
Opportunism is not the only reason why firms exist: why an explanatory emphasis on opportunism may mislead management strategy

Geoffrey M. Hodgson

The object of the present essay is neither to overturn transaction cost economics nor to deny the reality of opportunism. Instead, it is to question the explanatory role of the concept of opportunism in the transaction cost analysis of Oliver Williamson and his followers. Williamson has suggested that potential or actual opportunism emerges as the source of the 'transaction costs' involved in monitoring and enforcing contracts. On the contrary, it is shown here that there are several additional and likely sources of contract default or incomplete performance. Williamson's explanatory concentration on opportunism thus misidentifies the reasons for different hierarchical governance structures in the real world. As these additional sources are ignored, the emphasis on opportunism can impair a true understanding of the inner workings of the firm, and be a misleading guide for practical questions of organizational design and corporate strategy.

1. Introduction

Especially since the publication of Oliver Williamson's *Markets and Hierarchies* in 1975, the transaction cost approach to the study of firms and other governance structures has become established and recognized as a major theoretical and empirical field of study. Although several of its claims remain controversial, it is not the intention of this article to challenge transaction cost economics *per se*. Instead, the object of the present essay is to challenge the explanatory role of the concept of opportunism in the transaction cost analysis of Williamson and his followers.

Williamson has insisted that opportunism plays an essential role in his analysis. He argues that under conditions of imperfect information, all transactions are affected by the problem of 'self-interest seeking with guile'. Given the opportunity, agents may serve their own interests rather than those of the other party to the contract. Hence, according to Williamson, potential or actual opportunism emerges as the major source of the 'transaction costs' involved in monitoring and enforcing contracts.

For more than a quarter of a century, Williamson has repeatedly defended this line of argument. Many critics argue that, in the real world, agents are not so opportunistic as

Williamson would have it. Real-world firms, it is suggested, depend much more on commitment and trust. Some argue that human nature is not as Williamson describes. Be they right or wrong, these are not the critical arguments pursued here.¹

Instead, in this essay the *logical* centrality of the opportunism assumption to transaction cost analysis is challenged. It will be shown that, whether or not opportunism exists in the real world, it is not strictly necessary for Williamson's transaction cost argument. Importantly, there are major reasons additional to opportunism for the existence of governance structures. Having established these two key points, it is then argued that the *a priori* emphasis on opportunism in explaining governance can impair a true understanding of the inner workings of the firm, and can be a bad guide for practical questions of organisational design and corporate strategy.

Section 2 of this paper reviews Williamson's position on opportunism. Section 3 shows that opportunism is not the only reason for the existence of governance structures. Other functions and reasons for their existence are supplied. Section 3 considers how the excessive explanatory concentration on opportunism may mislead institutional analysis and management practice. Section 5 concludes the paper.

2. Williamson's explanatory emphasis on opportunism

Williamson is just part of a broader tradition of thought that emphasizes phenomena such as self-interest, opportunism and malfeasance. This 'neo-Hobbesian' tradition—as Bowles (1985) describes it—includes other prominent works such as those by Fama (1980), and Jensen and Meckling (1976). But Williamson has placed explanatory emphasis on the alleged centrality and necessity of the concept of opportunism in the analysis of transactions. His work is singled out for this reason.

Williamson (1975: 255) famously defined 'opportunism' as 'self-interest seeking with guile'. He argued that: 'Economic man is a much more subtle and devious creature than the usual self-interest seeking assumption reveals' (*ibid.*). Williamson (1985: 47) later elaborated the concept of opportunism in terms of 'the incomplete or distorted disclosure of information, especially to calculated efforts to mislead, distort, disguise, obfuscate, or otherwise confuse'.

His verdict concerning human nature is combined with a belief that opportunism is indispensable to transaction cost analysis. Williamson (1979: 234) wrote: 'there appears to be developing a general consensus' that 'opportunism is a central concept in the study of transaction costs'. This belief is repeated: 'the mitigation of opportunism plays a central role in transaction cost economics' (Williamson, 1993b: 115). Nevertheless, Williamson does not argue that all agents are opportunistic all the time. This point of clarification is also repeated in several works:

¹This does not of course deny the significance of these empirical issues. Questions of human nature, preferences or purposes are crucial for both the understanding and design of organizations. See, for example, Fehr and Gächter (2000).

Opportunism is a variety of self-interest seeking but extends simple self-interest seeking to include self-interest seeking with guile. It is not necessary that all agents be regarded as opportunistic in identical degree. It suffices that those who are less opportunistic than others are difficult to ascertain *ex ante* and that, even among the less opportunistic, most have their price. (Williamson, 1979: 234n.)

Similarly, Williamson (1993a: 98) later rejected a belief ‘in opportunistic practices most of the time. Rather, most economic agents are engaged in business-as-usual, with little or no thought to opportunism, most of the time.’ Hence Williamson’s *empirical* position concerning the existence of opportunism is difficult to challenge. In these quotations Williamson makes it clear that it is the threat of opportunistic behaviour, even by a small minority, that is significant. It is not that most people are opportunistic most of the time. This empirical stance is not being contested here.

