

Williamson's Theory of the Firm: Efficiency vs. Power

Giorgos Meramveliotakis

Department of Management Science & Technology
Hellenic Mediterranean University
Crete, Greece

Abstract

The purpose of this paper is to appraise the analytical usefulness of Williamson's theory of the firm by investigating the explanatory capacity of transaction cost concept. Particularly, the aim is to criticize Williamson's main tenet that the emergence of the hierarchical form of capitalist production system can be analysed in terms of efficiency considerations. Contrary to Williamson's rationale, I conclude that the concept of power seems, in contrast to Williamson's adherence to the efficiency principle, to be a more effective analytical tool in explaining the establishment of the factory system.

Keywords: Firms, Power, Transaction Costs, Efficiency.

JEL Classification Codes: B52, D02, P14

1. Introduction

In November 2009, Oliver Williamson was awarded the Nobel prize in economics.¹ This follows the award to Ronald Coase in 1991 and to Douglass North in 1993. Between them, Coase, Williamson and North, are the founders and most important representatives of new institutional economics.

In what follows, I am going to analyse the validity of the “market transaction cost” approach. In particular, I am going to examine if the transaction cost rationale, given the existence of an institutional environment (e.g. property rights, a functioning market etc.), can explain the origins of organisations, most notably the firm. The benchmark point is the analysis of Williamson (1975, 1985). The main premises on which his theory rests is that transaction cost economising is the most important explanatory variable underlying the emergence and evolution of hierarchies (e.g. the firm). The pervasive logic is that high transaction costs associated with market functioning mean that rational actors will do their best to avoid these costs by organising the production under a hierarchical structure involving lower transaction costs, hence the emergence of the firm. Thus, granted the transaction cost minimisation principle, the very existence of the firm implies that the latter is a more efficient contractual form than the market arrangement. Specifically, Coase (1937) was the first among new institutionalists who argues that the main function of the firm is to minimise transaction costs by replacing market transactions with the hierarchical organisation of work (Meramveliotakis & Milonakis, 2010). Williamson further develops Coase's argument by describing the scope of conditions under which a transaction would be more likely to take place within a firm than in the market. He also introduces the concept of transaction costs in dealing with the internal management of enterprises, focusing on the study of the internal organisational structure of the firm.

Initially, I will sketch the general logic of the new institutionalist theory of the firm, as represented by the classic contribution of Williamson. Then, a critical assessment of this contribution will be offered. It should be stressed from the outset that my critique is focused on the question of whether transaction costs can provide a reasonable rationale for explaining the emergence of the firm, and on this basis I will assess the theory of the firm itself within this framework. This is the reason why I am not going to analyse the other classical works within the new institutional economics' tradition, such as Alchian and Demsetz (1972) or to bring into picture any alternative theories of the firm such as the one by Penrose (1959).²

What Williamson (1985) is trying to explain is the emergence and evolution of the *capitalist* firm and not any form of historical organisation devoted to the production of goods or services. As is already made explicit in the title of his book, his theory is focused on the explanation of the institutions of capitalism. Accordingly, the firm is viewed as a hierarchical organisation governed by the employment relation or else by the capitalist-worker relation.

¹ Alongside Elinor Ostrom, a political scientist.

² For a critical analysis of Coase's theory of the firm see Meramveliotakis & Milonakis (2010).

Hence, my critique concentrates on the question of whether the transaction costs concept can offer a valid explanation of the origins of the hierarchical form of capitalist production. Particularly, here the focus is on the power relations and my argument is that the concept of collective power seems, in contrast to Williamson's adherence to the efficiency principle, to be a more effective analytical tool in explaining the establishment of the factory (i.e. the hierarchical form of capitalist production) system.

2. Williamson's Formulation of the Emergence of the Firm.

Coase (1993, p. 61) notices that his original formulation of the theory of the firm did not catch the economists' attention for a long period.³ Coming more than thirty years later, Williamson's (1975, 1985) works constitute the most comprehensive attempt to build upon Coase's theory of the firm. Williamson's work played a crucial role in establishing the "transaction cost economics" (TCE) line of investigation, where the emphasis is on the comparative costs incurred in "governing transactions". He also introduces the concept of transaction costs in dealing with the internal management of enterprises, focusing on the study of the internal organisational structure of the firm.

