growing extent by scope with an increased number of business lines, led to organizational innovations necessary to reach a better match between the growth of the enterprise and managerial efficiency. To quote, "the inherent weakness in the *centralized, functionally departmentalized* operating company…became critical only when the administrative load on the senior executives increased to such an extent that they were unable to handle their entrepreneurial responsibilities efficiently…This situation arose when the operations of the enterprise became too complex and the problems of coordination, appraisal, and policy formulation too intricate for small numbers of top officers to handle both long-run, entrepreneurial, and short-run operational administrative activities" (Chandler, 1966, pp. 382 – 383).

In his later work, *Scale and Scope* (Chandler, 1990) that looks into the dynamics of industrial capitalism in the US, Germany and the UK, Chandler observed that the leading companies, "as they added units abroad or in the related industries, modified the enterprise's administrative structure" (op. cit. p. 42). For cost saving strategies – that is, the utilization of economies of scale and scope, and economizing on transaction costs – the modification of administrative structure was distinctively driven by the reduction of administrative transaction costs[45]. Thereby, "only after extensive expansion

[45] Administrative costs are defined here as consisting of transaction costs and management costs (Demsetz, 1993). They approximate the costs of bureaucracy (Williamson, 1985), or generally, costs of hierarchy.

overseas did an enterprise adopt a multidivisional structure by which major geographical regions were administered through integrated area divisions (Chandler, 1990, p. 43)[46].
Corporate diversification into related industries brought about far more deep-going administrative restructuring, and diversifying companies adopted, some more quickly than others, a multidivisional structure. Within a multidivisional structure "each division became a replica of the enterprise's original centralized, functionally departmentalized organization except that the highest ranking officer in the division had become a middle manager reporting to the top executives in the corporate office…Each division competed functionally and strategically with the other firms or with the divisions of other firms within the same industry" (Chandler, 1990, p. 43).

The empirical research by Chandler has been instructive in the process of understanding corporate restructuring as a phenomenon to be examined through the lens of a contract (Williamson, 2003). Transaction cost economics maintains that such organizational innovations are central to an understanding of the "modern" corporation. The study of intra-organizational innovations requires, however, that the details of internal organizations are examined (Williamson, 1989). As with the ramifications of external contractual relations, the agenda with regard to corporate restructuring is to understand the comparative efficacy of internal governance processes better. The properties of functional and multidivisional forms are more systematically conveyed by the organizational theory, where corporate structuring is addressed in relation to the contingencies within a more general setting.

In response to greater amounts of task complexity, uncertainty and interdependence, "organizational forms are likely to exhibit increasing differentiation, structural flexibility and capability of coping with increased information processing demands" (Scott and Davis, 2003, p. 130). In the bottom of the ranking - in an ascending order - are the simple structure and bureaucracy, which are followed by the functional form. All these forms presume singularity of purpose and unity of command. The *functional or unitary form* (the U-form) captures the logic of centrally coordinated specialization, based on economies of scale and scope over corporate activities. A distinctive feature of the U-form is departmentalization around varying specialized activities which contribute to the common goals. It includes hierarchically organized "line" departments, involved in activities directly related to producing or distributing goods or services, as well as more independent staff departments, involved in support matters such as accounting, finance, and personnel (Scott and Davis, 2003).

More advanced organizational forms, including multidivisional and matrix forms[47], and to a higher extent adhocracies[48] and networks[49], accommodate multiple objectives and divided authority. They represent a shift from a reliance primarily on buffering tactics and sealing out or suppressing

---

[46] The reason why Chandler juxtaposes diversification and multinational growth in scale is that they serve similar strategic purposes for the corporation. In particular, while vertical integration and horizontal growth in the domestic markets are defensive to protect existing investments, i.e. the technical core as Thompson (1967) puts it, diversification and multinational expansion utilize the existing investments and their existing organizational capabilities, i.e. their facilities and skills, to move into new markets and into new businesses (Chandler, 1990). Supportive arguments have been presented by Penrose (1959) and Teece (1982) as well.
[47] The *matrix form* is a dual-hierarchical form that organizes work simultaneously by functional and project criteria. Hence, it is a combination of a unitary and a multidivisional form. "Workers associated with diverse functional departments are regularly assigned to project teams organized to produce particular products and services" (Scott and Davis, 2003, p. 131).
[48] *Adhocracy* is characterized by low formalization and centralization and relies heavily on highly trained, independent, self-organizing individuals, who move in and out of project teams (Scott and Davis, 2003).
[49] *Network* allows integration of activities across formal boundaries, both within and across organizations (Scott and Davis, 2003).

uncertainty and variety from the core, to the use of bridging tactics and expanding boundaries to incorporate uncertainty within the core activities (Scott and Davis, 2003; Thompson, 1967). The organizational logic of the multidivisional form is coupling divisional autonomy with centrally controlled performance evaluation and resource allocation. More specifically, the *multidivisional form* (the M-form) is based on groupings by products or markets overlaid on functional forms. Divisional units operate in a relatively autonomous manner from each other, and each contains departments organized along function lines. The superordinate corporate level oversees divisional performance and allocates resources accordingly (Scott and Davis, 2003). Hence, operational decisions reside within the division, while strategic decisions are allocated to corporate headquarters.

The referred taxonomy of the organizational forms outlined by Scott and Davis is instructive in several ways. First, it helps identify the optimal structure in response to the demands of technology and the contingencies arising from the environment. For that purpose the taxonomy reflects the evolutionary progress of managing resources, activities and uncertainty within "modern" business corporations. Second, it parallels implicitly with specific industry characteristics. As the unitary, the multidivisional and the matrix forms are characteristic of manufacturing industries and scale-intensive services with an extensive scope of markets and products, whereas the adhocracy and the network conform to the organizational needs of various knowledge-intensive service industries. Third, it can be concluded that the organizational forms actually implemented are inherently hybrids and some combinations of the basic forms. For instance, big corporations are often structured by a unitary, divisional and matrix form, but they may exercise adhocracy and networks in the service activities related to internal and external transactions.

Given the diversity of organizational forms to manage business activities and the suggested explanatory framework of organizational economics, the interest in the spirit of Chandler is in the focal question of why and in what circumstances the multidivisional structure outperforms the functional form as a feasible corporate structure. For the principles that guide the choice between the organizational modes there exists a clear distinction in emphasis between the organizational theory and the applied transaction cost economics. While the organizational theory assesses objectively the appropriateness of the organizational forms within the contingency framework, transaction cost economics takes a more normative stance and proclaims the superiority of the multidivisional structure with a lesser emphasis on the organizational fit. Thus, given these differences a closer inquiry into the argumentation is warranted.

Transaction cost economics asserts that once internalized, there is no guarantee that the transaction and the associated activity will be effectively organized, given the hierarchical structure of the firm (Teece, 1984; Williamson, 1981). Bounded rationality and opportunism are ubiquitous, and the problems presented by both vary with changes in the internal organization. The policy recommendation for the corporate management rests on the principle of hierarchical decomposition, which urges a balanced fit between market and hierarchy. Accordingly, "internal organization should be designed in such a way as effect quasi-independence between the parts, the high frequency dynamics (operating activities) and low frequency dynamics (strategic planning) should be clearly distinguished, and incentives should be aligned within and between components so as to promote both local and global effectiveness"(Williamson, 1981, p. 1550).

