

# THE CONCEPTUALIZATION OF 'TRUST' IN ECONOMIC THOUGHT

**Dominic Furlong\***

## **Summary**

Attempts to conceptualise trust has led to a confusing array of definitions across the social sciences. This paper focuses on the emergence of trust in economic thought and clarifies its conceptualisation as a term used to identify a purely calculative theory of behaviour in which agents form subjective probabilities regarding the future action of others. This conceptualisation of trust is distinguished by two assumptions: (i) boundedly rational agents seek to minimise the costs of acquiring information; and (ii) agents' subjective probabilities regarding the future action of others are 'Bayesian' in that they can be attached to any future event and every belief relevant for decision-making is captured therein. Alternatively, if it is assumed that boundedly rational agents find some calculations too difficult to perform even if the information is available, trust is better conceptualised as a complexity-reducing social norm. This alternative conceptualisation of trust is grounded in sociology, but depending on the importance given to social context in forming agents' predispositions towards trusting (and non-trusting) behaviour, it can be employed both as a substitute and as a complement to the calculative conceptualisation of trust found in economic discourse.

## **Contents**

I	INTRODUCTION: AN 'ELUSIVE NOTION'	1
II	THE EMERGENCE OF TRUST IN ECONOMIC THOUGHT	3
III	TRUST AS A CALCULATIVE PROCESS	7
IV	TRUST IN A SOCIAL CONTEXT	13
V	CONCLUSIONS: MARKING THE DISCIPLINARY BOUNDARIES	17

---

\* The author would like to thank John Humphrey and Hubert Schmitz for their helpful comments on earlier drafts.

## **I. Introduction: An ‘Elusive Notion’**

Across the social sciences the concept of ‘trust’ is gaining increasing attention. Recognition of its importance, however, is not a new phenomenon. As long ago as J.S. Mill (1891: 68) it was acknowledged that, [t]he advantage to mankind of being able to trust one another penetrates into every crevice and cranny of human life.’ Despite many such testaments as to the importance of trust in everyday situations, until recently few, if any; social scientists have attempted to unpack the concept. As Gambetta (1988: ix-x) argues, beyond acknowledgement of its importance, analysis of trust has been undermined by its pervasiveness in the face of the increasingly specialised social sciences:

[T]his very pervasiveness [of trust] seems to have generated less analysis than paralysis: in the social sciences the importance of trust is often acknowledged but seldom examined, and scholars tend to mention it in passing, to allude to it as a fundamental ingredient or lubricant, an unavoidable dimension of social interaction, only to move on to less intractable matters.  
*(ibid.)*

In the past decade, however, interdisciplinary discourse has grown and trust has emerged as an evocative theme upon which interest across the social sciences has converged (Swedberg 1987: 131). The literature on trust extends to the realms of philosophy (Williams 1988), political science (Dunn 1984); and social anthropology (Hart 1988), but its conceptualisation has commanded particular attention from economists and sociologists. Trust has arisen as a key issue in transaction cost economics (Williamson 1993), game theory (Dasgupta 1988), sociology (Luhmann 1988), and economic sociology (Granovetter 1985). Yet despite the plethora of material emerging on the subject, in the words of Gambetta (1988: ix), trust remains an ‘elusive notion’. As Shapiro (1987: 625) notes ‘the conceptualisation [of trust] has received considerable attention in recent years, resulting in a confusing potpourri of definitions applied to a host of units and levels of analysis’. If the pervasiveness of trust once deterred its analysis, it has now led to a proliferation of definitions each used in a different context. For example, trust is defined by some as a characteristic of a particular class of action, while others identify it as a precondition for any such action to take place. At the same time, some discuss trust with reference to governments and organisations, while others examine trust between individuals or people in particular roles (Sztompka, 1995: 5-8).

No one definition of trust is correct, but some are more refined than others. Definitions of trust differ not only between disciplines, but also within disciplines. The focus of this paper is the conceptualisation of trust within economic thought. Socio-economic notions of trust, however, are included where it is felt that they have been instrumental in pushing the issue into the mainstream of economic discourse. The aim of this paper is to strike a balance between clarity of economic argument on the one hand, and its positioning within inter-disciplinary debate on the other. As Coleman (1988: 527) argues,

Elements of these two intellectual traditions [economics and sociology] cannot be brought together in a pastiche. It is necessary to begin with a conceptually coherent framework from one and introduce elements of the other without destroying that coherence.

Although this paper does not seek to address fully how trust is perceived in sociology, it is hoped that it will serve as a useful complement to discussions of trust rooted in more sociological contexts.

The second reason for analysing trust from an economic perspective stems from the recent popularisation of the argument that under certain conditions trust enhances economic efficiency.<sup>1</sup> Once again, this hypothesis can be traced back to the likes of J. S. Mill (1891: 68-69) who emphasises that trust can reduce the transaction costs of enforcing honest behaviour. Of equal, if not more concern, are situations in which an absence of trust causes inefficiency. As Arrow (1975: 24) asserts,

[V]irtually every commercial transaction has within itself an element of trust, certainly any transaction conducted over a period of time. It can be plausibly argued that much of economic backwardness in the world can be explained by the lack of mutual confidence.

The absence of trust may be particularly prevalent in many developing and transition economics in which economic transactions are viewed as exploitative rather than mutually beneficial. Decision-makers in these countries often face considerable uncertainty fuelled by rapidly changing economic conditions and political instability (Leff, 1986: 2-5; Sztompka, 1995). The answers to questions such as 'why does trust exist between some economic actors but not others?' and 'what can be done to foster trust in economic relations?' are precluded without a tangible notion of the concept involved.

Section I of this paper describes why the conceptualisation of trust is topical, yet problematic for the social sciences. Section II traces the emergence of trust as a critical issue in economic thought. Trust is outlined as a calculative concept in Section III. Section IV discusses the extent to which the conceptualisation of trust as a purely calculative process can be reconciled with the importance of social context in determining the costs and benefits presented by specific situations. Section V concludes.