Instead, the focus here is on the general and more problematic, theoretical and explanatory role of opportunism in his analysis. Williamson (1985: 66) wrote that market contracting ‘would be ubiquitous in the face of nonopportunism’. Williamson is insistent that the existence of some opportunism is necessary to explain complex contracting phenomena and hierarchical control: ‘But for opportunism, most forms of complex contracting and hierarchy vanish’ (Williamson, 1993a: 97). Similarly, ‘investment hazards associated with the exchange . . . would vanish were it not for opportunism’ (pp. 99–100). Williamson subsequently argues that in the absence of opportunism, quite different forms of organization would be observed. This particular passage is worth quoting in detail:

the incentive, control, and contract law differences that define alternative modes of governance . . . all vanish if opportunism is zeroed out. Thus (1) no incentive differences will appear because all members of every group subscribe to the same ‘general clause’ and implement the same objective function in the same fully committed way; (2) all cost-effective regularities (practices and procedures) that are adopted by one group will also be adopted by another—whatever the nominal form of organization (private firm, public bureau, nonprofit, autonomous market) from which they start; and (3) contract law differences serve no purpose among groups all of which share the same purpose and converge to the same form. Note, moreover, that . . . conflict and haggling . . . will never appear in opportunism-free groups . . . every such group will work out its differences instrumentally. (Williamson, 1999: 1099)

These words are his strongest statement that different governance structures depend necessarily for their existence on potential or actual opportunism. It is clear from this passage that Williamson believes that opportunism is a necessary condition for ‘incentive differences’ and ‘conflict and haggling’. Furthermore, if opportunism were absent, Williamson argues, people would act ‘in the same fully committed way’. For

Williamson, if opportunism were 'zeroed out', then there would be no functional difference between governance structures.

Williamson (1999: 1099) seems to suggest that the disappearance of opportunism would lead not only to the evaporation of any *relative* advantage of any mode of governance, but also to *absolute* gains in terms of full commitment. It makes no difference to the argument here if the exclusive stress is placed on claims concerning relative differences between governance structures. According to the relativist claim, if opportunism were absent, firms and markets would be of equal economic advantage. This relativist interpretation neither alters nor undermines the argument herein, which shows that the advantages of firm organization over markets are not due to opportunism alone.

Characteristically, Williamson (2000: 605) singles out opportunism as the only reason why 'mere promise cannot be used to overcome noncontractability'. The view taken here is, of course, that mere promises are not enough to deal with contractual problems, but that the reasons why such contractual problems arise is not opportunism alone.

If the exclusively opportunism-based explanation of governance structures were defective, then it could have adverse theoretical and practical consequences, if applied in the real world. A faulty explanation would be likely to lead to faults in corporate strategy and in the design of governance structures. These faults could emerge, whatever the actual level of opportunism in the real world. For instance, even if opportunism were rampant, aspects of governance or strategy that have important functions that are not directly related to opportunism might be neglected.

Note also the general and *a priori* character of Williamson's argument. He attempts to establish the proposition that *all* governance structures arise principally to deal with opportunism. This is quite different from the (plausible) empirical claim that some (or many) aspects of governance structures have the function of dealing with opportunism. On the contrary, it is argued here that governance structures exist for other important reasons.

It is true that Williamson highlights factors other than opportunism in his works. In particular, he attributes some contractual and governance problems to bounded rationality. But the above quotations make it clear that, for Williamson, the explanation of differences in governance structure devolves exclusively on opportunism, allegedly because 'the incentive, control, and contract law differences that define alternative modes of governance . . . all vanish if opportunism is zeroed out'.²

To repeat: my argument with Williamson is not over the existence of opportunism in

²It may be suggested that Williamson's (1999: 1099) strong statement concerning opportunism is not representative of his work as a whole. Whether this is true or not, it is important to clarify the logical status of opportunism in his argument, and show that the stronger propositions can be removed without necessarily undermining other parts of Williamson's theoretical contribution.

reality. Williamson may (or may not) overestimate the extent of opportunism in reality; but that is not the point. Instead, I wish to counter his arguments of the following type:

- if opportunism were absent then people would act ‘in the same fully committed way’;
- if opportunism were absent, then ‘conflict and haggling . . . will never appear’;
- if opportunism were absent then ‘investment hazards associated with the exchange . . . would vanish’; and
- if opportunism were absent then the reason for the existence of different modes of governance to enforce contract compliance would disappear.

Against all these propositions, I shall argue that, even in the absence of opportunism, full commitment may be lacking, and conflict and haggling may still arise. Even more importantly, the absence of opportunism would not abolish ‘investment hazards associated with the exchange’ and modes of governance to enforce or maximise contract compliance would still be necessary. In this respect, some modes of governance would still retain advantages over others.

Hence opportunism is not a necessary condition for the phenomena and differences in structures that Williamson describes.³ Hence Williamson’s explanatory emphasis on opportunism is misplaced, *not* because opportunism is absent in the real world, but because there are *additional* and very important sources of the contracting problems that he identifies. The misidentification of opportunism as the principal source of these problems not only misunderstands the real basis of governance structures but also can promote potentially disastrous approaches to institutional design and management practice.

If the critical arguments in this paper are correct, then other aspects of transaction cost analysis may have to be re-examined, including the explanation of different governance structures and the existence of the firm. However, it is not the purpose of this paper to rebuild transaction cost theory, simply to criticize some of its fundamental assumptions concerning opportunism, and to show how the excessive explanatory emphasis on opportunism may mislead management practice. Some further implications of this critique will have to be spelt out at a later date.