The overriding logic behind the decomposition principle draws on the comparative efficacy of the alternative organizational modes in the face of corporate growth. The functional form, as noted by

Chandler (1990), is exposed to communication overload, and hence, bounded rationality as the congestion of managerial services compounds the "Penrose-effect". Decomposing and specialization of managerial responsibilities along product lines mitigates the costs of bounded rationality. The upshot according to Williamson (1989) is that the organizational innovation of the M-form, which has a mainly bounded rationality origin, has also unanticipated (side) effects on corporate purpose by attenuating sub-goal pursuit, i.e. opportunism.

The validity of the decomposition principle gains further support from the comparison with a third alternative, a holding company, the H-form. Whereas the H-form shows a similar pattern of decentralization as the M-form, its business units enjoy higher autonomy and are most often unrelated with respect of the markets and technology[50]. In contrast with the M-form, cash flows in the H-form are not reallocated between the competing divisions based on their relative performance (profitability), but instead are returned to the independent source divisions. Hence, the financial independence of the subsidiaries and the absence of effective cost control and market test is another source of inefficiency and opportunism, which entails a disadvantage relative to the M-form. From the corporate perspective both forms are decentralized and the subsidiaries hold assets, technology and organizational capabilities, which are specific to their current use, that is $k > 0$. Through the better utilization of internal capital markets and incentive alignment however, the M-form is able to safeguard the specific assets against hazards, which should lead to a higher organizational productivity relative to the H-form (Williamson, 1985).

To conclude, the competitive advantage of the M-form over the alternative forms draws on the "hierarchical separation and specialization of management, strategic planning and resource allocation capability supported by an efficient monitoring and control apparatus" (Williamson, 1985, p. 281). These capabilities enable resource reallocation from the less to the more productive uses within the corporation. In total, the adherence to the M-from by Williamson seems to be in consonance with his view on the individual transactions in the face of asset-specificity. Namely, market mechanism and high-powered incentives should be utilized as much as possible to attain production efficiency, whereas some degree of hierarchy is a necessary evil to delimit opportunistic sub-goal pursuit of the constituent parts of the corporation.

As noted by Hoskisson et al. (1993), the argumentation on corporate governance has inspired a whole stream of empirical studies in economics and strategic management, but the evidence supports the M-form hypothesis only to a qualified extent. In particular, no unambiguous conclusion can be drawn that enterprises with the M-form should systematically outperform enterprises that implement the U-form or the H-form. This points, *inter alia*, to the deduction that corporate performance is a function of the managerial capabilities and skills in matching the corporate strategy with the required structural form. The pursuit of the organizational match is the essence of the contingency argument raised within the strategic management discipline (Hoskisson et al., 1993).

Interestingly, the contingency analysis of corporate structuring by strategic management draws distinctively on the argumentation of transaction cost economics (Hoskisson et al., 1993). The contingency argument applied in corporate structuring is, however, most explicitly formulated by organizational theorists. They maintain that given the complexity, variation and unpredictability of the task environment, there exits no "one best way" to structure complex organizations. In particular, the

[50] The H-form is characteristic of conglomerates which pursue unrelated diversification.

more heterogeneous the overall task environment is, the greater are the constraints (anticipated variation) presented to the corporation, and the more dynamic (uncertain) the task environment is, the greater are the contingencies (unanticipated variation) presented to the corporation (Thompson, 1967).

In other words, the overriding principle of organizational "fit" suggests that the structure of a corporation cannot exclusively be determined by internal requirements of coordination and incentive alignments, as put forward by Williamson (1985) and Teece (1984). The organizational theory emphasizes that whereas internal coordination is essential for the effective operation of the technological core of corporations and managing of the interdependencies between the core activities, adjustment to constraints and contingencies, which are mostly beyond the control of the management, is the other side of the coin, and cannot therefore be ignored. The presence of environmental contingencies sets demands to the productive organization of boundary-spanning activities, which is the prerequisite for reaching the highest possible level of bounded rationality[51].

Given these premises and the strategic domain of the corporation, the existence of the U-form, the M-form, and the matrix-form[52], can be justified on the grounds of differing heterogeneity and dynamics of the task environment, internal interdependencies (organizational technologies), as well as consequent needs for organizational differentiation, flexibility and information-processing demands (Scott and Davis, 2003)[53]. For this purpose the main proposition on organizational rationality and structure by Thompson is instructive. Namely, "under norms of rationality organizations facing heterogeneous task environments seek to identify homogenous segments and establish structural units to deal with each" (Thompson, 1967, p. 70). The main dimensions of such heterogeneity are geography, i.e. the number of markets served, and social composition of the environment, which refers to the variety of customers (segments), inputs, and other organizations the corporation is dealing with.

When the task environment is heterogeneous, which is a reasonably realistic assumption for all kinds of corporations, but *stable*, organizational efficiency assumes several functional divisions (specialized production, distribution, procurement etc) capable of coping with the diversity of the constraints. The adaptation of the functional divisions to the environment is based on standardized responses and rules[54], which enable utilization of economies of scale and scope as well. Given the heterogeneity-stability conditions Thompson proposes that "when technical-core (production) and boundary-spanning activities can be isolated from one another except for scheduling, organization under norms of rationality will be centralized with an overarching layer composed of functional divisions" (1967, p. 75). Under these the conditions the U-form is the most productive structure of corporate governance, where organizational productivity draws principally on *scale-driven efficiency*.

Other things being equal it can be concluded that the uncertainty of the task environment dictates the feasibility and organizational efficiency of the M-form. That is, when the task environment is both

---

[51] The boundary-spanning activities of a firm can be defined as the opposite ends of the value chain internalized by the firm (Porter, 1985). The internal activities at the opposite ends are linked to external activities of other firms and organizations. For instance procurement is linked to the external sales of the supplying firms and sales are linked to the external procurement of the customers of the firm.

[52] As a hybrid of the M-form and the H-form a closer analysis of the matrix form is left for forthcoming analysis of the productivity of the universal banking industry.

[53] "Process specialization will be carried furthest in stable environments…under rapidly changing circumstances specialization will be sacrificed to secure greater self-containment of separate programs" (March and Simon, 1958, p.159).

[54] Standardized rules are characteristic of internal coordination of pooled interdependence as well.

heterogeneous and dynamic, the adaptation based on rules is displaced by a continuous monitoring the environment and planning responses, which calls for a decentralized corporate structure. Accordingly, "under conditions of *complexity* (heterogeneity and dynamics), when the major components of an organization are reciprocally interdependent, these components will be segmented and arranged in self-sufficient clusters, each cluster having its own domain" (Thompson, 1967, p. 73).

For the organizational theory, the M-form is an organizational response to the dual needs to adapt to a heterogeneous and shifting environment, and manage reciprocal interdependence between the technical core and boundary-spanning activities. Whereas the division of the technical core and boundary-spanning components in clusters of separate profit centers is an effective means of managing bounded rationality, the overall rationality of the M-form in comparison with the U-form becomes eventually more constrained. This is just because the technical core, which represents the rational system itself, cannot be separated from the boundary-spanning activities in the M-form. As by definition the boundary-spanning activities follow the open system logic (Thompson, 1967; Scott and Davis, 2003), the technical core of the M-form is more exposed to the external contingencies. To conclude, in contrast with the U-form, the organizational productivity of the M-form draws principally on *effectiveness and adaptation to the locally differentiated markets*.