## **II. The Emergence of Trust in Economic Thought**

The neoclassical paradigm of the perfectly competitive market obviates any space for the discussion of trust, effectively keeping the concept outside the domain of economics. Trust only enters the discursive arena through departure from neoclassical orthodoxy. Transaction cost theory provides an

---

<sup>1</sup> For example see Lorenz (1990); Menkoff (1991); and Sako (1992: 221-240).

alternative paradigm according to which exchange is not costless, but where it does occur, it is supported by the most economically efficient institutional arrangement. Recently, however, there has been much discussion concerning the limitations of transaction cost theory in explaining the behaviour of agents in situations of exchange. A new consensus is growing that emphasises the importance of the system of social relations underpinning many transactions. Attention has converged on one particular feature of social relations which it is argued, economises on the costs of exchange. This feature is commonly referred to as 'trust'. The section below outlines how recent developments in economic discourse have led to the emergence of trust as a critical issue in mainstream economic thought.

According to the neoclassical paradigm, the perfectly competitive market is populated by large numbers of anonymous, price-taking buyers and sellers who meet for an instant to exchange standardised goods. All actors in the perfectly competitive market seek to maximise their own welfare, but this does not include the pursuit of self-interest through acts of deceit or the withholding of relevant information. Even if agents are permitted to have a desire to behave dishonestly, the assumption of perfect competition includes that of perfect information, and so the opportunity for dishonest behaviour never arises. Under the assumptions of the perfectly competitive market, there is no need to trust (or distrust): the economic system is rendered transparent to all agents thereby negating trust as an issue (Platteau 1994: 540-1).

On the rare occasions that social relations between economic actors surface in neoclassical discourse, it is in the form of obstructions to the workings of the competitive market. This perception has its origins in classical economics and is evident in Adam Smith's well-known complaint that, 'people of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public or in some contrivance to raise prices' (Smith 1776: 232). In short, neoclassical economists have 'tied their own hands' when it comes to analysing the system of social ties (of which trust is one form) which underpin much economic activity in the real world (Hirschman 1982: 1473). Economists wishing to examine the concept of trust cannot do so from within the idealised model of perfect competition. Markets in the real world are not perfect and opportunities for dishonest behaviour abound. Without a framework that incorporates these features, the concept of trust remains outside the scope of economic analysis.

An alternative to the neoclassical interpretation of the co-ordination of economic activity is provided by transaction cost theory. This school of thought forms the cornerstone of 'new institutional economics' and is most notably associated with Williamson (1975; 1985). Unlike neoclassical economists, he does not treat transactions as costless, but recognises that costs are incurred in 'obtaining relevant information, the cost of bargaining and making decisions, and finally the costs of policing and enforcing contracts' (Hodgson, 1988: 180). In some cases the transaction costs may be

negligible, but in others they outweigh the benefits of exchange. If the latter condition applies, rational agents do not proceed with the transaction. Transaction costs can be explained in terms of the factors that give rise to them. These factors may be divided into two categories: human attributes and transaction-specific characteristics.

Williamson (1975; 1985) describes economic actors in terms of two key human attributes. The first is 'bounded rationality' - a cognitive assumption, according to which actors are intendedly rational but their capacity to receive, process, store and retrieve information is limited (Williamson, 1975: 21). The most important implication of bounded rationality is that it is costly for agents to contemplate and contract for all future contingencies that might arise over the course of a transaction. Indeed, the costs may be so great that agents fail to undertake the contemplation necessary to foresee some future contingencies. In addition, there may be contingencies that are unforeseeable (i.e. 'acts of nature') that within the context of Williamson's analysis may be regarded as having infinite transaction costs (Kreps, 1990: 744).

Market exchange provides ample opportunity for economic actors to pursue what Williamson (1975: 26) identifies as their second key attribute, 'opportunism':

Opportunism extends the conventional [neoclassical] assumption that economic agents are guided by considerations of self-interest to make allowance for *strategic* behaviour. This involves, self-interest seeking with guile...

If it will be to their own advantage, opportunistic agents will 'selectively reveal and distort information, even provide false information' to trading partners, if they can do so without later penalty (McGuinness, 1991: 68). The crucial point here is that not all agents are cynically assumed to behave opportunistically, but that it is costly to find out who will and who will not (Knorringa, 1994: 76). Thus, agents face a problem of how to keep opportunistic behaviour in check with information being imperfect and costly to acquire (Knorringa, 1992: 4).

Transactions differ in their complexity and their scope for opportunism. Consequently, it is the combination of the human attributes described above and nature of transaction-specific characteristics which together determine transaction costs. Williamson (1985) identifies three transaction-specific characteristics of importance. The first is asset-specificity - the extent to which transaction-specific investment in the form of money, time or energy, cannot be readily used in transactions with others (Grabher, 1993: 3). Second, is the degree of uncertainty involved in a particular transaction. Uncertainty arises as a consequence of both imperfect information and the opportunistic behaviour of others. Finally, the frequency of a transactions has a bearing on the relative costs of alternative institutional arrangements or 'governance structures' under which the transaction is undertaken:

When a transaction between two parties recurs frequently, the two parties can construct a special governance structure for the transaction, even if the governance structures are costly, since the cost of the structure can be amortised over many transactions. But when the transaction is a one-time-only transaction or recurs only infrequently, then it is generally more costly to put into place specialized mechanisms for this particular transaction...

Kreps (1990: 749)

Williamson's framework provides an explanation for the existence of alternative organisational arrangements in the economic system. He argues that boundedly rational, opportunistic agents seek to reduce their transaction costs and that it is this concern that determines the governance structure under which transactions occur. Economic transactions that require substantial transaction-specific investment, are of uncertain outcome, or recur frequently, are more likely to be conducted within hierarchically organised firms. Those that require no transaction-specific investment, are of more certain outcome, or are one-off in nature, are more likely to take place between firms in the marketplace (Granovetter, 1985: 493). In other words, the governance structure that is the most efficient in the reduction of transaction costs is adopted for duration of that transaction.

Transaction cost theory, however, is criticised on a number of points. It is argued that even in societies where legal sanctions are enforceable, the complexity of many transactions means that not all possible contingencies in the course of a transaction can be covered. The transaction costs of policing some contracts may in any case, be greater than the benefits from exchange (Lazerson 1988: 333). Consequently, in the absence of mechanisms other than contracts, a potential efficiency loss arises because agents forego mutually beneficial exchange (Leff 1986: 5). Moreover, what prevents the individuals who comprise the very components of the policing mechanisms from behaving opportunistically? As Arrow (1973: 24, quoted in Fiewal 1987: 594) asserts,

It is not adequate to argue that there are enforcement mechanisms, such as the police and the courts; these are themselves services that must be bought and sold, and it has to be asked why *they* will in fact do what they have contracted to do.