³Hodgson (1988: 205) and Langlois (1984, 1995) have previously proposed—but without the full explanation and analysis provided in Section 3 below—that opportunism is not a necessary element in the explanation of governance modes. Coase (1988: 45) also became convinced that ‘opportunism’ was not a general problem and should not be included in any general theory of the firm. Also Alchian and Woodward (1988: 66) briefly remarked that: ‘different but honest perceptions can lead to disputes that are costly to resolve’. See also Madhok (1996), Stephen (1996), Demsetz (1997: esp. p. 21) and Hendry (2002). It is the intention of the present essay to begin to give these neglected matters of perception and interpretation much more analysis and attention. This essay also underlines for the first time some of the possibly deleterious practical consequences of a faulty governance analysis based on opportunism alone.

3. Why opportunism is not the only reason for the existence of differences in governance structures

Henceforth, we shall exclude the unrealistic case where parties have common aims that are consistent with the same objective function.⁴ In reality, management and governance problems arise in contexts where individuals have some different purposes, even if some aims are shared. The more people involved the less likely it is that everyone has the same aims. Some differences in objectives are assumed here. Contrary to Williamson (1993a, 1999), it is argued that for individuals that have different objectives but are nevertheless non-opportunistic, modes of governance to enforce contract compliance would still be necessary.

This section examines some ubiquitous sources of contracting problems, other than opportunism. The analysis is quite general, in that it applies to all principal-agent relationships with non-identical aims. However, the intention is to show that governance structures can arise for different reasons, not to provide a general theory of their origin. At the start the focus is on codified communication. The possibility of limits on codifiable communication will be raised later. It is also assumed for convenience that the principal is consistent in his or her aims at this specific time. (The argument below is strengthened if this consistency assumption is dropped.) Five steps in the communication process between principal and agent are now considered.

1. Principal *P* has an intention consistent with codified instructions X_1 .
At this first stage, the principal's intentions are formed and they are assumed to be consistent with a set of written or verbal instructions X_1 (whether or not these instructions are actually expressed explicitly).
2. Principal *P* codifies intention X_1 in terms of codified instructions X_2 and communicates these to agent *A*.
However, the possibility exists that the principal will not be able to express these intentions well, or make mistakes or omissions in doing so. Most typically, the principal will be obliged to give a summary of the required task and be unable, due to lack of time or specific expertise, to express it in full detail. Hence X_2 may be a summary, or even a partial distortion, of X_1 .
3. Agent *A* interprets codified instructions X_2 in perceived terms X_3 .

⁴Would different governance structures exist if the parties to a contract shared common aims, wished to maximize the same 'objective function', and also knew that they shared precisely the same aims? In such an extreme case, mutual consultation or perhaps the delegation of authority to one of the two parties might be all that is necessary to ensure contract compliance. No more than minimal governance would be necessary. But note that this argument is quite independent of the question of opportunism. If two parties wish to maximise the same objective function then the 'self-interest' of one is also the 'self-interest' of the other. They could still be pursuing their 'self-interest with guile' but the necessary degree of governance would be nevertheless reduced. This outcome results from the sharing of objectives, not from a reduction of opportunism, unless opportunism is redefined in a way to exclude the possibility that the interest of one party could be the same as the interest of the other. Williamson does not restrict the meaning of opportunism in this way.

The agent receives this codified information and interprets it in terms of his or her understandings of the terms and meanings involved. All communication involves interpretation and depends on the conceptual framework of meanings held by the receiver. Because all people differ in their cognitive frameworks, the agent's understanding of all the categories or concepts involved will never be identical to that of the principal. Often these differences may be insignificant, but sometimes they will not. This is quite independent of any possible wilful misinterpretation or disobedience.

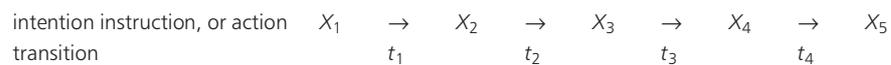
4. Because of differences in objectives and limited monitoring, agent *A* wilfully adapts perceived instructions X_3 to outcomes that are instead consistent with codified instructions X_4 .

This step admits the possibility of wilful disobedience or adaptation of the principal's perceived instructions, for whatever reason. It is at this point, *and only at this point*, that opportunism may have an effect. But a difference between X_3 and X_4 may occur for reasons other than opportunism. For whatever reason, the agent may disagree with the instructions, or see them as counter to his or her own interests. The instructions are thus changed or adapted to the extent that is possible without the agent incurring penalty or dismissal. The extent to which this occurs is the difference between X_3 and X_4 . It is possible that this gap could be insignificant or insignificant.

5. Agent *A* attempts to act in accordance with X_4 but, due to a limited capacity to foresee the outcomes of his or her actions, he or she in fact performs in a way that is consistent with instructions X_5 .

Acting in a way that is consistent with X_4 does not imply this outcome. Actions are not always consistent with intentions. Due to complex, misunderstood or unforeseen circumstances, the result of pursuing X_4 might be quite different, namely X_5 .

No less than *four* potential shifts of meaning are involved in the transitions from the principal's intention X_1 to the agent's actions X_5 . These may be represented as follows:



where t_1 , t_2 , t_3 and t_4 denote the four transitions and possible shifts of meaning of the instructions.