Starting, for "expositional convenience", from the theoretical assumption that "in the beginning there were markets" (1975, p. 20), Williamson creates a scheme similar to Coase's "market vs hierarchy" type of analysis. Once more, markets and firms are seen as alternative modes of organising the very same transactions and whether a transaction occurs under an authority rule or through a market exchange is a matter of efficiency considerations: "the general approach to economic organisation employed here can be summarised compactly as follows: (1) Markets and firms are alternative instruments for completing a related set of transactions; (2) whether a set of transactions ought to be executed across markets or within a firm depends on the relative efficiency of each mode" (1975, p. 8). The unit of analysis is the notion of transaction. Following Commons' (1934) lead, Williamson (1985, p. 3) states that "the transaction is properly regarded as the basic unit of analysis".⁴ Economic relations are, therefore, viewed in contractarian terms, as organised through the establishment of contracts. Consequently, for Williamson (1985, p. 20), the study of economic organisations of capitalism is conducted in terms of a contract theory, i.e. individuals establish "governance structures", which represent institutional modes of contacting to protect themselves from various hazards associated with exchange.

Williamson's theory of the firm is based on two behavioural assumptions. The first draws on Simon's (1961 [1947], p. xxiv) notion of "bounded rationality", according to which human behaviour is "*intendedly* rational, but only *limitedly* so", because of the cognitive limits of the human mind to perform all the necessary calculations even if the relevant information is available. The second is "opportunism", which is extended beyond the hypothesis of pure selfishness, indicating "self-interested seeking with guile" (Williamson, 1975, p. 26), i.e. agents try to satisfy their interest by using blatant forms of behaviour, such as lying, stealing and cheating (Williamson, 1985, p. 47).⁵ In close relation with these two behavioural assumptions is Williamson's additional concept of asset specificity, which refers to a human or physical asset that is specialised to a single use or to a single transaction (p. 55). Hence "... asset specificity only takes on importance in conjunction with bounded rationality/opportunism and in the presence of uncertainty ... Absent this condition [asset specificity], the world of contracting is vastly simplified; enter asset specificity, and non-standard contracting parties quickly disappear" (p. 55).⁶

³ Coase (1993, p. 62) notes that the number of references to "The Nature of The Firm" is significantly larger after 1966. He regards the popularity of his other celebrated article "The Problem of Social Cost", which has originally published in 1960, as the main reason for the switch of the academic interest to his 1937 article.

⁴ Hodgson (1999, pp. 218-219), however, stresses correctly that Williamson's use of transaction, as the basic unit of analysis, lies at the very opposite of Commons' view. While Commons adopts a structural analysis, Williamson treats the notion of transaction in a "given, atomistic and individualistic concept of thought".

⁵ It is, therefore, the simultaneous existence of bounded rationality and opportunism that causes the costs of transacting and make each contract unavoidably incomplete (Williamson, 1985, pp. 30-32). Otherwise, complete contracts are feasible if it is supposed that one of the two behavioral attitudes is vanished: if we assume that individuals are perfectly rational then the assumption of opportunism is empty of meaning, since agents could ex ante foresee and deal with any potential opportunistic behaviour. On the other hand, if we treat human beings as non-opportunistic, there would be no reason for deviation from the initial terms of the contract. But with the simultaneous existence of these attitudes the aim should be to "*organize transactions so as to economize on bounded rationality while simultaneously safeguarding them against the hazards of opportunism*" (1985, p. 32).

⁶ Williamson (1985, pp. 52-61), also, discerns two additional dimensions that characterise each transaction: uncertainty and frequency. He argues, however, that asset specificity is the most important.

Given these conceptions, Williamson proposes the general thrust of his analysis: “the economic institutions of capitalism have the main purpose and effect of economizing on transaction costs” (1985, p. 17). In his view, efficiency and cost considerations are central to the emergence of the firm. Transaction costs create market failures, which, in turn, lead to an inefficient allocation of resources. Firms, then, are devices that internalise these failures, leading the whole economic system into an efficient allocation of resources.⁷ Williamson further indicates that organisational arrangements could be described along a spectrum, with “classical market” and “centralized hierarchical organization” at the two poles. Between the two poles there exist a variety of “mixed modes of firm and market organization”. The movement from market to hierarchy “mainly involves a comparative institutional assessment of discrete institutional alternatives”, a tradeoff between alternative modes for organising transactions (1985, p. 42).