In the extreme, if a sufficient number of agents behave in an opportunistic manner, the organisation of exchange would operate at a very basic level:

If people were regularly to behave in an opportunistic fashion market exchange would become risky, rare, and confined largely to carefully structured face-to-face transactions in which exchanges of many goods and services would be made simultaneously, and with the close scrutiny of the quality of the goods and services provided and of security arrangements. In such circumstances, exchange would remain primitive, and market society as we know it would not exist.

(Moore, 1994: 819).

Of course many economies today have developed to a level of sophistication far in advance of primitive spot markets. Yet, as argued above, by itself transaction cost theory cannot explain fully all forms of the co-ordination of economic activity found in the real world. A new consensus is forming that at least some of the explanation has to do with the fact that a great deal of economic activity is underpinned by social relations.<sup>2</sup> It is suggested that dyadic relations as well broader social structures generate some degree of 'trust' between agents which lower the transaction costs of exchange. The existence of trust between parties permits transactions to occur without the rigidity and expense of hierarchical organisation, while at the same time minimising the risk from opportunistic behaviour in the marketplace.

Lorenz (1988: 198), for example, perceives trust as a distinct but complementary mechanism to the transaction costs of monitoring and enforcing sanctions:

I would like to suggest that the transaction cost literature...can tell us something about the role of trust in the economy....Trust enters into the argument because the presence of these [transaction] costs is directly linked to the possibility that economic actors will behave opportunistically....If transactions are thought of as friction in the economy, then trust can be seen as an extremely effective lubricant.

Similarly, Moore (1994: 819) argues that it is the generation of trust that makes agents who are 'incompletely protected by law willing to dispense with detailed personal policing of every transaction, and put great trust in exchanges involving large apparent risks'. If the costs involved in a potential transaction are prohibitive, the presence of trust between trading partners may reduce expenditure on policing mechanisms to a level where both parties would benefit from exchange. It is the notion that trust can lower the costs of exchange that has pushed the issue into the mainstream of economic debate.

This section has described how economic thought has moved away from neoclassical orthodoxy towards considerations of the co-ordination of economic activity in situations where markets are imperfect. It is argued that transaction cost theory alone cannot account for the co-ordination of all forms of economic activity since much economic exchange is embedded in social relations. While the idea that trust between trading partners acts to lower the cost of transaction has been introduced, what do economists mean by the term 'trust'? The following section addresses the conceptualisation of trust in economic discourse as an entirely calculative process.

---

<sup>2</sup> The seminal paper which prompted this consensus is by Granovetter (1985).

### III. Trust as a Calculative Process

As mentioned in the introduction, the definition of trust differs not only between disciplines but also within them. This section identifies elements common to definitions of trust within the domain of economics. Not all the definitions referred to here are put forward by economists, however. Some contributions from political scientists and sociologists are included where it is felt that they conceptualise trust in an 'economistic' manner, that is, as an entirely calculative process.

Associated with most definitions of trust is an element of risk. Dasgupta (1988: 51), for example, excludes from his analysis of **Trust as a Commodity** usages that do not involve risky action (such as 'I trust your family is well'). Luhmann (1988: 97) suggests that trust 'presupposes a situation of risk'. Risky action as a particular characteristic of trust, however, is most explicit in the work of Coleman (1990: 91) who describes situations of trust as 'a subclass of those involving risk'.

The risk involved in situations of trust derives from uncertainty regarding the future actions of others.<sup>3</sup> According to this view, trust is warranted when an agent expects the benefit from making him/herself vulnerable to another agent (whose behaviour is not under his/her control) to be positive, but not otherwise (Lorenz 1988: 197; Williamson 1993: 463). Through this calculus, an agent allows uncertainty regarding a future action of another agent to affect his/her own choice of action before he/she can monitor that action (Dasgupta, 1988: 51). If an agent decides to expose him/herself to the risk of opportunistic action by another agent, this is taken to imply the presence of trust (Coleman 1990: 105; Williamson 1993: 463).

It is only meaningful to trust someone to do **something**. Hardin (1993: 506) states, 'Trust is a three-part relation: A trusts B to do X'. It is possible for A to trust B to do X, but to distrust B in matters Y. In this sense, trust is always specific to a particular action of another and as such, is distinct from the more generic concept of trustworthiness. As Dasgupta (1988: 53-54) points out,

'Trustworthiness' concentrates on a person's overall disposition, his motivation, the extent to which he awards importance to his own honesty. Being able to trust a person to do what he said he would, on the other hand, requires us to know not only something of his disposition, but also something of the circumstances surrounding the occasion at hand. If the incentives are 'right', even a trustworthy person can be relied upon to be untrustworthy.

---

<sup>3</sup> It is important to distinguish between uncertainty arising from the strategic behaviour of others, and that arising from environmental uncertainty - so-called 'acts of nature'. Trust in the latter context is unrelated to agents' concerns regarding possible opportunism in transactions, but can provide an opening for opportunistic behaviour (Lorenz 1990: 4; Gambetta 1988: 218).

Trust arises only when agents have a choice. Agents on both sides of a transaction must be free to co-operate. If co-operation is enforced and asymmetrical, it is completely different from co-operation under conditions of trust (Knorringa, 1992: 5). As Lorenz (1990: 4; 1988: 197) argues,

[T]he action [exchange based on trust], and hence the risks of being abused, must be perceived as avoidable. Being able to avoid the relation is fundamental. If you could not, you might say something like: 'I have no choice but to trust this person, institution, etc.' Clearly, where there is no choice, we do not have to invoke trust to explain our behaviour.

Trust, then, relates to a particular future action of another and manifests itself in an agent's choice to expose him/herself to the risk and uncertainty associated with that action. Coleman (1990) incorporates these elements into his formalisation of the calculus that agents face when deciding whether to trust or not to trust:<sup>4</sup>

$$(1) \quad G \cdot p > (1-p) \cdot L$$

where,  $G$  is the potential gain if the other actor behaves honestly;  
 $p$  is the probability that the other actor behave honestly;  
 $L$  is the potential loss if the other actor behaves dishonestly.