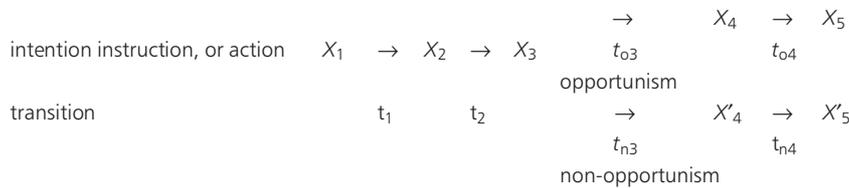
A real-world example of a shift from X_1 to X_2 and from X_2 to X_3 is illustrative of the kind of problems involved. A British relative of the author is a highly capable personal assistant who works for a company in America. She brought a one-page document to the attention of her boss who then instructed her to 'post it'. While Americans use the word 'mail' to refer to postal services, the British more commonly use the term 'post' in the same context, also as a noun or a verb. The instruction was thus interpreted as one to put the document in an envelope and to mail it to its named addressee. What the boss in fact meant was that the document should be posted on the company notice board. However, the instruction was completely misinterpreted, not because of any

opportunism but partly because of inadequate instructions (step t_1) and partly because of differences in use of language (step t_2).

With a little imagination we can conceive of other such possibilities of misinterpretation based on inadequate instructions and no doubt they are plentiful in practice. Note that it is not the burden of this paper to provide empirical evidence of the extent of such misinterpretations. We only have to assume a significant *potential* for misinterpretation to undermine Williamson’s argument that without opportunism firms would not exist.

Let us now focus on t_3 . This adaptation, from X_3 to X_4 , can occur for several reasons. One of which is opportunism, as defined by Williamson. In the following diagram, t_3 represents the ‘opportunism’ reason for changing X_3 to X_4 . Clearly, even if opportunism is rampant, it does not explain the degree of change involved in the whole transition from X_1 to X_5 .

Furthermore, reasons other than opportunism may be responsible for the particular transition from X_3 to X_4 . The agent may have a moral objection to X_3 and choose X'_4 for this reason. Examples of moral objections might be that instruction X_3 is interpreted as racist or sexist, or involves an objectionable political purpose, or is perceived as contributing to global warming or environmental degradation. Other non-opportunistic reasons for changing X_3 are also possible. For example, the agent may believe that X_3 is not in the interests of the company, or goes against an agreed or published corporate strategy. In the following diagram, t_{n3} represents ‘moral objection’ and all other possible non-opportunistic reasons for changing X_3 to X'_4 .



It is now clear that in his argument for the centrality of opportunism, Williamson concentrates exclusively on transition t_{o3} and ignores all other potentially distorting transitions, namely t_1 , t_2 , t_{n3} , t_{o4} and t_{n4} .⁵ In contrast to Williamson, any one of these transitions can give rise to outcomes that are different from the intentions of the principal. Hence governance structures have to be developed that minimize the scope and effect of all potentially distorting transitions, not simply of opportunism. Furthermore, governance modes that deal with opportunism may not be so effective in dealing with other sources of imperfect compliance.

⁵In at least one passage Williamson (1975: 123) does mention problems of communication and cognition that might arise in contracting relationships. However, Williamson does not draw the correct inference (developed here) that problems of communication and cognition are additional reasons for the establishment of governance structures.

Notice also that even if there were descriptively complete contracts, differences of interpretation could exist. Completeness of description neither guarantees interpretative concordance nor excludes misunderstandings. Hence some monitoring might be required even if contracts were descriptively complete. Of course, the notion of 'complete contract' could be defined as contracts that are both descriptively complete and incapable of multiple interpretations or misunderstandings. In which case, any contract non-compliance could be addressed by breach-of-contract proceedings rather than by monitoring. However, it is well known (in philosophy and linguistics) that such an unambiguous description is generally impossible. The words and terms used in such a description can never be completely and unambiguously defined, and they always depend for their meaning on an extended network of other terms and concepts (Wittgenstein, 1953; Quine, 1960). No real-world attempt to move towards a completely specified contract can itself remove any need for monitoring. If monitoring can be eliminated, then it will be for different reasons.

The introduction of non-codifiable communication reinforces, rather than undermines the above conclusions. If information is tacit or uncodified, then the scope for misinterpretation is even greater. In addition, the principal may not know enough about the detailed production processes to communicate detailed instructions. These may have to be inferred by the agent. Hence there is an enhanced possibility of errors of inference, as well as errors of interpretation. Arguably, following Polanyi (1967) and others, tacit knowledge is the indispensable foundation of all knowledge. Accordingly, problems of codification of some tacit knowledge, and of transmitting the uncodifiable residual, are ubiquitous in social and economic life.

Williamson's failure to emphasize potential mismatches of interpretation betrays an empiricist treatment of information. Implicitly, Williamson assumes that information is readily and unambiguously understood, despite the different cognitive frameworks of the communicators. The perception, interpretation and cognitive framing of information is ignored (Hodgson, 1988). This partly accounts for his extraordinary statement that people would all act 'in the same fully committed way' if it were not for opportunism.

Transaction costs can involve costs of contract formulation, interpretation, monitoring and enforcement. Williamson sees these problems as giving no advantages to hierarchy if opportunism were absent. On the contrary, the position advanced here is that the origin of these costs is even more general and ubiquitous than opportunism. The essential and general problem is one of interpretation. The communication of an instruction always carries the possibility of default, because it can always be interpreted in a different way. Communication itself is costly, and there is a trade-off between the principal providing a more complete set of instructions, thus reducing (but never entirely eliminating) the possibility of misinterpretation, and the provision of a shorter instruction that is not so time consuming in its articulation.