But what prevents the firm from growing to the extent that it replaces all exchange and market relations? Williamson argues that the internalisation of production is not without costs and these costs are associated with a reduction of the incentives that the market provides through the competition principle: “... the transfer of a transaction out of the market into the firms is regularly attended by an impairment of incentives ... The market is a marvel ... because of its remarkable capacity to present and preserve high-powered incentives ... As compared with internal transactions, market mediated transactions rely more on high-powered incentives and less on the administrative process (including auditing) to accomplish the same result” (1985, pp. 161-162). In this way, Williamson determines the boundaries of the firm at the point where the merits from the minimisation of transaction costs are offset from the lack of market incentives. Thus, “... firms and markets coexist ... because in a context of costly transactions, neither can solve the whole allocative problem efficiently” (Ankarloo and Palermo, 2004, p. 416).

3. Efficiency vs. Power

Having shown the explanatory inadequacy of the transaction cost rationale for the emergence of the firm elsewhere (Meramveliotakis & Milonakis, 2010; Meramveliotakis 2018), I now turn to criticise a decisive part of Williamson’s analysis, that the evolution of the hierarchical organisational form is justified on the grounds of efficiency, i.e. the minimisation of transaction costs. In fact, this is the logic that dominates Williamson’s historical explanation of the development of the factory system. In his article on “The Organization of Work”, Williamson (1980) attempts to demonstrate that the evolution of the firm from the putting out system to the modern factory has taken place primarily on efficiency grounds. His core argument is that the factory system emerges as a result of an attempt by the individuals involved to minimise transaction costs. It should be noted that the transition from the putting-out to the factory system involves a transition from one hierarchical form to another and not a transition from the market to a hierarchical organisation of work, since the putting-out system was itself a hierarchical form (Pitelis, 1991, p. 28). As such, Williamson’s analysis refers to the way the hierarchical form evolves over time and not to the origins of the hierarchy itself.⁸

The use of the transaction costs rationale ostracises from the analysis any thought of power considerations involved in the setting-up of capitalist firm relations. Indeed, for Williamson, the legitimacy of the authority relation lies in a process of voluntary contracting in the employment relationship. The resulting authority relation is assumed to be in the interests of both workers and employers because organisational performance is enhanced through the minimisation of transaction costs. Williamson maintains that hierarchies grow and change in response to the dynamics of transactions, not in response to the dynamics of power: “‘power relations’ result when the organization sacrifices efficiency to serve special interests. We concede that this occurs. But we do not believe that major organisational changes in the commercial sector are explained in these terms” (Williamson and Ouchi, 1981, p. 364).⁹

⁷ Recently, Williamson (1990, p. 67) has argued that economising on the sum of production costs and transaction costs is a more general statement of his hypothesis. However, as Furubotn and Richter (1991, p. 8) observe, “this distinction between transaction costs and production costs can be particularly hard to make”. According to Dietrich (1994, p. 28) it is very doubtful if a theory can simultaneously adopt these two different concepts, since transaction costs are relying on the persistence of bounded rationality, a term totally ignored in the neoclassical production cost theory.

⁸ Landes (1966, p. 12), quoted in Pitelis (1991, p. 29), observes that although most domestic weavers owned their looms and neilers their forges, “they were not, however, independent entrepreneurs selling their products in the open market; rather they were hirelings, *generally tied to a particular employer*, to whom they agreed to furnish a given amount of work at a price stipulated in advance”.

⁹ In a later work, Williamson (1985, p. 238) stresses that power is “so poorly defined that power can be and is invoked to explain virtually anything”, thus reversing the critique originally targeted against the transaction cost concept.

Following this, Williamson's (1980, pp. 11-12) method is to compare the transaction costs of alternative modes of organising work. He argues that the historical progression from the putting-out to the factory system is best explained in terms of the superior efficiency properties of the latter. In particular, putting-out suffered from three basic transaction costs drawbacks: "embezzlement and like deceits", work flow and quality work deficiencies, and weak incentives structures. Dispersed workstations made chronic theft and the substitution of lower quality supplies difficult to counter. Merchants were often reliant on the work pace and product quality supplied by a small number of outworkers. This made it difficult for merchants to respond quickly to fluctuations and changes in the composition of product demand. The factory system economised on each of these transaction costs: leakages were prevented, work flow and product quality were controlled, and workers could be dismissed for "malingering" (pp. 21-30).

The implication is that the factory system emerges because it is a more efficient form of organisation. On such an assessment, the factory system appears to be superior, in terms of efficiency, to the putting out system. However, a particular problem of method in undertaking such a comparison concerns the definition of efficiency employed. In other words, is it efficient in Pareto terms? And efficient for whom?