An agent will enter into a transaction provided that the expected benefit if the other actor behaves honestly, is greater than the expected loss if he/she behaves dishonestly.<sup>5</sup> Equation (1) is adapted by Humphrey and Schmitz (1996: 5) to include the costs of monitoring and enforcing sanctions:

$$(2) \quad (G-S) \cdot p > (1-p) \cdot L$$

where,  $S$  represents the costs of safeguards and monitoring.

According to equation (2), an agent enters into a transaction only if the expected *net* gain conditional on the other agent behaving honestly, is greater than the expected loss if he/she behaves dishonestly. As argued in Section II, an element of trust in the expected behaviour of others is still necessary, since not all possible contingencies can be covered in a contract, nor sanctions in the event of non-fulfilment enforced.

The trust calculus espoused by Coleman (1990) is algebraically analogous to the definition expressed by Gambetta in his role as editor of the collection of papers, **Trust: Making and Breaking Cooperative Relations** (1988):

---

<sup>4</sup> Coleman's formalisation of the trust calculus is adapted from Coleman (1990) and Craswell (1993: 489-90).

<sup>5</sup> These gains and losses are calculated in terms of utility, not monetary values.

[T]here is a degree of convergence on the definition of trust which may be summarised as follows: trust...is a particular level of a subjective probability with which one agent assesses that another agent or group of agents will perform a particular action, both before he can monitor such action (or independently of his capacity to monitor it) and in a context in which it affects his own action...When we say we trust someone or that someone is trustworthy, we implicitly mean that the probability that he will perform an action that is beneficial or at least not detrimental to us is high enough for us to consider engaging in some form of cooperation with him. Correspondingly, when we say that someone is untrustworthy, we imply that that probability is low enough for us to refrain from doing so.

(Gambetta 1988: 217-18).

The conceptualisation of trust so far described in this section lends itself well to game theoretic explication. Game theory has been developed as a tool in order to analyse how perfectly rational people would behave in situations of interactive choice. As with any game, it is necessary to understand first the rules. Each game has a set of specific assumptions concerning the number of players, the strategies available to them, the scope for bargaining, the payoffs to all combinations of strategies, the information available to each player, and extent to which this is common knowledge (Lyons, 1992: 95-97). In the gaming context, as with the preceding definitions by Coleman (1990) and Gambetta (1988), trust is conceptualised as an *ex ante* probabilistic calculus by an agent with regard to a particular future action of another. The remainder of this section considers two games commonly used to elucidate the notion of trust.

The first of these two is a particular type of game known as the ‘Assurance Game’ (Platteau, 1994: 757-758; Dasgupta, 1988: 56-57). It takes the form of a two-person, one-shot game in which the payoffs are symmetrical for each player. Both players derive greater utility from mutual co-operation than from bilateral cheating. The utility obtained from all possible strategic combinations is shown in the payoff matrix of Figure 1.

**Figure 1: The Assurance Game**

		player 2	
		H	C
player 1	H	20, 20	5, 15
	C	15, 5	10, 10

Players have a choice between two strategies: to be honest (H), or to cheat (C). The choice of strategy is determined simultaneously. Each player prefers to transact honestly if the other behaves

honestly, but prefers to cheat if the other cheats. It is assumed that from a large group, pairwise individuals come together both at random and anonymously, to transact in each period. Thus, the optimal strategy for each player is determined by the probability with which he/she assesses that the other will behave honestly. With the particular payoff structure given in Figure 1, each player's optimal choice is *H* (to be honest) if he/she assesses the probability that the other player will behave honestly to be greater than 1/2, and is *C* (to cheat) otherwise.<sup>6</sup> Whether or not the more favourable joint payoff derived from both behaving honestly prevails, is dependent on the level of subjective probability with which players assess that each other will behave honestly.

Both Dasgupta (1988) and Platteau (1994) introduce further elements into the game in order to demonstrate which of the two equilibria (HH and CC) prevail under different assumptions. For the purposes of this paper, however, the important point is that the level of subjective probability with which a player assesses that another will behave honestly is taken as a measurement of trust (Platteau 1994: 758). Thus, once again trust is conceptualised as an *ex ante* probabilistic calculus by an agent with regard to the future action of another, which has an impact on the agent's own choice of action.

The Assurance Game provides a useful example of how trust is conceptualised in a one-shot situation, but what about trust in a dynamic context? As Dasgupta (1988: 59) argues,

For trust to be developed between individuals they must have repeated encounters, and they must have some memory of previous encounters. Moreover, for honesty to have a potency as a concept there must be some *cost* involved in honest behaviour. And finally, trust is linked with reputation, and reputation has to be acquired.

Dynamic considerations in the conceptualisation of trust can be most simply illustrated through a modification of the best known scenario in game theory, that of the Prisoner's Dilemma, into an infinitely repeated game. In its original form, the Prisoner's Dilemma is a two-player, one-shot game (played once only), in which each player has a choice between two strategies: not to confess and confess. The analytical framework originally referred to the behaviour of two suspects being interviewed separately about their involvement in a major and a minor crime. As both suspects are prevented from communicating with one another, there is no possibility of bargaining between parties. The payoff matrix for the one-shot game is shown in Figure 2.

---

<sup>6</sup> If the probability is denoted by  $p$ , then for each player to be indifferent between H and C,  $p$  must satisfy the equation  $20p + 5(1-p) = 15p + 10(1-p)$ , i.e. when  $p = 1/2$  (Platteau 1994: 804). Under the assumption that uncertainty regarding the future action of others has an impact on the decision of whether to trust someone or not, situations in which agents are indifferent to the future actions of others are excluded (Gambetta, 1988: 218).