Essentially, in order to admit costly contract enforcement, all that transaction cost economics requires is a possibility of contract default. This possibility can arise because

of wilful misinterpretation or because of the existence of different cognitive frameworks or because of differences in knowledge. Opportunism may be part of reality but it is not the only explanation of the existence of different governance structures.⁶

4. Why the explanatory emphasis on opportunism may mislead management strategy

It is clear that although Williamson does not regard every agent as opportunistic, he does regard opportunism as the principal source of possible contract non-compliance. This general theoretical argument has important analytical and policy consequences. If opportunism is the principal source of all contract non-compliance, then all such shortfalls are mainly caused by opportunism. Furthermore, if opportunism is the main reason why one governance structure has advantages over others, then any economic justification of hierarchy over markets rests principally on the existence of opportunism. Consequently, the sole justification of management hierarchy and monitoring is seen as opportunism, and it is proclaimed that management's principal role is to keep it in check. Although not everyone is opportunist, Williamson is nevertheless obliged by the logic of his own argument to identify opportunism as a likely reason for any contractual shortfall. For Williamson, the special advantage of corporate hierarchies (when compared, for example, with markets) is that they are little more than gendarmes against opportunist behaviour.

Needless to say, this is a highly contestable conception of corporate management and leadership. It is not the purpose of this paper to argue an alternative conception, simply to point out that Williamson's conception is denied by a whole range of studies of modern organisations, from Argyris and Schön (1996) through March (1988) to Senge (1990) and countless others. In particular, Conner (1991), Kogut and Zander (1992), and Conner and Prahalad (1996) have attempted to construct a theory of the firm that does not rely on incentive problems and opportunism. The validity or otherwise of these alternative accounts is partly an empirical question. In contrast, Williamson proposes a general theory of governance structures in which opportunism is always the compelling reason for the existence of one type of structure rather than another.

The aim here is not to establish a full alternative explanation of corporate governance, but to use the existing literature to establish a *prime facie* case that an exclusive explanatory emphasis on opportunism may lead to misunderstandings concerning the role of management and thereby mislead management strategy. The rectification of the flaw in Williamson's argument concerning the role of opportunism brings other factors

⁶In a concordant identification of problems of potential misinterpretation, based on some interesting case studies, Helper *et al.* (2000: 445) emphasize 'the ambiguity and provisional nature of all understanding, from the simplest verbal exchange to the most complex co-development project. As a consequence, interlocutors and partners must cooperate in pursuit of mutual intelligibility as a condition of self-understanding.'

into consideration, and the relative and practical importance of these factors is largely a matter of empirical investigation.

Alternative conceptions of corporate governance, based on the counter-argument elaborated here, would recognize the reality of opportunism but place no exclusive stress upon it. On the contrary, the role of hierarchical governance structures in directing employees may not simply be to minimize the effects of opportunism but also to minimize all distortions in cognitive and communicative transitions, namely t_1 , t_2 , t_{n3} , t_{o4} and t_{n4} . Accordingly, the role of management, even if we are concerned simply with the question of carrying out instructions, may not be principally a gendarme, but equally if not more an educator, concerned with the minimization of misunderstandings and the development of a common corporate culture where—as much as possible—shared aims and conceptions dominate sectional interests. Importantly, the creation, transmission and interpretation of knowledge within the firm must build on these shared values and conceptions. It is perfectly feasible that some types of governance structure may have advantages in these respects over others, in some contexts. If so, it may be a major reason for the existence of the firm.⁷

To some degree, the management roles of gendarme and educator may conflict with one another. The cultivation of a common corporate culture may require a degree of trust between employee and employer and between employees themselves. However, habits of trust may be undermined by the excessive use of the gendarme role and the presumption by management that employees are potentially opportunistic. As Jacoby (1990: 334) reports:

as industrial studies have repeatedly shown, the presumption of innate opportunism is fatal to trust. . . . It leads to a proliferation of control structures—supervision, rules, and deferred rewards—intended to inhibit opportunism. These create resentment and distrust among employees, who correctly perceive the controls as expressions of their employer's distrust.

A manager inspired by Williamson's analysis would concentrate principally on the installation of monitoring structures to minimize opportunistic behaviour. But he or she would place less emphasis on other, more educative, measures that would attempt to minimize possible communicative and cognitive distortions in transitions t_1 , t_2 , t_{n3} , t_{o4} and t_{n4} . The deleterious result of such a policy could be to undermine trust in the organization, reduce free and open communication about problems, and increase the possible scale of these other distortions. As a result, the application of Williamson's theory to management practice can have negative results.

Note, however, that the above argument still assumes a basis for traditional, hierarchical roles. It assumes a 'Taylorist' model of management where the managers are in full possession of all the relevant information and the role of employees is to carry

⁷This is consistent with the empirical evidence provided by Wernerfelt (1997) that the lower costs of contract adjustment within the employment relationship is an important reason for the existence of the modern firm.

out instructions. It is shown that, even in this Taylorist case, Williamson's analysis is defective and its implications for corporate practice can be misleading.