The move from the putting-out to the factory system engages a further loss of autonomy on the part of the workers to the capitalist employers, a point also admitted by Williamson (1985) himself. But if this is so, why then do rational workers voluntarily agree to establish more hierarchical relations of work (i.e. the factory), given to relative sacrifice in their autonomy vis a vis the employer? Thus, it could be presumed that the development of the factory system does not represent a Pareto efficient movement, since one party (i.e. the employer) becomes better off, while the other party (i.e. the worker) is made worse off, at least in terms of autonomy loss, assuming his/her remuneration does not change.

A possible counter-argument may be that workers also become better off since they can gain an increase (yet uncertain) in their wage, as their share in the efficiency gains grow, and hence it is in their interest to abandon their work independence. This seems to be Stiglitz's (1975, p. 571) assumption when he argues that "workers voluntarily undertake to be supervised ... They submit to being compelled to work harder than direct incentives provide for, because the consequence is a higher expected utility". This assertion, however, is very dubious for two reasons. First, as Marglin (1991, p. 238) puts it "we cannot talk first about the size of the pie (efficiency) and then about how to slice it (distribution), as mainstream economics are apt to do". The two are intrinsically interdependent. Second, it is very doubtful whether the pecuniary cost-benefit analysis could be extended to incorporate "psychological" costs. If so, it has to be explicitly shown rather than be taken for granted, that the gains from an increase in worker's wage exceed the costs from losing his/her independence. Indeed, it is highly questionable whether such a comparison is feasible, given the non-quantifiable nature of psychological and other non-pecuniary costs. In short, the only certain and indisputable fact in moving from the putting-out to the factory system is the cost in terms of the autonomy loss incurred by the worker. The rest, regarding higher remuneration resulting from efficiency gains, cannot be taken for granted as they depend crucially on the power of employees vis a vis their employers to enforce such increase in their earnings.

Thus, Williamson's preoccupation with (Pareto) efficiency seems to revolve around a fallacy arising from an attempt to separate artificially production from the distributional aspects of the problem, in studying economic organisations. It could, therefore, be argued that the emergence of the factory system represents not a step forward in the direction of Pareto improvement, but rather an augmentation of the notion of "class efficiency" (Bhaduri, 1991, p. 56), that is the ability of some powerful parties (i.e. capitalists employers) to ensure a sufficiently favourable distribution of income for themselves through a certain organisational arrangement, without being concerned with the efficiency of those institutions. As Pitelis (1998, p. 1010) argues, "the driving force has been efficiency for profit (power-distribution) and not efficiency for efficiency's sake (economizing) as the Coase-Williamson neoclassical new institutionalism suggests". As a result, the Pareto criterion becomes misleading when applied to the development of the hierarchical relations of work, which are shaped largely by the power relations among different classes involved in the organisation of production. Particularly, the emergence of the factory system serves the interest of employers by being a more profitable form of production irrespective of the effect on the worker's welfare. It reflects a series of choices made by powerful employers over work organisation, which serves to maximise their own private return. There is no presumption that these choices are organisationally efficient: "*the central hypothesis ... is that capitalist work organization does not maximize the size of the economic pie but the size of the boss's slice*" (Marglin, 1991, p. 238).¹⁰

¹⁰ In his classic article, Marglin (1974) argues that it was precisely the power to appropriate the returns from the organisation of production in the factory, which spurred capitalist employers to embrace this system. Supervision and control were arguably the key to the development of the factory as a central feature of the capitalist system of production. The development of the factory system was not in the interests of workers, who, being left with no alternative but the freedom to starve, were forced to comply. Yet, it did prove to be a far more efficient form of organisation. However, this efficiency refers more to the dynamic innovative efficiency associated with productivity increases and technological innovation, rather than its static Paretian version. Furthermore, this could only be argued ex post, and as such cannot form part of the *explanan* of the emergence of the factory system, other than on functionalist grounds.