4 Services, Trust and Rationality

The above discussion on the premises and theoretical refinements of transaction cost economics is related to the broad issue of transactional efficiency and organizational productivity. In this regard the analysis is equally well applicable for service industries and manufacturing industries. In both cases the internalization of activities embraces the transactions of the productive "services" provided by the specialized assets and idiosyncratic investments. Moreover, it was demonstrated above that the phenomenon of asset-specificity is conceptually closely linked to the efficiency-effectiveness framework, which constitutes the theoretical basis of the research on service productivity and management.

4.1 Service externalization

Aside from the considerations taken up above, applied transaction cost economics is not explicitly focused on the specific characteristics of service industries or technologies *per se,* least of all the organizational ramifications of transactional efficiency. The theoretical and empirical negligence of services notwithstanding, transaction cost considerations should be of high importance for the issue of service productivity as well, and some initiatives to provoke further research will be made here. Interestingly, such considerations are evoked from the critical assessments and relaxation of the behavioural assumption of the standard transaction cost analysis.

In effect, the marked development for the last 40 years has been the outsourcing of service activities by industrial corporations as well as small and medium size enterprises. This illustrates the strategically important but scientifically less examined facet of the make-or-buy issue. In a longer perspective, outsourcing of service activities has played a prominent role in the creation of new service businesses and the growth and development of service industries and sectors in advanced economies (Gadrey and Gallouj, 2002; Viitamo, 2003). Undoubtedly, changes in the determinants of transactions – asset specificity, frequency and uncertainty - have each contributed to vertical disintegration in the manner predicted by transaction cost economics.

From a broader perspective, the assets of in-house service activities have, through technological development and industry evolution, become less firm-specific, though at the same time their specificity with regard to the service activity has increased. As a result, the strategic importance of in-house service activities relative to the core capabilities of the firm has diminished (Teece, 1986a). These developments have fostered the advantages of specialization and division of labour, skills and human capital, accordingly. In these circumstances transaction cost economics predicts that market coordination should replace hierarchy as an efficient governance mode in organizing service transactions. Diverse demands of outsourcing enterprises can be aggregated, which supported by higher frequency of market transactions, creates economies of scale and scope appropriable by specialized service suppliers. The evolving competitive markets of business services have diminished uncertainty in the availability and quality of the services.

4.2 Professional services revisited

In concordance with Miles (2003) and Løwendahl (2005), it is suggested here that transaction cost analysis may provide helpful insights and conceptualization for the examination of the characteristics

of services, professional services in particular. Pointing out to organizational implications, Teece (2003) notes that there exists a requirement for both complementary and supporting services in many engagements, which has important ramifications for how expert work should be organized. "As with traditional firm, coordination must be achieved and conflicts must be expunged [in professional service firms]" (Teece, 2003, p. 897). As put forward by the present author (Viitamo 2007; Viitamo 2008), the industry characteristics of professional services represent dimensionally the opposites of standard manufacturing. This makes it theoretically reasonable to focus on the issues related to the professional services which, as suggested here, should generate useful insights on service businesses and related activities in a more general setting.

From the perspective of transaction cost economics, the specific characteristics of professional services are particularly appealing. The services are produced and delivered by highly qualified individuals, who invariably retain high-powered incentives irrespective of the organizational setting (Williamson, 1985). "As expert talent becomes more important to problem-solving, decision-making, and dispute resolution, new organizational forms are emerging to cater to the needs of both experts and clients" (Teece, 2003, p. 909). In general, professional services are highly customized, which implies that a significant proportion of the human assets (knowledge) involved in the transaction, including the service itself, are specific to the transaction (Williamson, 1985; Masten 1982). The service transaction is based on intensive and often long-standing interaction between the producer and the client, which implies that the service transaction, in the face of uncertainty, is subject to the costs of bounded rationality (Miles, 2003), strategic behaviour (Masten, 1986) and opportunism (Williamson, 1985). The problems are compounded by quality assessment, which is inherently subjective, and externalities[55].

The transactions of professional services are invariably associated with information asymmetry, mainly in favour of the supplier, which is the major source of verification problems, measurement cost and moral hazard (Miles, 2003; Løwendahl, 2005). For instance, the client may find it hard to establish the competence and experience of a specific supplier. Similarly, the client may not be able to assess accurately the kind or level of skills required to deal with the specific problems it faces, nor to match them to the offering of the supplier (Miles, 2003). As the service rendered is often specific and complex, it may be difficult to agree on the characteristics of the service and the criteria of assessing its quality (Miles, 2003; Løwendahl, 2005)[56]. Finally, the impact and effectiveness of professional services may be co-influenced by other known or unknown factors, making it difficult, if not impossible to isolate the productivity impact of the specific service on the client's performance. By and large, these issues are also addressed by the service management literature, but without explicit reference to transaction costs or implications for efficient governance modes.

Interestingly, what Løwendahl (2005) defines as the "value creating process" of professional services carries distinct similarities with the "fundamental transformation" discussed by Williamson (1985), which leads to the suppression of market mechanism and fosters the creation of bilateral dependency. Merging these two perspectives, there seem to exist three phases of transactions characteristic for professional services. Assuming competitive bidding at the initial stage (*ex ante phase*) the winning

[55] Løwendahl (2005) notes that for the professional service firm the problem is reversed, as opportunism may also occur on the client side. In corporate finance this challenge is identified as the "double moral hazard" problem.

[56] In terms of service quality, the higher the required specificity of the service is, the more difficult it will be for the service firm to guarantee the quality ex ante. "What is perceived to be high quality service depends very much on the expectations of the client, and the less clear the client is about what is expected of the supplier, the more difficult it is to deliver what is expected" (Løwendahl, 2005, p. 42).

supplier is able to sell a credible promise to the client. The more innovative and idiosyncratic the service is, and the larger the knowledge gap, the more complex the *ex ante* phase is likely to be. This implies high *ex ante* transaction costs associated with the service contract. The credibility of the promise depends on the seller's reputation signalled by the success in previous projects, and the assigned professionals (Løwendahl, 2005). Like the contract itself, the reputation of the supplier works as a credible commitment (Williamson, 1985), which safeguards the productive value of the required assets and the service delivery against opportunistic behaviour of the supplier in the *interaction phase*.

The interaction phase involves the set of activities required to deliver what has been promised and involves both the client firm and the professionals assigned. At this phase post-contractual opportunism and bounded rationality, if the safeguards fail, are the sources of *ex post* transaction costs. The supplier is concerned with the actual quality of what is delivered, the perceptions of quality by the client, and the efficiency of the delivery process. Hence, irrespective of the chosen strategy (domain[57]) of the service firm, it has to resolve *anew* the dilemma (trade-off) of how to apply standard procedures (efficiency) to unique problems and still deliver a customized and idiosyncratic solution that is appropriate for the client (Løwendahl, 2005; Viitamo, 2008). Note that if the efficiency of the delivery is over-emphasized, transaction costs of misalignment and opportunism should rise, and conversely, an extensive customization leads ultimately to inefficiencies in the delivery process.