The problems for Williamson's analysis become even greater when we move away from a Taylorist model of work organization. In some circumstances, monitoring would be ineffective, even with access to all information and in the absence of opportunism. Minkler (1993) investigates the internal organization of the firm where entrepreneur and worker differ in knowledge, interpreted as different understandings of production possibilities. Sometimes principals cannot monitor agents even when information on behaviour is costless, because they do not have the necessary knowledge about production processes or outcomes. In some situations, the principal would not even be able to decipher what was observed. Minkler argues that this situation of dispersed knowledge requires a more participatory governance structure for the firm. This provides greater incentives for the worker to use his or her specialised knowledge, and greater opportunity to do so.

In another work, I have argued that the thrust of capitalist development involves increasing complexity, variety and specialization of skills (Hodgson, 1999). In these circumstances, it is less and less likely that an employer will have the same specialist skills as many workers in his or her employment. As acknowledged by other writers (Demsetz, 1988; Conner and Prahalad, 1996), it is possible that the employer will often lack some particular skills and knowledge held by the employees. In such circumstances, monitoring by an employer cannot be effective, even if opportunism is absent. The problem of management within the firm becomes less one of communicating orders and more one of enabling learning and innovation. Within broad boundaries, employment becomes less a matter of contract compliance and more a matter of the growth of capabilities and knowledge. If these developments are born out in reality, then the emphasis on purely contractual aspects of employment is somewhat one-sided.

Williamson's general theory is also insensitive to differences in (historically or geographically) specific governance structures and the explanations of their efficacy. For instance, addressing the performance of the Japanese economy, Williamson (1985: 122) wrote: 'the hazards of trading are less severe in Japan than in the United States because of cultural and institutional checks on opportunism'. But this is, at best, a one-sided explanation. The efficacy of Japanese corporate structures might have a lot more to do with the capacity to minimise distortions arising from transitions, namely t_1 , t_2 , t_{n3} , t_{o4} and t_{n4} . Japanese corporate culture might be more effective in nurturing a common aim and sharing of knowledge and technical understanding (Nonaka and Takeuchi, 1995). Contrary to Williamson, explanations of economic success in organizations do not have to devolve principally on the minimization of opportunism.

This argument is valid even if we find cases where actual and potential opportunism is indeed a major factor in contract formulation, enforcement and outcomes. We would have relatively little difficulty in listing many case studies that show that the anticipation of opportunistic behaviour is important in some contexts. Risk aversion, self-interest,

greed and untrustworthiness are part of modern corporate reality. But so too are the cognitive and communicative problems identified in this essay.

To repeat, the existence and potential significance of opportunism is not denied here. The point of contention is whether opportunism is the *necessary* reason for the existence of different governance structures and whether hierarchical governance is *always* explained adequately by the threat of opportunism. What is being denied here is a universal inference, not an empirical statement about the extent to which opportunism exists in reality.

Assume that actual or potential opportunism is rampant in a particular context. In this case it would be appropriate to place this issue at the centre of contract and governance design. Management policy would rightly take this into account. However, it is argued here that managers and other business actors should also take into account the potential problems of communication and cognition as well. Even if opportunism were widespread, an explanatory emphasis on opportunism in matters of governance and corporate strategy would not be warranted.

Crucially for management practice, the extent and intensity of potential or actual opportunism may vary from institution to institution and culture to culture. Accordingly, there is a danger that opportunism might be underestimated or overestimated in some contexts. The determination of the actual and potential extent of opportunism is a matter of investigation and analysis of specific institutional and governance structures. What has been shown here is that, in general, there are *additional* potential sources of contract non-compliance. Their importance and impact is also a matter of investigation into specific structures and contexts. But these additional potential sources of non-compliance cannot be ruled out *a priori*, and in general it is illegitimate to concentrate on opportunism alone. Guidelines for management strategy and practice should be developed with this warning in mind, and with attention to the specific institutions and structures that pertain.

There are costs and benefits involved in dealing with opportunism, just as there are costs and benefits with any educative management strategy. These costs and benefits have to be evaluated in specific contexts. There is no *a priori* reason to assume that the net costs of opportunism will generally dominate, even if they do so under specific circumstances.

It is reasonable to argue that the creation of shared values and common understandings can occur in other interactive settings, outside the firm. What about markets and networks? The correction of faulty interpretations, the cultivation of shared norms, the reduction of communication costs, and the development of shared modes of understanding may also occur through social mechanisms in markets or networks. Perhaps the extent to which this may happen will depend on the longevity of social contacts. But essentially this is a different issue, and the outcome does not affect the central argument in this paper. To see this, assume that markets or networks could sometimes be more efficient than the firm in this respect. The central argument of this section would still remain valid, because an explanatory emphasis on opportunism within the firm would

still lead to the neglect of deficiencies of communication, interpretation and understanding within that organisation. The explanatory emphasis on opportunism would still mislead management practice *within the firm*. However, if the extension of markets and the contraction of the domain of management control would help to further reduce these shortcomings, then managers would *also* have to concentrate on questions of possible outsourcing, divestment or restructuring. What has been shown here is that the explanatory emphasis on opportunism is generally misleading, whatever other considerations may pertain.

5. Conclusion

The central purpose of this paper is not to present an alternative analysis but to identify and explore the theoretical and practical implications of a logical flaw in Williamson's theory. Williamson insists that, but for opportunism 'most forms of complex contracting and hierarchy vanish'. He suggests that without the possibility of opportunism, transacting difficulties would be greatly reduced and a high degree of common commitment would prevail.