**Figure 2: The Prisoner's Dilemma<sup>7</sup>**

		prisoner 2	
		not confess	confess
prisoner 1	not confess	1,1	10, 0.25
	confess	0.25, 10	8,8

If both suspects do not confess, there is enough evidence to convict each of the lesser crime for which they receive a short sentence. Neither can be convicted of the major crime without a confession from the other. If either suspect confesses while the other does not, the one who does not will receive a long sentence for the major crime, and the other a very short one for co-operating with the police. If both confess they each receive medium-term sentences. Each suspect ranks the four possible outcomes: it is best to confess and not be implicated by the other suspect (a very short sentence); next is not to confess and not to be implicated (a short sentence); then comes confessing and being implicated (a medium-term sentence); and worst of all is not confessing and being implicated by the other suspect (a long sentence). The suspects know the strategies and payoffs available to each other, but information is imperfect as decisions have to be made before the other's choice is revealed. The payoffs are such that most favourable joint outcome is when neither suspect confesses. Since whatever one suspect does the other is better off by confessing, however, both confess (do not trust) and subsequently receive medium-term sentences.

The one-shot Prisoner's Dilemma described above can be adapted to an infinitely repeated game in which the same two suspects are unfortunate enough to be placed in the same situation on an infinite number of occasions (or at least for a finite but indefinite number of periods). Now say, for example, that the rules of the game are altered such that before being separated for questioning one suspect announces to the other his/her intention to adopt a 'trigger' strategy. This strategy is defined by Lyons (1992: 115) as the decision to 'not to confess (co-operate) in the first round and all subsequent rounds unless the other player confesses, in which case confess in the next and all subsequent rounds'. If the other suspect believes this intention, he/she will adopt an identical strategy since the utility obtained from both parties not confessing over indefinitely repeated encounters, is greater than

---

<sup>7</sup> The payoffs are in years and for all combinations of strategies, the payoff to prisoner 1 is on the left, to prisoner 2 on the right.

that from confessing. Thus, neither party confesses in the first round and trusting behaviour is sustained in subsequent periods.<sup>8</sup> It is the degree of belief by one player regarding the other's stated intention to behave honestly which is labelled 'trust'.

In the repeated Prisoner's Dilemma it is the extent to which one player believes in the other's stated intention to behave honestly, that is taken as a measurement of trust. Despite different rules, the conceptualisation of trust in this dynamic context is essentially the same as that in the one-shot Assurance Game. The rules of the repeated Prisoner's Dilemma, however, allow players to have not only repeated encounters, but to have some memory of previous encounters. Moreover, the payoff structure of the repeated Prisoner's Dilemma is such that not confessing is costly to individuals in the short-run (rationally each suspect would choose to confess in a one-off encounter), but when encounters are indefinitely repeated the strategy of not confessing provides greater utility overall. Thus, the repeated Prisoner's Dilemma demonstrates the important links between trust, repeated encounters, experience, and reputation.<sup>9</sup>

The games described above show how trust is conceptualised as an entirely calculative process. As an analytical tool, however, game theory does have its limitations in that its rules are exogenous to central decision-making process. For example, all games start with the premise that the players have a certain disposition towards co-operation (trusting behaviour) relative to non co-operation (opportunistic behaviour). The inclusion of this assumption in the rules of the game renders irrelevant consideration of how players come to be predisposed in a particular manner in the first place. As Kreps (1990: 771-772) observes,

[A] weakness of game theory [is that it takes] institutions (the rules of the game) as given exogenously, without explaining where the institutions come from. We can view transaction cost economics as giving us a bit of a lead on this, insofar as we maintain the hypothesis that institutions will be created with a general view towards minimising transaction costs (or more precisely, with a view towards maximising the net benefits of transactions). But important pieces of transaction cost economics depend on the society in which the transaction is embedded - on the framework of laws and customs in particular....the solution to the problem of multiple equilibria...lies in the individual's strategic expectations and social and normative environment.

In real life, social relations, cultural factors, and experience shape the context in which people choose to trust (or not to trust). The next section discusses the extent to which this fact can be reconciled with the conceptualisation of trust as an entirely calculative process.

---

<sup>8</sup> It is assumed that payoffs in all periods are valued equally.

<sup>9</sup> For a more expansive treatment of the concept trust in dynamic games see Dasgupta (1988: 59-71) and Platteau (1994: 760-764).

#### IV. Trust in a Social Context

It is argued in Section II that the co-ordination of economic activity can be explained, at least in part, by degree of trust within social relations in which many market transactions are embedded. The 'social embeddedness' of market transactions refers to the fact that economic activity is affected by the dyadic social interactions of agents and the overall network of relations (Granovetter and Swedberg, 1992; Granovetter, 1985). The conceptualisation of trust as an entirely calculative process in Section III, however, treats the social context of decisions over whether or not to trust as given exogenously. This section considers the extent to which the conceptualisation of trust as a probabilistic calculus can be reconciled with the importance of social context in forming agents' preferences towards trusting behaviour.

The preceding section conceptualises trust as abstracted from social context through 'dyadic reductionism': on the one hand economic agents are treated as atomistic and undersocialised, but on the other, they are oversocialised by the assumption that exogenous social influences condition actors once-and-for-all. Dyadic reductionism therefore renders all ongoing social relations irrelevant (Grabher 1993: 4; Wrong 1961). Williamson (1993: 475), however, in his article **Calculativeness, Trust, and Economic Organization**, employs a calculative conceptualisation of trust but at the same time recognises that 'man, after all, is a 'social animal" and that 'socialization and social approvals and sanctions are also pertinent'. In order to accommodate social context in his calculative framework, he suggests that social factors are best regarded as part of the environment within which all transactions are conducted:

[T]he need for transaction-specific safeguards (governance) varies systematically with the institutional environment within which transactions are located. Changes in the condition of the environment are therefore factored in - by adjusting transaction-specific governance in cost effective ways. In effect, institutional environments that provide general purpose safeguards relieve the need for added transaction-specific supports. Accordingly, transactions that are viable in an institutional environment that provides strong safeguards may be nonviable in institutional environments that are weak - because it is not cost-effective for the parties to craft transaction-specific governance in the latter circumstances.