The third (*ex post*) phase of the transaction is concerned with learning from the project and institutionalizing the learning to the extent that it can be utilized for both improved service quality and improved efficiency that the current and future clients are provided with[58]. The actual learning occurs at the interaction phase, but the point made by Williamson (1985) is that learning and accumulation of transaction-specific assets and knowledge for the supplier and the client is entailed in the fundamental transformation. That is, if the transaction and the associated bidding contest for *similar* services are recurrent, the original winner with the accumulated assets gets a competitive advantage over the (potential) competitors and is thus expected to take it all. "Faceless contracting is thereby supplanted by contracting in which the pair-wise identity of the parties matter" (Williamson, 1985, p. 62). Williamson goes on by noting that "both institutional and personal trust relations evolve…and individuals located at the interface may refuse to be part of the opportunistic efforts to take advantage of (rely on) the letter of the contract when the spirit of the exchange is emasculated" (ibid. p. 62).

4.3 Behavioural paradox

Despite the higher productivity of the transaction-specific assets and the organizational efficiency that the fundamental transformation should bring about, Williamson takes a more dismal stance. "Given the unenforceability of general clauses (of the contract) and the proclivity of human agents to make false and misleading statements, the following hazards must be confronted: joined as they are in a condition of bilateral monopoly, both buyer and seller are strategically situated to bargain over the disposition of any incremental gain whenever a proposal to adapt is made by the other party. Although both have a long-term interest in effecting adaptation of a joint profit-maximizing kind, each also has an interest in appropriating as much of the gain as he can on each occasion to adapt. Efficient adaptation that would otherwise be made thus results in costly haggling or even goes unmentioned, lest the gains be

[57] The domain or strategy of a service company here refers to a chosen combination of efficiency and effectiveness, which characterize the service production and delivery of the firm (Viitamo, 2008).
[58] "The current clients" added here to the original proposition by Løwendahl (2005) should not alter the main conclusion.

dissipated by costly sub-goal pursuit…governance structures that attenuate opportunism and otherwise infuse confidence are evidently needed" (Williamson, 1985, p. 63).

Given the high potential for a virtuous circle, which would also remedy bounded rationality, it is justified to ask whether such a vicious circle predicted by Williamson is inevitable and universal[59]. For Granovetter (1985) for instance, this is not the case[60]. More generally, with regard to the behavioural assumptions of transaction cost economics, bounded rationality is an implicit assumption made by the service management and productivity literature as well. In that context the key question is how to induce the actions of the supplier and the client within reciprocal interdependence to make the service more innovative and profitable for each. The general answer provided by the service marketing research is deepening contractual co-operation (see. e.g. Ravald and Grönroos, 1996). Through effective learning and reciprocal routines, collaboration makes the parties more rational, and the transaction more productive, respectively. Rooted in the spirit of empiricism, service management research makes few allowances for opportunistic behaviour as it appears in transaction cost analysis.

Taking into account of what was presented above, the behavioural assumptions and organizational implications of transaction cost economics do not match unquestionably with the empirical evidence of professional services. For instance, Miles (2003) notes that the full array of complexities of a business relationship associated with the governance of professional service transactions would be hard to capture within the formal framework of transaction cost analysis. The assessment of Miles is justified with some qualifications. Whereas transaction cost economics is methodologically far from formal, which has been a major reason for criticism against it, the behavioural assumptions it rests on, are so, and thus causes inconsistencies in the analysis. In particular, the analytical focus of service management research indicates that bounded rationality and opportunism are context-specific, and may influence adversely to what transaction cost economics *seems* to predict. "As expert talent becomes more important to problem-solving, decision-making, and dispute resolution, new organizational forms are emerging to cater to the needs of both experts and clients…Traditional hierarchical structures are likely to decline, to be replaced by more decentralized quasi self-organized organizations employing at-will contracts with performance measurement down to the individual level" (Teece, 2003, p. 909).

Bounded rationality, as it appears in the organizational theory and transaction cost economics, delimits the cognitive capacity of human actors to achieve the maximum productivity, which approximates the outcome of neoclassical perfect rationality. This is why organizations are needed to coordinate action and process information in the search for the second best solution. While bounded rationality has been the object of extensive assessments (Fransman, 1998; Radner, 2000) there are and will be insurmountable difficulties to operationalize the concept for robust theorizing. For instance, it remains unresolved how to distinguish between different sources of bounded rationality, i.e. collecting enough information, collecting the right information, and making right decisions based on the information. Owing to the informational limitations, transaction cost economics asserts that economic agents cannot write complete contracts. Yet, economic agents are assumed to be rational enough to assess the characteristics (e.g. asset specificity) of the transaction, and capable of electing the fittest mode of

[59] Actually, Williamson does not claim that all actors behave opportunistically. He merely suggests "that the probability of opportunism occurring increases as the asset specificity increases" (Hill, 1990, p. 502).

[60] The view of transaction cost economics of human nature is "under-socialized", portraying people as if they were social atoms unsentimentally pursuing their pecuniary interest with little regard for the social connections around them, which are mostly a source of friction (Granovetter, 1985).

governance. This suggests a major paradox, as the existing organizations represent the most productive forms only by chance. The question of how distant organizations are from the first best is displaced by the assumption that organizations tend to calibrate (evidently forever) towards it.

Equally problematic are the occurrence of opportunism and the assumed superiority of hierarchy as a universal solution to the vicious circle of opportunism. In particular it seems that opportunism and bounded rationality are interrelated in that "irrational" opportunism is mostly a reflection of bounded rationality. If profitable opportunities for a bilateral exchange exist for both parties for any length of time, the only plausible explanation for why these opportunities are degraded by opportunism is that the rationality of the transacting parties is somehow limited. It can be demonstrated by the game-theoretic setting of the prisoner's dilemma (Friedman 1986) that the seller and buyer tend to behave opportunistically in the short run, if they know that the relationship will not last indefinitely. Opportunistic strategy dominates even if the mutual benefit from co-operation (trust) for the length of the contract is higher than the benefit from acting opportunistically. In this case the rationale for opportunistic strategy stems from the impossibility to demonstrate trustworthiness and reputation[61] for the other party. Conversely, absence of trust results in costly safeguards to conduct the transaction, which is manifested in the presence of bounded rationality.

It is justified to assume, however, that in a majority of exchanges "both parties enter with the expectation that they may interact again in the future, although neither party can predict or know exactly how many times this will occur" (Hill, 1990, p. 505). So, if the prisoner's dilemma is planned to be played for an indefinite time, it can be shown that co-operation and trust is a viable strategy. The implication for transaction cost economics is that even if the fundamental transformation has created a bilateral monopoly based on specific assets, a simple market contract is the most efficient governance mode, if the time horizon is unlimited and the discounted net benefits from the co-operation exceed the one-time benefit from opportunistic behaviour. In the absence of exogenous assumptions on the probability of opportunism[62], reputation of trustworthiness can be demonstrated in subsequent periods, which will suffice to keep the relationship going. Again, any opportunistic act should reflect the presence of bounded rationality.

Accordingly, a more analytical inspection of bounded rationality and opportunism assumptions is needed to explain the persistence of long-term contractual relations in the professional service industry. An effective demonstration of faultless reputation is elementary to mitigate the information asymmetry between the supplier and the buyer, and thus a prerequisite for the survival of the supplier on the market. Reputation in turn creates trust, which is another indispensable asset of competitiveness of any professional service firm. Accordingly, "both the reputation of the firm and personal relationships with specific professionals are important factors for the client" (Løwendahl, 2005, p. 40). Opportunistic behaviour of the supplier is further suppressed by the *professional norm of conduct.* This includes setting the client needs higher than the profits of the firm and respecting the absolute standards of professional expertise[63].