Against these propositions, it has been argued here that, even in the absence of opportunism, full commitment may be lacking, and conflict and haggling may still arise. The absence of opportunism would not abolish investment hazards associated with exchange; modes of governance to alleviate incomplete contract compliance would still be necessary. Furthermore, even if opportunism were absent, there are still reasons why some forms of governance structure retain advantages over others, in specific contexts.

Note also that the argument here does not depend on any assumption of 'irrational' behaviour. Agents can still be regarded as rational in the sense that they are maximizing their own utility. Even if they act in the interests of others then this behaviour can be seen as utility maximization (Collard, 1978). No particular assumption is made here about selfish or non-selfish behaviour. The arguments in this paper remain valid even if self-interest and opportunism are ubiquitous. The key proposition here is that likely and significant sources of contract non-compliance are misinterpretation, misunderstanding or disagreement. None of these assumptions is logically inconsistent with utility maximization. Whether utility maximizing behaviour is an adequate assumption is a quite separate matter (Hodgson, 1988, 1998).

It has been shown here that opportunism is not a necessary condition for the governance phenomena and structures that Williamson illuminates. As a result, Williamson's explanatory emphasis on opportunism is misplaced, not because opportunism is absent or unimportant in the real world, but because there are other likely and prominent sources of the contracting problems that he identifies. A misdiagnosis of contracting difficulties, as emanating largely from opportunism or malfeasance, not only misunderstands the basis of governance structures but also promotes defective approaches to institutional design and management practice.

The central argument in this paper would not be undermined even if opportunism were widespread in the real world. It has been shown that in general there are additional reasons for the existence of different governance structures, notably to deal with ubiquitous problems of communication, cognition and interpretation.

Note that these conclusions do not depend on any 'utopian' denial of the possibility of opportunism. At least for the purpose of this argument, Williamson's statements about the extent of opportunism have not been denied. No differences of estimate of the extent of real world opportunism are relevant here. If such differences of estimate exist then they must be resolved through empirical study. Instead, the main point here is that Williamson ignores several other reasons for the existence of hierarchical management structures. The undue emphasis criticised here is not on the *existence* of opportunism but on the identification of opportunism as the *principal cause* of management hierarchy. This has serious implications for both analysis and management strategy.

A principal conclusion of this paper is that matters of cognition and interpretation are vitally important in securing contract compliance and their study deserves a much more substantial place in the literature on corporate governance and strategy.⁸ Essentially, to show that the argument in the present paper is incorrect, it would have to be demonstrated that the problems of potential miscommunication and misinterpretation that have been elaborated here do not and cannot exist in reality. That too would be an example of a 'utopian fantasy' of a type against which Williamson (1999: 1099) has himself warned.

This paper points to an extensive agenda of empirical enquiry where the actual, potential and relative importance of opportunism, alongside problems of communication, cognition and interpretation, are all assessed in different organisations and cultures. Management practice should be informed by actual and likely empirical realities, not by a false universal and *a priori* assumption that without opportunism, no governance structures would have any advantage over another.

Acknowledgements

The author is very grateful to Nicolai Foss and anonymous referees for critical and helpful comments on earlier versions of this essay.

Address for correspondence

Malting House, 1 Burton End, West Wickham, Cambridgeshire CB1 6SD, UK. Email: g.m.hodgson@herts.ac.uk.

⁸Fransman (1994), Ghoshal and Moran (1996), Helper *et al.* (2000), Nooteboom (2000) and Witt (2000) emphasize matters of communication, cognition and interpretation in the firm. Future research could build on these and other precursors, and link with the 'competences' or 'capabilities' theory of the firm, at least in its stress on knowledge and learning (Dosi and Marengo, 2000).

References

- Alchian, A. A. and S. Woodward (1988), 'The firm is dead; long live the firm: a review of Oliver E. Williamson's *The Economic Institutions of Capitalism*,' *Journal of Economic Literature*, **26**, pp. 65–79.
- Argyris, C. and D. A. Schön (1996), *Organizational Learning II: Theory, Method, and Practice*. Addison-Wesley: Reading, MA.
- Bowles, S. (1985), 'The production process in a competitive economy: Walrasian, Neo-Hobbesian, and Marxian models,' *American Economic Review*, **75**, 16–36.
- Coase, R. H. (1988), 'The nature of the firm: origin, meaning, influence,' *Journal of Law, Economics, and Organization*, **4**, 3–47. Reprinted in Williamson, O. E. and S. G. Winter (eds) (1991), *The Nature of the Firm: Origins, Evolution, and Development*. Oxford University Press: Oxford and New York.
- Collard, D. (1978), *Altruism and Economy: A Study in Non-selfish Economics*. Martin Robertson: Oxford.
- Conner, K. R. (1991), 'A historical comparison of resource-based theory and five schools of thought within industrial organization economics: do we have a new theory of the firm?,' *Journal of Management*, **17**, 121–154.
- Conner, K. R. and C. K. Prahalad (1996), 'A resource-based theory of the firm: knowledge versus opportunism,' *Organization Science*, **7**, 477–501.
- Demsetz, H. (1988), 'The theory of the firm revisited,' *Journal of Law, Economics, and Organization*, **4**, 141–62. Reprinted in Williamson, O. E. and S. G. Winter (eds) (1991), *The Nature of the Firm: Origins, Evolution, and Development*. Oxford University Press: Oxford and New York.
- Demsetz, H. (1997), *The Economics of the Business Firm: Seven Critical Commentaries*. Cambridge University Press: Cambridge.
- Dosi, G. and L. Marengo (2000), 'On the tangled discourse between transaction cost economics and competence-based views of the firm: some comments,' in N. Foss and U. Mahnke (eds), *Competence, Governance and Entrepreneurship*. Oxford University Press: Oxford, pp. 80–92.
- Fama, E. F. (1980), 'Agency problems and the theory of the firm,' *Journal of Political Economy*, **88**, 288–307.
- Fehr, E. and Gächter, S. (2000) 'Fairness and retaliation: the economics of reciprocity,' *Journal of Economic Perspectives*, **14**, 159–181.
- Fransman, M. (1994), 'Information, knowledge, vision and theories of the firm,' *Industrial and Corporate Change*, **3**, 713–757.
- Ghoshal, S. and P. Moran, (1996), 'Bad for practice: a critique of the transaction cost theory,' *Academy of Management Review*, **21**, 13–47.
- Helper, S., J. P. MacDuffie and C. Sabel, (2000), 'Pragmatic collaborations: advancing knowledge while controlling opportunism,' *Industrial and Corporate Change*, **9**, 443–487.
- Hendry, J. (2002), 'The principal's other problems: honest incompetence and the specification of objectives,' *Academy of Management Review*, **27**, 98–113.