(Williamson, 1993: 476)

Williamson considers that different social contexts can alter the utility derived by economic agents from their choice of behaviour when faced with situations of exchange. *Ceteris paribus*, in some social contexts, the utility obtained from trusting behaviour in economic exchange is higher than that from cheating, in others the converse is true. In other words, it is recognised that it is the social context which forms the predispositions of agents towards trusting behaviour. Williamson (1993: 476) distinguishes between six types of contextual features which influence the co-ordination of economic activity: "[S]ocietal culture, politics, regulation, professionalization, networks, and corporate culture. Each can be thought of as institutional trust of a hyphenated kind: 'societal-trust,'

political-trust,' and so forth". In terms of explanation, however, the conceptualisation of trust put forward by Williamson does not extend beyond that described in Section II, except to set the probabilistic calculus in an undetermined social context which exhibits some form of 'hyphenated trust'.<sup>10</sup>

The conceptualisation of trust as an entirely calculative process is reliant on the behavioural assumptions of bounded rationality and opportunism as they were stated by Williamson in 1975 and summarised in Section II. Much of Williamson's work, as is well known, derives from that of Coase (1937) and Simon (1957). Coase argues that economic activity takes place within different governance structures in order to minimise the costs of transacting. Williamson (1993; 1985; 1975) synthesises this Coasian argument with his own interpretation of Simon's concept of bounded rationality and argues that agents seek to satisfy their objectives even though they are not fully informed. In other words, in situations of exchange agents act as 'satisficers' employing calculative rules of thumb to economise on the costs of information, rather than optimising (Hargreaves-Heap, 1992: 17-18; Hodgson, 1991: 198-199). It is on the basis of this 'economistic' interpretation of bounded rationality that the concept of trust has thus far been portrayed as a simple probabilistic calculus in which individual preferences are unchanged by the social environment.

In the wider social sciences, however, it is argued that 'satisficing behaviour' arises not only because of imperfect information, but also because bounded rationality renders some calculations too difficult for agents to perform even if the information is available. Furthermore, 'satisficing' amounts to more than just the cost-minimisation objective Williamson emphasises.<sup>11</sup> As Hodgson (1991: 199) notes,

[If] 'satisficing' was essentially a matter of minimising costs then it would amount to maximising behaviour of the orthodox type....Whilst Williamson recognises some of the informational problems, the fact that the cost calculus remains supreme in his theory means that he has not broken entirely from the orthodox assumption of maximisation.

Under this alternative view of bounded rationality, not all rules of thumb that agents follow can be reduced to a simple cost-minimising calculus because they might take the form of shared norms. From this more sociological perspective, consideration of the social environment as something more than just the sum of interacting cost-minimisers is a prerequisite to understanding the action of agents

---

<sup>10</sup> Williamson (1993: 479) argues that it is 'mind-boggling to contemplate the absence of calculativeness' in matters of trust except for the kind of trust found in close personal relations which, he advises should not be explained on a calculative basis.

<sup>11</sup> In its strict sense, satisficing behaviour derives from "a theory of the firm that postulates that firms typically not only seek to secure 'satisfactory' profits rather than maximum profits, as depicted in the traditional theory of the firm, but that other objectives such as increasing sales, market share of the size of the firm may be accorded equal or greater prominence than profits" (Pass and Lowes, 1993: 484).

in situations of exchange. Hargreaves-Heap (1992: 18) provides the following example to illustrate the point:

Consider the rule which says 'give way to traffic coming from the right' at a road intersection. When this rule is shared by all road users it will coordinate the traffic at road intersections perfectly. This is the sort of expected rule which, when widely followed, helps to generate the information about other people's behaviour that is crucial for instrumental calculations to proceed. However, the rule...is not a second best device which substitutes for scarce computing capacity. No amount of computational power will enable you to anticipate what other people will do when their behaviour depends on what they expect of you: the circle of expectations merely becomes wider the more you calculate, as they anticipate that you anticipate, and so on. In this form, the situation is characterised as uncertainty as opposed to risk,...and what transforms it into one where expectations can be formed is the existence of the shared rule.

The above example shows how boundedly rational agents can use a shared norm as a rule to reduce complexity. This interpretation of bounded rationality puts greater emphasis on people's limited computational capacity as opposed to Williamson's concern with the costs of information. The implication is that the conceptualisation of trust cannot always be abstracted from the social environment and reduced to a probabilistic calculus. Moreover, the conceptualisation of trusting (or non-trusting) behaviour as a shared norm provides an explanation for the behaviour of individuals in its own right.<sup>12</sup> By contrast, the conceptualisation of 'trust' as an entirely calculative process employs the term as a label to identify a particular class of behaviour for which an explanatory theory based on subjective probability is then invoked (Craswell, 1993: 488-489). Crucial to the latter notion of 'trust' is the fact that agents can attach a subjective probability to any event, and that every belief about a particular event that is relevant for decision-making, can be captured in the single dimension of 'probability'. This is the view of probability associated with Bayes on which the standard theory of rational choice is based (Sugden, 1992: 46).<sup>13</sup> Henceforth, to distinguish between the two alternative conceptualisations, trust as an entirely calculative process is referred to as 'Bayesian trust' and trust as a shared norm is referred to as 'non-Bayesian trust' since trust in the latter sense cannot be reduced to a particular level of subjective probability.

As both Bayesian and non-Bayesian trust are unobservable, it is possible only to examine *ex post* their influence on human behaviour *ex ante*. It might therefore be argued that it is not important whether agents act on Bayesian or non-Bayesian trust, so long as the outcomes of their actions are **as if** made on the basis of Bayesian trust. As Craswell (1993: 494-5) argues,

---

<sup>12</sup> For a discussion of trust as a non-calculative social norm see Luhmann (1988).

<sup>13</sup> See Cyert and DeGroot (1987) for a detailed analysis of why the Bayesian approach to decision-making is an effective methodology for handling uncertainty in economic theory.

Individuals may believe that they and others are acting for noncalculative reasons, but if their actions always turn out to be those that a calculative person would take, then the calculative theory provides a more parsimonious account of their behaviour, and the individuals' internal mental states can be disregarded. For Williamson's purposes, then, 'as if' is all that matters.