---

[61] Specifically, the value created by exchanges involving actors of questionable reputation is significantly reduced by the need to set up safeguards to limit opportunism (Hill, 1990).

[62] Williamson makes this assumption more explicitly.

[63] Løwendahl (2005) notes that the primary characteristic of a professional service is "altruistic service to clients", meaning that "in cases of conflicts of interest between what is profitable to the supplier and what will be the best solution for the client, the latter alternative must be chosen" (op. cit. p. 22).

4.4 Trust versus opportunism

The analysis above is not meant to invalidate the basic premises of transaction cost economics. The purpose here is to outline the circumstances under which the occurrence of opportunistic behaviour is likely or unlikely, given the hold up -situation of idiosyncratic exchange. It also suggests that when reputation and trust are "sensitive" competitive assets exemplified by professional services, the linkage between the behavioural assumption and organizational ramifications deserve a closer examination. Hence, based on the game-theoretic reasoning that accounts for future business opportunities, opportunism is likely to occur when a) the outcome of productive action is uncertain, b) reputation is difficult or costly to establish, c) the payoff from opportunistic action in the present period seems to outweigh the discounted present value of future co-operation that is put in jeopardy by such action, and d) when opportunism is difficult to detect (Hill, 1990).

This leads to the conclusion that the principal source of transaction costs and the need for hierarchical governance is not opportunism triggered off by asset specificity as asserted by Williamson, but uncertainty and asymmetric information. Hence, any sort of opportunistic behaviour should be consequence of uncertainty, and if any profitable business opportunities or other future benefits are therefore lost, it will be an outcome of bounded rationality (Radner, 2000, p. 22)[64]. In the absence of uncertainty, Hill (1990) concludes, markets would be an efficient governance mode, and all safeguards needed to check against opportunism, including hierarchy, should reflect cost inefficiency. In the long run *invisible hand* of market selection (Smith, 1776) favours actors whose behavioural repertoires are biased towards co-operation and trust. For this interpretation trust manifests rational behaviour.

The second issue influencing the occurrence of opportunism is "the wider social context within which economic transactions are embedded" (Hill, 1990, p. 501). Most visibly the argument has been raised by Granovetter (1985) who maintains that the wider social context, social norms and culture have a profound impact on the behaviour of economic actors. In particular, social networks and interpersonal relationships between the parties of an exchange are a mechanism for attenuating opportunism in situations where transaction cost economics would predict that opportunism is likely and therefore hierarchy would be necessary (Hill, 1990). Hence, given the arguments on the importance of social context, it is contended here that opportunism will occur in a situations where, according to the propositions of transaction cost theory, it should not.

A further justification for these considerations follows from the biased view of transaction cost economics to regard markets and hierarchies as opposite forms of governance and assimilate the firm with the highest economic authority. "Rather, they [market and hierarchies] should be seen, conceptually, as mutually creating and mutually reinforcing aspects of any economic systems…Markets and hierarchies exist in a dynamic, creative tension with each other. To conceptualize them as opposites obscures their mutuality" (Hamilton et al., 1998, pp. 106-107). Different market and hierarchy configurations, such as industrial clusters and agglomerations (Porter, 1990) exist empirically and vary historically and geographically. The key point made by Hamilton et al. (1998) is that hierarchy is a relative concept and it includes several layers which extend across

[64] Bounded rationality as interpreted here includes inconsistency in decision procedures, ambiguity and vagueness in the cognitive frames by which agents attribute causal relations amongst observed variables, and related failures in mappings between environmental states, actions and consequences in ever-expanding state-spaces (Dosi et al 2005; Radner, 2000).

boundaries of single economic units. It should be noted that the argument is supportive of the open systems perspective of the organizational theory as well (Scott and Davis, 2003).

Drawing on the Weberian view on power and hierarchy, Hamilton et al. (1998) propose that hierarchy should be defined in terms of authority and economically effective action, which implies that boundaries of economic organizations are "determined by the reach of authoritative power and not arbitrarily by equated with the firm" (ibid., p. 116). This formulation has significant ramifications for productivity analysis and the coordination of professional service activities. Namely, "to the extent that a network of people or firms are linked by the exercise of binding norms, then that network functions as an economic organization…and to the extent that networks of people or firms are linked together only by their individual economic interest, then that network does not constitute an organization in its own right" (ibid., p. 116). Irrespective of the legal boundaries of the firm, from the hierarchy perspective, the network in the former case is an appropriate unit for productivity analysis, whereas in the latter case it is clearly not. Yet, the organizational design of professional services firms (legal, accounting, management consulting etc.) may be based on both types of networks.

Aside from the analysis of the externalities associated with the team production, the focal point in this approach is the participants' [professionals] subjective recognition that "they [professionals] are bound to the authoritative norms of the organization, that they are not formally free to act in other ways, and that there is a coercive means to enforce the normative rules" (Hamilton et al., 1998, p. 117)[65]. Normative influences of this sort are of course a matter of degree. Yet, if hierarchy is assumed to suppress opportunism as predicted by transaction cost economics, collaboration and trust may guide the inter-firm relations irrespective of the degree of asset specificity. To conclude, inter-firm networks that rest on strongly normative social bonds are better understood as economic organizations in their own right instead of a residual or intermediate category, as suggested by transaction cost economics (Hamilton et al., 1998).

Given these considerations and the inter-organizational affiliations of individuals, it is easy to see why opportunism of highly educated individuals is a persistent managerial challenge and difficult to suppress in a professional service firm. As the main parts of the "effective network" of employed professionals are located outside the firm, and consist of colleagues in competing companies and client companies, the conflict of interests with the employer is barely unavoidable. Accordingly, in contrast with the predictions of transaction cost economics, a professional service firm is not necessarily effective in transforming high-powered incentives of professionals into low-powered incentives, i.e. to harmonise interests and lower the risk of opportunism. With the degree of specificity and the strategic importance of human assets, these ramifications are contingent upon technological characteristics - such as the role of team work – required for service production and provision.

Suffice it to say, "very little can be owned by the professional service enterprise" (Teece, 2003, p. 900). Much of the firm's assets reside with key individuals, and are quite transportable beyond the boundaries of the firm as "capital goes down the elevator every night" (ibid. p. 900). Therefore professionals tend to protect their core capability, which implies that the reputation they are concerned with is their core capability such as personal relations with the customers. Given the embedded risk of opportunism within the professional firm, "a trigger of a conflict can lead to mass defections and

[65] The authors note that participants often recognize that they are bound by organizational authority and act in a self-interested way in violation of the organizational norms.

destruction of enterprise value, even more assuredly than in an industrial enterprise" (Teece, 2003, p. 900). Hence, design of treaties including e.g. partnership and incentive schemes that reduce the risk of conflict induce co-operation and reward performance simultaneously is of a greatest value to the professional firm.

To sum up, relaxing the axiomatic and biased treatment of human behaviour does not invalidate the basic propositions of transaction cost economics. The inclusion of bounded rationality and opportunism more analytically into the explanatory framework is instead contributory to the refined theory on organizational productivity. With the prevalence of knowledge and human capital for the competitiveness of a modern business enterprise, innovative strategies on organizational adaptation are increasingly important. In reference to Teece (2003) "traditional hierarchical structures [of a firm] are likely to decline, to be replaced by more decentralized quasi self-organized organizations employing at-will contracts with performance measurement down to the individual level" (ibid. p. 910). Hierarchies cannot be abandoned, but contingencies associated with knowledge-based services are most economically contained by adhocracy and networks (Scott and Davis, 2003).