4. Conclusions

The main implication is that the concept of collective power seems, in contrast to Williamson's adherence to the efficiency principle, to be a more effective analytical tool in explaining the establishment of the factory system. This also challenges the idea stemming from transaction cost economics, that hierarchical relations are the consequence of a voluntary contracting, i.e. that it is in the interests of both employers and workers to enter into a hierarchical organisation of work as a result of their optimising decision-making. The conclusion is that there is not sufficient theoretical evidence to substantiate the core proposition that higher forms of hierarchical organisation develop in response to the efficiency principle. This applies with even greater force to the question of the actual emergence of the hierarchical form itself. It follows that claims to the efficiency of the resultant forms of work organisation do not suffice as an explanation of their emergence. For a sufficient explanation of the emergence of the firm the issue of collective power should become an essential ingredient of the analysis and not just an optional extra.

References

- Alchian, Armen and Harold Demsetz (1972) "Production, Information Costs and Economic Organization", *American Economic Review*, vol. 62, no. 5, pp. 777-795.
- Ankarloo, Daniel and Giulio Palermo (2004) "Anti-Williamson: a Marxian critique of New Institutional Economics", *Cambridge Journal of Economics*, vol. 28, no. 3, pp. 413-429.
- Bhaduri, Amit (1991) "Economic Power and Productive Efficiency in Traditional Agriculture" in Gustafsson (ed.), pp. 53-68.
- Coase, Ronald, (1937), "The Nature of the Firm", *Economica*, vol. 4, no. 4, pp. 386-405, reprinted in Coase (ed.) pp. 33-55.
- Coase, Ronald (1993) "The Nature of the Firm: Influence", in Oliver Williamson and Sidney Winters (eds.) (1993) *The Nature of the Firm: Origins, Evolution, and Development*, Oxford: Oxford University Press, pp. 61-74.
- Commons, John (1990) [1934] *Institutional Economics: Its Place in Political Economy*, New Brunswick and London: Transaction Publishers.
- Dietrich, Michael (1994) *Transaction Cost Economics: Towards a new economics of the firm*, London: Routledge
- Gustafsson, Bo (ed.) (1991) *Power and Economic Institutions: Reinterpretations in Economic History*, Cheltenham: Edward Elgar.
- Hodgson, Geoffrey (1999), *Evolution and Institutions: On Evolutionary Economics and the Evolution of Economics*, Cheltenham: Edward Elgar.
- Landes, David (1966) *The Rise of Capitalism*, London: Macmillan.
- Marglin, Stephen (1974) "What Do Bosses Do? The Origins and Functions of Hierarchy in Capitalist Production", *Review of Radical Political Economics*, vol. 6, no. 2, pp. 60-112.
- Marglin, Stephen (1991) "Understanding Capitalism: control versus efficiency", in Gustafsson (ed.), pp. 225-252.
- Meramveliotakis, Giorgos (2018) "New Institutional Economics: A Critique of Fundamentals and Broad Strokes Towards an Alternative Theoretical Framework for the Analysis of Institutions", *Asian Journal of Social Sciences Studies*, vol. 3, no. 2, 50-64.
- Meramveliotakis, Giorgos and Dimitris Milonakis (2010) "Surveying the Transaction Cost Foundations of New Institutional Economics: A Critical Inquiry", *Journal of Economic Issues*, vol. XLIV, no. 4, pp. 1045-1071.
- Penrose, Edith (1959) *The Theory of The Growth of the Firm*, Oxford: Basil Blackwell.
- Pitelis, Christos (1991) *Market and Non-Market Hierarchies: Theory of Institutional Failure*, Oxford: Blackwell Publishing.
- Simon, Herbert (1961) [1947] *Administrative Behaviour*, New York: Mcmillan.
- Stiglitz, Joseph (1975) "Incentives, Risk, and Information: Notes Toward a Theory of Hierarchy", *Bell Journal of Economics*, vol. 6, no. 2, pp. 552-579.
- Williamson, Oliver (1975) *Markets and Hierarchies: Analysis and Antitrust Implications*, New York: Free Press.
- Williamson, Oliver (1980) "The Organization of Work: A Comparative Institutional Assessment", *Journal of Economic Behavior and Organization*, vol. 1, no. 1, pp. 5-38.
- Williamson, Oliver (1985) *The Economic Institutions of Capitalism*, New York: Free Press.
- Williamson, Oliver (1990) "A Comparison to Alternatives Approaches to Economic Organisation", *Journal of Institutional and Theoretical Economics*, vol. 146, no. 1, pp. 61-71.
- Williamson, Oliver and William Ouchi (1981) "The markets and hierarchies program of research; Origins, implications, prospects" in Andrew Van de Ven and William Joyce (eds.) (1981) *Perspectives on Organisation Design and Behaviour*, New York: Wiley, pp. 347-370.