- Hodgson, G. M. (1988), *Economics and Institutions: A Manifesto for a Modern Institutional Economics*. Polity Press and University of Pennsylvania Press: Cambridge and Philadelphia.
- Hodgson, G. M. (1998), 'The approach of institutional economics,' *Journal of Economic Literature*, **36**, 166–192.
- Hodgson, G. M. (1999), *Economics and Utopia: Why the Learning Economy is not the End of History*. Routledge: London and New York.
- Jensen, M. C. and W. H. Meckling (1976), 'Theory of the firm: managerial behavior, agency costs and ownership structure,' *Journal of Financial Economics*, **3**, 305–360.
- Jacoby, S. M. (1990), 'The new institutionalism: what can it learn from the old?,' *Industrial Relations*, **29**, 316–359.
- Kogut, B. and U. Zander (1992), 'Knowledge of the firm, combinative capabilities, and the replication of technology,' *Organization Science*, **3**, 383–397.
- Langlois, R. N. (1984), 'Internal organization in a dynamic context: some theoretical considerations,' in M. Jussawalla and H. Ebenfield (eds), *Communication and Information Economics: New Perspectives*. North-Holland: Amsterdam, pp. 23–49.
- Langlois, R. N. (1995), 'Capabilities and coherence in firms and markets,' in C. A. Montgomery (ed.), *Resource-based and Evolutionary Theories of the Firm: Towards a Synthesis*. Kluwer: Boston, MA, pp. 71–100.
- Madhok, A. (1996), 'The organization of economic activity: transaction costs, firm capabilities and the nature of governance,' *Organization Science*, **7**, 577–590.
- March, J. G. (1988), *Decisions and Organizations*. Basil Blackwell: Oxford.
- Minkler, A. P. (1993), 'Knowledge and internal organization,' *Journal of Economic Behavior and Organization*, **21**, 17–30.
- Nonaka, I. and H. Takeuchi (1995), *The Knowledge-creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press: Oxford and New York.
- Nooteboom, B. (2000), *Learning and Innovation in Organizations and Economies*. Oxford University Press: Oxford and New York.
- Polanyi, M. (1967), *The Tacit Dimension*. Routledge and Kegan Paul: London.
- Quine, W. van Orman (1960), *Word and Object*. Harvard University Press: Cambridge, MA.
- Senge, P. M. (1990), *The Fifth Discipline: The Art and Practice of the Learning Organization*. Doubleday: New York.
- Stephen, F. H. (1996), 'Lawyers, transaction costs and opportunism,' *Journal of Institutional and Theoretical Economics*, **152**, 146–153.
- Wernerfelt, B. (1997), 'On the nature and scope of the firm: an adjustment cost theory,' *Journal of Business*, **70**, 489–514.
- Williamson, O. E. (1975), *Markets and Hierarchies: Analysis and Anti-trust Implications: A Study in the Economics of Internal Organization*. Free Press: New York.
- Williamson, O. E. (1979), 'Transaction-cost economics: the governance of contractual relations,' *Journal of Law and Economics*, **22**, 233–261.
- Williamson, O. E. (1985), *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. Macmillan: London.

- Williamson, O. E. (1993a), 'Opportunism and its critics,' *Managerial and Decision Economics*, **14**, 97–107.
- Williamson, O. E. (1993b), 'Transaction cost economics and organization theory,' *Industrial and Corporate Change*, **2**, 107–156.
- Williamson, O. E. (1999), 'Strategy research: governance and competence perspectives,' *Strategic Management Journal*, **20**, 1087–1108.
- Williamson, O. E. (2000), 'The new institutional economics: taking stock, looking ahead,' *Journal of Economic Literature*, **38**, 595–613.
- Witt, U. (2000), 'Changing cognitive frames—changing organizational forms: an entrepreneurial theory of organizational development,' *Industrial and Corporate Change*, **9**, 733–55.
- Wittgenstein, L. (1953), *Philosophical Investigations*. Basil Blackwell: Oxford.