5 Conclusion

Organizations are ubiquitous. They are dominant characteristics of modern societies and actors of their own right. Organizations are social structures created by individuals to support collaborative pursuit of specific goals. Organizations must define and refine their objectives, and they must induce the participants to deliver services, which have to be controlled. As with business enterprises, resources have to be garnered from the environment and the products and services dispensed. Characteristically, a substantial proportion of the capacity of the resources has to be expended to maintain the organization itself, in which case the means become the ends (Scott and Davis, 2003).

Distinctively, the general theory on organizations parallels with the theory of strategic and corporate management. The *ends* or the objective in the case of a business firm are production and profitability, whereas the *means* are reflected by the organizational and administrative structure that enables profitable production of products and services. As with production, the organization entails costs, which any rationally behaving manager and entrepreneur pursue to minimize. From the strategic management perspective of the firm, the production and organizational costs are not strictly distinguished, as the firm's activities together constitute its production technology, and the *production function*.

In the microeconomic analyses of industrial organization in contrast, the distinction between organizational costs and production costs is made. Whereas the neoclassical theory of the firm deals exclusively with the cost of production and the implications for the firm's optimal choices, transaction cost economics focuses principally on the costs or governance, where the choice with a given technology has to be made between the available (the most economizing) modes of organizing the conduct of business. While each approach concentrates on one of the two focal dimensions of a firm's productivity, their conclusions are therefore unavoidably subject to some degree of arbitrariness and biases caused by the simplified reality.

To facilitate a more comprehensive analysis of a firm's productivity from the perspective of growth and integration, a viable option is a combined approach. Whereas the neoclassical theory is capable of identifying the conditions under which technological integration of the adjacent stages production is feasible, i.e. conducive to higher economic efficiency or market power, the characteristics of technology and markets are not sufficient rationales for a unified (integrated) ownership. Consequently, the degree and mode of unified control should be determined by the transaction costs associated with the alternative modes of governance (Viitamo, 1996; Teece, 1986)[66]. This way the insistence of effective governance should bring added realism to the abstract neoclassical analysis.

It is necessary to take a more objective view on how the shift in emphasis from the neoclassical to the organizational approach actually affects the balance of *realism*. The allowance of uncertainty, imperfect rationality and moral hazard within the organizational framework is undoubtedly conducive to a more accurate depiction of the real world. Such an increment in realism is attained at the expense of higher obscurity of the economic objectives, and consequent imprecision of the analysis. More

[66] Note that a simplistic "division of labour" in the overall optimization may encounter problems as well. It assumes that technological design of production can be made independently of the organizational context. If the organization is an essential input to the production process, and in part determines the characteristics of the service and product offered, then a separate analysis leads to a biased sub-optimization.

generally, significant imperfection and strive for a more efficient organization may be realistic assumptions at the systemic level, but they fail to account for the variations in calculating behaviour at the individual level. Hence, realism increases in some sense and diminishes in another.

For the sake of realism, it is essential to stress that organizational choices of firms may - and more often than not they do - reflect the co-influence of economizing on organizational costs and strategizing in the pursuit of higher market power and barriers to competitive entry. The isolation of these effects, as well as the distinction between production and transaction costs is difficult, if not impossible to make in practise. Given these qualifications, transaction cost economics proves to be a viable approach for the analysis of a firm's productive performance and service productivity, in particular. Most services are characterized by temporary bilateral exchange, where the client participates actively in the production and delivery of the service. The service transaction itself is often a multi-stage process embedded in the co-production by the supplier and the client.

From a broader perspective, the agenda of transaction cost economics is connected to service productivity in two complementary ways. Regarded as a costly input to any kind of processes, organizational pattern provides productive *services* (Penrose, 1959) in the production function of a firm. Conceptually, organizational productivity can be understood as the outcome of the organizational services relative to the associated costs of organization. Moreover, any organization encompasses complementary activities and assets, which provide services internally to the other activities of the organization, and externally to the clients the organization is supposed to serve. How these internal activities transact mutually and with the external activities of other organizations, determines the overall productivity of individual activities and assets.

The first of the above mentioned interpretations is applicable at the strategic level of corporate structuring and governance, while the second is reflected in the make-or-buy issue of specific transactions and productive assets. More specifically, the issue of make-or-buy is related to the feasibility of incremental growth through vertical, horizontal, and diversified integration. In transaction cost economics, integration is defined in terms of unified control which is a matter of degree indicated by the continuum of governance between markets and hierarchies. On balance, both agendas of corporate structuring and make-or-buy follow the principle of contingency developed by the general organization theory. The organization is a choice variable, which at the optimum should *match* with the strategy, as well as with technological and environmental contingencies.

Of the three dimensions of a transaction, asset-specificity, frequency and uncertainty, the degree of asset-specificity is the key determinant of vertical integration. Yet, transaction cost economics fails to explain convincibly, why to invest in specific assets in the first place. It is contended here that asset-specificity reflects the type of productivity the owner of the asset is pursuing. Non-specific assets enable the utilization of scale and scope conducive to efficiency, whereas transaction-specific assets are capable of providing higher effectiveness and customized quality. This suggests the existence of a trade-off between efficiency and effectiveness also emphasized by service management literature. The existence of quality-induced externality is another rationale for integration, and thereby a major concern for effective service management.

In the case of diversification and horizontal integration, the issue of make-or-buy in vertical integration is reversed to the problem of make-or-sell. In effect, the internalized asset of a firm exhibits excess capacity in the services it can provide, in which case these services find alternative uses outside, but

related to the current production of the firm. If the markets for the excess services are thin or otherwise exposed to a moral hazard and externalities within the bilateral exchange, diversification is a feasible growth strategy. A prominent example is the difficult-to-trade knowledge, which explains the prevalence of multi-product firms in various professional service industries (Teece, 1982). Combining the rationales of vertical and diversified integration, the framework of *profiting from innovation* (Teece, 1986a) provides a generalized theory on the markets and hierarchies in the utilization of innovation through inter- and intra-firm transactions of specific assets and the associated services.

Corporate governance and restructuring shift the focus from the make-or-buy problem in the context of individual transactions, onto the comprehensive issue of organizing internally, i.e. how-to-make. Here the organizational services are provided by specific "patterns" of governance, the choice of which is strategic, comparable with the choice of other means with respect to the given ends of the corporation. In general, transaction cost economics indicates an unqualified preference of the multidivisional form over the functional form of governance. This stems from the axiomatic belief that market governance should be relied upon as much as possible to attain the efficiency on production, while hierarchy should be used only to attain aligned incentives and productive use of corporate assets. The organizational theory, instead, addresses the issue more objectively and favours the logic of contingency in the pursuit of an organizational "fit". In contrast with the U-form, where productivity draws on *scale-driven efficiency*, the M-form builds more on *effectiveness and reciprocal interdependence*[67].