Even if the economic, calculative conceptualisation of trust does provide a more accurate description of the resulting behaviour, Williamson (1993) does not attempt to apply this rationale to explain the origins of the social norms (he describes as 'hyphenated-trust') which influence agents' preferences towards trusting behaviour. In his framework, norms of trusting embedded in the social environment remain undetermined: they could be either Bayesian or non-Bayesian. The conceptualisation of social norms of trusting as Bayesian is incomplete, however. As Hargreaves-Heap (1992: 15-17) elaborates,

[W]hat could be simpler than following the rule of investing in information up to the point at which the marginal benefit in terms of additional utility exactly matches the marginal cost in terms of utility forgone in other activities which might have been undertaken instead? However, this formulation begs the question. How is the individual to know what the marginal benefits of further information acquisition are without knowledge of the full information set?...Of course, it is always possible to reply that the individual has a subjective beliefs about these benefits. But this introduces an arbitrary element into the description of action, unless those subjective beliefs can themselves be justified as rational. And to reply that these beliefs are rational because the agent has acquired information on these beliefs, up to the point at which marginal beliefs equal marginal costs, merely begs the question again at one stage higher in the structure of beliefs. At each stage there will be a question about the rationality of belief, which can be answered in an instrumentally rational [calculative] fashion only though begging a question about the rationality of the beliefs about beliefs.

Thus, the basis of social norms of trusting cannot be satisfactorily explained through the invocation of Bayesian trust. Attempts to do so, only create an infinite spiral as the basis of beliefs about beliefs is questioned.

Alternatively, it can be argued that social norms of trusting are in fact non-Bayesian. From this perspective, the Bayesian and non-Bayesian conceptualisations of trust would be complements rather than substitutes. The non-Bayesian notion of trust explains the origins of shared norms of trusting that, with other factors, provide the background for the decision-making process associated with any particular action. Once agents' behavioural predispositions with regard to a particular action are established, the Bayesian notion of trust can then be used to explain how agents behave when faced with that choice of costs and benefits (Craswell, 1993: 496).

In summary, even if it is accepted that trust is best conceptualised as a particular level of subjective probability with which an agent *ex ante* assesses that another will perform a particular action, **the level of subjective probability at which the agent considers engaging in such action, varies according to social context. If social context, however, is to be incorporated into the**

**conceptualisation of trust, it is necessary to go beyond purely calculative notion of trust found in economics and consider the notion of trust as a social norm as is often treated in sociology. Social norms of trusting behaviour cannot always be reduced to a sum of agents who all interact on a purely calculative basis.<sup>14</sup>**

## **V. Conclusions: Marking the Disciplinary Boundaries**

Given the confusion surrounding the notion of trust across the social sciences, this paper has sought to clarify its conceptualisation within at least one discipline, that of economics. Trust has emerged as a critical issue in economic thought because of a growing consensus that transaction cost theory cannot by itself explain the behaviour of agents in situation of exchange. It is suggested that social relations underpin much economic exchange and it is the degree of trust between agents generated by these social relations that can lower the costs of transacting which result from bounded rationality and opportunism, to the point at which parties consider engaging in exchange.

Within economics trust is conceptualised as a particular level of subjective probability with which an agent *ex ante* assesses that another will perform a particular action which has implications for his/her own choice of behaviour. Trust in this sense is used as a label to identify a particular class of action for which a calculative explanation is then invoked. This conceptualisation does not explain how a certain level of trust is reached, but only that once reached it may be high enough to result in co-operative behaviour depending on the costs and benefits specific to any situation. These costs and benefits are treated as exogenous to the central decision-making process in which it is assumed that a subjective probability can be attached to the likelihood of any future event occurring, and that every belief about a particular event relevant for decision-making can be captured by the probability attached. In other words, this 'Bayesian' view of probability is used as a calculative rule of thumb by boundedly rational agents to economise on the costs of acquiring information.

'Bayesian trust' is reliant on the cost-minimising interpretation of bounded rationality common in economics. Not all rules of thumb, however, can be reduced to a simple cost-minimising calculus because even if relevant information is available the calculations may be too difficult to perform. This alternative view of bounded rationality frequently put forward by sociologists, gives scope for the conceptualisation of trust as a complexity-reducing social norm which, unlike Bayesian trust, can be put forward to explain the behaviour of individuals when faced with situations of exchange in its own right. Even if it is argued that the calculative, economic conceptualisation of trust provides a more parsimonious account of the outcome of human behaviour, Bayesian trust cannot be invoked to explain how the costs and benefits presented by specific situations of potential exchange differ with

---

<sup>14</sup> For a discussion of different bases of trust see Humphrey and Schmitz (1996).

social context. Non-Bayesian trust, however, can explain how agents come to be predisposed towards (or against) trusting behaviour with regard to any action. Once the social context of any particular exchange is taken into consideration as formed in part by norms of trusting, the calculative notion of trust can then be employed to explain how agents behave when presented with that choice of costs and benefits. In this way, Bayesian and non-Bayesian conceptualisations of trust can be complements as well as substitutes.

Without considering the social context in which particular transactions are set, the concept of 'trust' in economic thought is nothing more than a label used to identify a purely calculative theory of behaviour based on subjective probabilities. It is the consideration of the social context that marks the boundary between economic discourse and the domain of sociology. Whether or not crossing this boundary proves a fruitful exercise remains to be seen, but the issue of trust is sure to fuel interdisciplinary debate for a long time to come.