Given the versatility and suitability of the transaction cost approach to the analytical purpose here, the lack of applied research on services and service productivity is particularly striking. The negligence is in this respect general, as the field of industrial organizations has not showed any specific interest in services. Dominated by the neoclassical doctrine, the industrial organization is aligned with the *assimilation* perspective (Viitamo, 2007), which posits that there are no accountable differences between manufacturing and services with regard to innovations and productivity. Another plausible explanation stems from the available evidence on service transactions, which do not give an unqualified support for the basic propositions of transaction cost economics. If the rigid assumptions on bounded rationality and opportunism are relaxed and made more context-specific, new insights on service productivity, as well as realistic refinements in organizational analysis can be made.

[67] Reciprocal interdependence prevails between the internal activities of the corporation, and between the boundary-spanning activities of the corporation and the environment.

References

Arrow, K.J. (1971), Essays in the Theory of Risk bearing, Chicago, IL.

Alchian, A.A. and H. Demsetz (1972), Production, Information Costs and Economic Organization, American Economic Review, 62, pp. 777- 795.

Van Ark, B. and G. Jong (2004), Productiviteit in Dienstverlening, Deel 1: Wat Het Is En Waarom Het Moel? Studie Uitgevoerd in Opdracht van Stichting Management Studies (SMS), Haag.

Ashby, W.R. (1968), Principles of Self-Organizing System, pp. 108-118, in Buckley, W. (eds) (1968), Modern Systems Research for the Behavioural Scientist, Chicago: Aldine.

Barzell, Y. (1982), Measurement Cost and the Organizations of Markets, Journal of Law and Economics, 25, 4, pp. 27-48.

Baumol, W.J. (1982), Contestable Markets: an Uprising in the Theory of Industry Structure, American Economic Review, 72, pp. 1-15.

Chandler, A.D. (1990), Scale and Scope, The Dynamics of Industrial Capitalism, The Belknap Press of Harvard University Press, Cambridge, Massachusetts, London, England.

Chandler, A.D. (1977), The Visible Hand; The Managerial Revolution in American Business, Cambridge, Massachusetts: Belknap Press.

Chandler, A.D. (1962), Strategy and Structure: Chapters in the History of American Industrial Enterprise, MIT Press, subsequently published in (1966), New York: Doubleday and Co.

Coase, R. (1937), The Nature of the Firm, Economica, 4, pp. 386-405.

Commons, J.R. (1934), Institutional Economics, Madison: University of Wisconsin Press.

Demsetz, H. (1993), The Theory of the Firm Revisited, in Williamson O.E. and S.G. Winter (eds) (1993), The Nature of the Firm, Origins, Evolution and Development, New York, Oxford, Oxford University Press.

Dosi, G., D.J. Teece and J. Chytry, (eds) (1998), Technology, Organization and Competitiveness, Perspectives on Industrial and Corporate Change, Oxford University Press.

Donaldson, L. (1995), American Anti-Management Theories of Organization: A Critique of Paradigm Proliferation, Cambridge: Cambridge University Press.

Friedman, J.W. (1986), Game Theory with Applications to Economics, Oxford, England: Oxford University Press.

Fransman, M.(1998), Information, Knowledge, Vision, and Theories of the Firm, pp. 147-192, in Dosi, G., D.J. Teece and J. Chytry (eds) (1998), Technology, Organization and Competitiveness: Perspectives on Industrial and Corporate Change, Oxford University Press.

Gadrey, J. and F. Gallouj (eds) (2002), Productivity, Innovation and Knowledge in Services. New Economic and Socio-Economic Approaches, Edward Elgar, Cheltenham, UK

Guillén, M.F. (1994), Models of Management: Work, Authority and Organization in Comparative Perspective, Chicago: University of Chicago Press.

Granovetter, M. (1998), Coase Revisited: Business Groups in the Modern Economy, in Dosi, G., D.J. Teece and J. Chytry, (eds) (1998), Technology, Organization and Competitiveness, Perspectives on Industrial and Corporate Change, Oxford University Press.

Granovetter, M. (1985), Economic Action and Social Structure: The Problem of Embeddedness, American Journal of Sociology, 91, pp. 481-510.

Grant, R.M. (1991), The Resource-Based Theory of Competitive Advantage: Implication for Strategy Formulation, California Management Review, Spring 1991, 33, 3, pp. 114-135.

Grossman, S.J and O.D. Hart (1986), The Cost and Benefits of Ownership: A Theory of Vertical and Lateral Integration, Journal of Political Economy, 94, 4, pp. 691-719.

Hamilton, G.G., and R.C. Feenstra (1998), Varieties of Hierarchies and Markets: An Introduction, pp. 105-146), in Dosi, G., D.J. Teece and J. Chytry (eds) (1998), Technology, Organization and Competitiveness: Perspectives on Industrial and Corporate Change, Oxford University Press.

Grönroos, C. and K. Ojasalo (2004), Service Productivity: Towards a Conceptualization of the Transformation of Inputs into Economic Results in Services, Journal of Business Research, 57, pp. 414– 423.

Hill, C.W.L. (1990), Cooperation, Opportunism, and the Invisible Hand: Implications for Transaction Cost Theory, Academy of Management Review, 15, 3, pp. 500-513.

Hoskisson, R.E., C.W.L Hill and H. Kim (1993), The Multidivisional Structure: Organizational Fossil or Source of Value? Journal of Management, 19, 2, pp. 269-198.

Jensen, M.C. and W.H. Meckling (1976), Theory of the Firm: Managerial Behaviour, Agency Costs, and Ownership Structure, Journal of Financial Economics, 3, pp. 305-360.

Jones, G.R and C.W.L. Hill (1988), Transaction Cost Analysis of Strategy-Structure Choice, Strategic Management Journal, 9, pp. 159-172.

Kay, N. (2000), Searching for the Firm: The Role of Decision in the Economics of Organization, Industrial and Corporate Change, Oxford, Dec 2000, 9, 4, pp. 683-703.

Klein, B., R.A. Crawford and A.A. Alchian (1978), Vertical Integration, Appropriable Quasi-Rents, and Competitive Contracting Process, Journal of Law and Economics, 21, pp. 297-326.

Kreps, D.M. (1990), A Course in Microeconomic Theory, Harvester Wheatsheaf, New York, London, Toronto, Sydney, Tokyo.

Lawrence, P.R. and J.W. Lorsch (1967), Organization and Environment: Managing Differentiation and Integration, Boston: Graduate School of Business Administration, Harvard University.

Løwendahl, B.R. (2005), Strategic Management of Professional Service Firms, Copenhagen Business School Press, 3rd Edition.

Macneil, I.R. (1978), Contracts: Adjustments of Long-Term Economic Relations under Classical, Neoclassical, and Relational Contract Law, Northwestern University Law Review, 72, pp. 854-906.

March, J.G and Simon H.A. (1958), Organizations, New York: John Wiley.

Masten, S.E. (1986), Institutional Choice and Organization of Production: The Make-or-Buy Decision, Journal of Institutional and Theoretical Economics, 142, pp. 493-509.

Masten, S.E. (1982), Transaction Costs, Institutional Choice and the Theory of the Firm, Ph.D. Dissertation, University of Pennsylvania.

Metcalfe, S. and I. Miles (2006), Service Productivity in Europe, Part 1: Investigation and Comparison of Methods for Productivity Measurement, July 2006, Price Waterhouse Coopers.

Miles, I. (2003), Knowledge intensive services' suppliers and clients, The Ministry of trade and industry Finland, Research reports 15/ 2003.

Nelson, R.R. (2006), Reflections of David Teece's "Profiting from Technological Innovation…", Research Policy, 35, pp. 1107-1109.

Nelson, R.R and S.G. Winter, (1982), An Evolutionary Theory of Economic Change, The Belknap Press of Harvard University Press, Cambridge, Massachusetts, and London, England.

Panzar, J.C. and R.D. Willig (1981), Economies of Scope, American Economic Review, 71, 2, pp. 268-272.

Penrose, E. (1959), The Theory and Growth of the Firm, Oxford: Oxford University Press, 3rd Edition.

Pisano (2006), Profiting from Innovation and the Intellectual Property Revolution, Research Policy, 35, pp. 1122-1130.

Porter, M.E. (1998), On Competition, A Harvard Business Review Book.

Porter, M.E. (1985), Competitive Advantage: Creating and Sustaining Superior Performance, New York: Free Press.

Porter, M.E. (1980), Competitive Strategy: Techniques for Analyzing Industries and Competitors, New York: Free Press.

Radner, R. (2000), Costly and Bounded Rationality in Individual and Team Decision-making, pp. 3-35, in Dosi, G., D.J. Teece and J. Chytry (eds) (2005), Understanding Industrial and Corporate Change, Oxford University Press.

Ravald, A. and C. Grönroos (1996), The Value Concept and Relationship Marketing, European Journal of Marketing, 30(2), pp. 19-30.

Riordan, and O.E. Williamson (1985), Asset Specificity and Economic Organization, International Journal of Industrial Organization, 3, pp. 365-378.

Rumelt, P.R., D. Schendel and D.J. Teece (1991), Strategic Management and Economics, Strategic Management Journal, 12, pp. 5-29.

Salter, A. and B.S. Tether (2006), Innovation in Services: Through the Looking Glass of Innovation Studies, Background paper for Advanced Institute of Management (AIM) Research's Grand Challenge on Service Science.

Scherer, F.M. (1980), Industrial Market Structure and Economic Performance, Rand McNally, Boston, MA, 1970, Second Edition 1980.

Schumpeter, J.A. (1947), Capitalism, Socialism and Democracy, Second Edition, New York: Harper & Row.

Scott, W.R. and G.F. Davis, (2003), Organizations and Organizing, Rational, Natural, and Open System Perspectives, Pearson, Prentice Hall, Upper Saddle River, New Jersey 07458.

Smith, A. (1776), An Inquiry into the Nature and Causes of the Wealth of Nations, London: Strahan and Caddel, vol. 1 and 2.

Simon, H.A. (1961), Administrative Behaviour, 2nd Edition, Macmillan: New York.

Simon, H.A. (1957), Models of Man, John Wiley & Sons: New York.

Teece, D.J. (2006), Reflections on "Profiting from Innovation", Research Policy, 35, pp. 1131-1146.

Teece, D.J. (2003), Expert Talent and the Design of (Professional Services) Firms, Industrial and Corporate Change, Oxford, Aug 2003, 12, 4, pp. 895-916.

Teece, D.J. and G. Pisano (1998), The Dynamic Capabilities of Firms: Introduction, in Dosi, G., D.J. Teece and J. Chytry (eds) (1998), Technology, Organization and Competitiveness: Perspectives on Industrial and Corporate Change, Oxford University Press.

Teece, D.J., G. Pisano and A. Shuen, (1997), Dynamic Capabilities and Strategic Management, Strategic Management Journal, 18, 7, pp. 509-533.

Teece, D.J. (1986a), Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy, Research Policy, 15, pp. 285-305.

Teece, D.J. (1986b), Transaction Cost Economics and the Multinational Enterprise, An Assessment, Journal of Economic Behaviour and Organization, 7, pp. 21-45.

Teece, D.J. (1984), Economic Analysis and Strategic Management, California Management Review, XXVI, 3, Spring 1984, pp. 87-110.

Teece, D.J. (1982), Towards an Economic Theory of the Multiproduct Firm, Journal of Economic Behaviour and Organization 3, pp. 39-63.

Teece, D.J. (1980), Economies of Scope and the Scope of the Enterprise, Journal of Economic behaviour and Organization, 1, pp. 223-247.

Thompson, J.D. (1967), Organizations in Action, Social Science Bases of Administrative Theory, Transaction Publisher, New Brunswick USA, and London UK.

Vargo, S.L and R.F. Lusch (2004), The Four Service marketing Myths, Remnants of a Goods-based, Manufacturing Model, Journal of Service Research, 6, 4, pp. 324-335.

Varian, H.R. (1984), Microeconomic Analysis, Second Edition, W. W. Norton & Company, New York, London.

Viitamo, E. (2008), On Service Productivity – Strategic Management Perspectives, Lappeenranta University of Technology, Faculty of Technology and Management, Department of Industrial Management, Research Reports, 205.

Viitamo, E. (2007), Productivity of Business Services – Towards a New Taxonomy, Lappeenranta University of Technology, Faculty of Technology and Management, Department of Industrial Management, Research Reports, 188.

Viitamo, E. (2004), Knowledge-Intensive Services and Competitiveness of the Forest Sector – Case Finland, IIASA, Interim Report IR-03-057.

Viitamo, E. (1996), Clustering from the Perspective of Transaction Cost Economics – The Case of the Finnish Paper Industry (In Finnish), Unpublished Licentiate Thesis, Helsinki University, Department of Economics, Faculty of Social Sciences.

Williamson, O.E. (2003), Examining Economic Organization through the Lens of Contract, Industrial and Corporate Change, 12, 4, pp. 917-942.

Williamson, O.E. (1998), Transaction Cost Economics and Organization Theory, pp. 17-66, in Dosi, G., D.J. Teece and J. Chytry (eds) (1998), Technology, Organization and Competitiveness: Perspectives on Industrial and Corporate Change, Oxford University Press.

Williamson, O.E. (1993), The Logic of Economic Organization, pp. 90-116, in Williamson O.E. and S.G. Winter (eds) (1993), The Nature of the Firm, Origins, Evolution and Development, New York, Oxford, Oxford University Press.

Williamson, O.E. (1991), Strategizing, Economizing and Economic Organization, Strategic Management Journal, 12, pp. 75-94.

Williamson, O.E. (1989), Transaction Cost Economics, Chapter 3, in Schmalensee, R. and R.D. Willig (eds) (1989), Hanbook of Industrial Organization, Volume 1, Elsevier Science Publishers B.V.

Williamson, O.E. (1985), The Economic Institutions of Capitalism: Firms, Markets and Relational Contracting, New Your: Free Press.

Williamson, O.E. (1981), The Modern Corporation: Origins, Evolution, Attributes, Journal of Economic Literature, 19, pp. 1537-1568.

Williamson, O.E. (1975), Markets and Hierarchies: Analysis and Antitrust Implications, New York.

Winter, S.G. (2006), The Logic of Appropriability: From Schumpeter to Arrow to Teece, Research Policy, 35, pp. 1100-1106.

Winter, S.G. (1993), On Coase, Competence, and the Corporation, in Williamson O.E. and S.G. Winter (eds) (1993), The Nature of the Firm, Origins, Evolution and Development, New York, Oxford, Oxford University Press.

Lappeenrannan teknillinen yliopisto
Digipaino 2008
ISBN 978-952-214-710-3 (Paperback)
ISBN 978-952-214-711-0 (PDF)
ISSN 1459-3173