## References

- Arrow, K., 1975, 'Gifts and exchanges' in E.S. Phelps (ed), **Altruism, Morality and Economic Theory**, New York: Russel Sage Foundation
- Coase, R., 1937, 'The nature of the firm', **Economica**, Vol 4: 386-405
- Coleman, J.S., 1990, **Foundations of Social Theory**, Harvard University Press
- \_\_\_\_\_, 1988, 'Social capital in the creation of human capital', **American Journal of Sociology**, Vol 94: s95-s120
- Craswell, R., 1993, "On the uses of 'trust': comment on Williamson, 'calculativeness, trust , and economic organization", **Journal of Law and Economics**, Vol 36 No 2: 487-500
- Cyert R.M., and DeGroot, M.H., 1987, **Bayesian Analysis and Uncertainty in Economic Theory**, London: Chapman and Hall
- Dasgupta, P., 1988, 'Trust as a commodity' in D. Gambetta (ed), **Trust: Making and Breaking Cooperative Relations**, Oxford: Basil Blackwell: 49-72
- Dunn, J., 1988, 'Trust and political agency' in D. Gambetta (ed) **Trust: Making and Breaking Cooperative Relations**, Oxford: Basil Blackwell: 73-93
- Fiewal, G.R., 1987, **Arrow and the Foundations of the Theory of Economic Policy**, London: Macmillan
- Gambetta, D., 1988, **Trust: Making and Breaking Cooperative Relations**, Oxford: Basil Blackwell
- Grabher, G., 1993, **The Embedded Firm: On the Socioeconomics of Industrial Networks**, Routledge
- Granovetter, M., 1985, 'Economic action and social structure: the problem of embeddedness, **American Journal of Sociology**, Vol 91: 481-510
- Granovetter, M. and Swedberg, R., 1992, **The Sociology of Economic Life**, Oxford: Westview Press
- Hardin, R., 1993, 'The street-level epistemology of trust', **Politics and Society**, Vol 21 No 4: 505-30
- Hargreaves-Heap, S., M. Hollis, B. Lyons, R. Sugden, and A. Weale, 1994, **The Theory of Choice: A Critical Guide**, Oxford: Basil Blackwell
- Hart, K., 1988, 'Kinship, contract, and trust: the economic organization of migrants in an African city slum' in D. Gambetta (ed) **Trust: Making and Breaking Cooperative Relations**, Oxford: Basil Blackwell
- Hirschman, A.O., 1982, 'Rival interpretations of market society: civilising, destructive, or feeble ?', **Journal of Economic Literature**, Vol 24 No 4: 1463-84
- Hodgson, G.M., 1991, **After Marx and Sraffa**, London: Macmillan
- Hodgson, G.M., 1988, **Economics and Institutions**, Cambridge: Polity Press
- Humphrey, J. and H. Schmitz, 1996, 'Trust and economic development', mimeo, Institute of Development Studies, University of Sussex, January

- Knorringa, P., 1994, 'Lack of interaction between traders and producers in the Agra footwear cluster' in P.O. Pedersen, A. Sverrisson and M.P. van Dijk (eds), **Flexible Specialization**, London: Intermediate Technology
- Knorringa, P., 1992, 'Adaptive capacities in the Agra footwear industry', mimeo, Department of Development Economics, Free University of Amsterdam
- Kreps, D., 1990, **A Course in Microeconomic Theory**, London: Harvester Wheatsheaf.
- Lazerson, M., 1988, 'Organizational growth of small firms: an outcome of markets and hierarchies', **American Sociological Review**, Vol 53: 330-42
- Leff, N.H., 1986, 'Trust, envy, and the political economy of industrial development: economic groups in developing countries', **Columbia First Boston Series in Money, Economics and Finance, Working Paper** FB 86-38, November
- Lorenz, E.H., 1990, 'Trust, co-operation, and flexibility: international comparisons', **Working Paper** 149, The Helen Kellogg Institute for International Studies, University of Notre Dame
- Lorenz E.H., 1988, 'Neither friends nor strangers: informal networks of subcontracting in French industry' in D. Gambetta (ed), **Trust: Making and Breaking Cooperative Relations**, Oxford: Basil Blackwell: 194-210
- Luhmann, N., 1988, 'Familiarity, confidence, trust, problems and alternatives', in D. Gambetta (ed), **Trust: Making and Breaking Cooperative Relations**, Oxford: Basil Blackwell.
- Lyons, B., 1994, 'Game theory' in Hargreaves-Heap, S., M. Hollis, B. Lyons, R. Sugden, and A. Weale (eds), **The Theory of Choice: A Critical Guide**, Oxford: Basil Blackwell: 93-129
- McGuinness, T., 1991, 'Markets and managerial hierarchies' in G. Thompson, J. Frances, R. Levacic, and J Mitchell, **Markets, Hierarchies & Networks: The Coordination of Social Life**, London: Sage: 66-81
- Menkoff, T., 1992, 'Xingyang or how to trust trust ? Chinese non-contractual business relations and social structure: the Singapore case', **Internationales Asienforum**, Vol 23: 261-288
- Mill, J.S., 1891, **Principles of Political Economy**, London: Longmans
- Moore, M., 1994, 'How difficult is it to construct market relations ? A commentary on Platteau', **Journal of Development Studies** Vol 30 No 4: 818-30
- Pass, C. and Lowes, B., 1993, **Collins Dictionary of Economics** (Second Edition), Glasgow: Harper Collins
- Platteau, J.P., 1994, 'Behind the market stage where real societies exist - Part I: the role of public and private order institutions', **Journal of Development Studies**, Vol 30 No 3: 533-77
- \_\_\_\_\_, 1994, 'Behind the market stage where real societies exist - Part II: the role of moral norms', **Journal of Development Studies**, Vol 30 No 4: 753-817
- Sako, M., 1992, **Prices, Quality and Trust: Inter-firm Relations in Britain and Japan**, Cambridge: CUP.
- Shapiro, S., 1987, 'The social control of impersonal trust', **American Journal of Sociology**, Vol 93 No 3: 623-658

- Simon, H.A., 1957, **Models of Man: Social and Rational**, Chichester: Wiley
- Smith, A., 1776, **The Wealth of Nations**, reprinted in 1979, London: Penguin
- Sugden, R., 1994, 'How people choose', in S. Hargreaves-Heap, M. Hollis, B. Lyons, R. Sugden, and A. Weale, **The Theory of Choice: A Critical Guide**, Oxford: Basil Blackwell: 36-50
- Swedberg, R., 1987, 'Economic sociology: past and present', **Current Sociology**, Vol 35 No 1: 1-221.
- Sztompka, P., 1995, 'Trust: the missing resource of post-communist society', mimeo, Institute of Sociology, Jagiellonian University, Krakow
- Williams, B., 1988, 'Formal structures and social reality' in D. Gambetta (ed) **Trust: Making and Breaking Cooperative Relations**, Oxford: Basil Blackwell: 3-13
- Williamson, O., 1993, 'Calculativeness, trust, and economic organization', **Journal of Law and Economics**, Vol 36 No 2: 453-486
- \_\_\_\_\_, 1985, **The Economic Institutions of Capitalism. Firms, Markets, Relational Contracting**, New York: The Free Press
- \_\_\_\_\_, 1975, **Markets and Hierarchies: Analysis and Antitrust Implications**, New York: Free Press
- Wrong, D.H., 1961, 'The oversocialized conception of man in modern sociology', **American Sociological Review**, Vol 26 No 2: 183-92