Tilburg University

Adaptive build-up and breakdown of trust

Gorobets, A.; Nooteboom, B.

Publication date:
2005

CentER

Discussion Paper

ADAPTIVE BUILD-UP AND BREAKDOWN OF TRUST: AN AGENT BASED COMPUTATIONAL APPROACH

By Alexander Gorobets, Bart Nooteboom

This is a revised version of CentER Discussion Paper 2005-39 dated February 2005

Revised version dated April 2006

ISSN 0924-7815

Adaptive build-up and breakdown of trust: An agent based computational approach.[1]

Revised version, April 2006

Alexander Gorobets, Sevastopol National Technical University, Ukraine
Bart Nooteboom, Tilburg University, the Netherlands

Abstract

This article employs Agent-Based Computational Economics (ACE) to investigate whether, and under what conditions, trust is viable in markets. The emergence and breakdown of trust is modeled in a context of multiple buyers and suppliers. Agents develop trust in a partner as a function of observed loyalty. They select partners on the basis of their trust in the partner and potential profit, with adaptive weights. On the basis of realized profits, they adapt the weight they attach to trust relative to profitability, and their own trustworthiness, modeled as a threshold of defection. Trust and loyalty turn out to be viable under fairly general conditions.

JEL code: C63, L14, L24, L22, P13

Keywords: Agent-based computational economics, Inter-firm relations, Transaction costs, Governance, Trust, Complex adaptive systems.

1. Introduction

1.1 Research question

The viability of trust between firms in markets is a much-debated issue (for a survey, see Nooteboom, 2002). Economics, in particular transaction cost economics (TCE), doubts the viability of trust, on the argument that under competition, in markets, firms are under pressure to opportunistically utilize any opportunity for profit (Williamson, 1993). TCE proposes that people organize to reduce transaction costs, depending on conditions of uncertainty and specific investments, which yield switching costs and a resulting risk of hold-up.

In this article we employ TCE logic, but we also deviate from TCE in two fundamental respects. First, while TCE assumes that optimal forms of organization will arise, yielding maximum efficiency, we consider that problematic. The making and breaking of relations between multiple agents with adaptive knowledge and preferences may yield complexities and path-dependencies, with unpredictable patterns of making and breaking relations, that preclude the achievement of maximum efficiency. The methodology of Agent Based Computational Economics (ACE) is well suited to model complexities of multiple interactions, and to see to what extent theoretical benchmarks of maximum efficiency can in reality be achieved.

Second, while TCE assumes that reliable knowledge about loyalty or trustworthiness is impossible (Williamson, 1975), so that opportunism must be assumed, we expect that loyalty may be inferred from observed behaviour (Six, 2005). In contrast with Williamson (1993) we

[1] The authors wish to acknowledge the work on earlier versions of the model used in this article by Tomas B. Klos and Martin Helmhout.

propose that in markets loyalty may be viable, in yielding profit. While loyalty may carry opportunity costs it may also yield benefits of stable relationships that yield economies in learning by ongoing interaction. To investigate this, the methodology of ACE enables us to take a process approach to trust (Gulati, 1995; Zucker, 1986; Zand, 1972), by modeling the adaptation of trust and trustworthiness in the light of experience in interaction.

The analysis is conducted in the context of transaction relations between multiple buyers and suppliers, where buyers have the option to make rather than buy, which is the classical setting for the analysis of transaction costs.

In our model we include incremental improvement, in exploitation, as a result of ongoing collaboration, but we do not include radical innovation or exploration (March 1991, Nooteboom 2000). The reason for this is that we expect to contradict the proposition from TCE that trust is not viable in markets, and to make a strong case we bias the model in favour of TCE. TCE is most likely to be valid under static conditions. As Williamson (1999) acknowledged, TCE does not cover conditions of radical innovation. Under innovation, uncertainty is greater, hence contracts are more problematic, and the need for trust increases. Also, in static conditions competitive pressure of price competition is greater than in innovation, and the greater competitive pressure is, the less room there is for loyalty and trust. Hence, if we take a static context rather than one of innovation, the claim of TCE that trust is not viable gets the best shot.

We proceed as follows. First we discuss previous agent based models of trust, and we indicate how our model compares to them. Second we describe our model and indicate how it fits in the categorization of complex adaptive systems and agents proposed by Holland (1996). Third, we specify the experiments. Fourth, results are presented and discussed, first for some illustrative time paths of individual relationships and next for averages across agents (buyers, suppliers). We draw conclusions concerning the viability of trust in markets. Finally, we indicate shortcomings of our model and avenues for further research.

1.2. Other ACE models of trust

Many earlier attempts have been made at agent-based modeling of trust and related issues, to analyze emergent properties of complex interaction that would be hard or impossible to tackle analytically. The purpose of these models varies widely. Some study the effectiveness of sanctions and/or reputation mechanisms and agencies to support them, e.g. in information systems or supply chains (Zacharia et al., 1999; Meijer and Verwaart, 2005; Diekmann and Przepiorka, 2005), or in artificial societies (Younger, 2005). Some study self-organization, e.g. in the internalization of externalities in a common pool resource (Pahl-Wostl and Ebenhöh, 2004), the emergence of leadership in open-source communities (Muller, 2003), or the emergence of cooperative social action (Brichoux and Johnson, 2002). Others investigate the working of decision heuristics (Pahl-Wostl and Ebenhöh, 2004; Marsella et al., 2004).

The general set-up is that of multiple agents who can profit from each other but are uncertain about the quality or competence that is offered, sometimes allowing for multiple dimensions of quality, and dependencies between them (Maximilien and Singh, 2005). Other studies focus on the benevolence or intentions of agents, i.e. absence of cheating, in free-ridership, defection or expropriation of knowledge or other resources, and many look at both competence and intentions (Castelfranchi and Falcone, 1999; Pahl-Wostl and Ebenhöh, 2004; Breban, 2002; Muller, 2003; Gans et al., 2001). This is line with the distinction made in the trust literature between competence trust and intentional trust (see e.g. Nooteboom, 2002).

Mostly, agents are oriented only towards their self-interest, such as maximum profit, but some studies also allow for fairness and equity as objectives or dimensions of value (Pahl-Wostl and Ebenhöh, 2004). Mostly, trust is measured as a number between zero and one, and, following Gambetta (1988), is often interpreted as a subjective probability that goals will be achieved or no harm will be done. Mostly, conduct is individual, but sometimes allowance is made for coalitions (Breban, 2002).

Few studies of defection explicitly model both sides of the coin: the expectation of defection by others (trust) and one's own inclination to defect (trustworthiness). Also, most studies treat trust as of purely extrinsic value, in the achievement of profit, and do not include the possible intrinsic value of trust. Notable exceptions are Pahl-Wostl and Ebenhöh (2004) and Marsella et al. (2004).

Trust is generally updated on the basis of experience, sometimes only one's own experience in interaction, sometimes (also) on the basis of reputation mechanisms, sometimes with the services of some ´tracing agency´ (Zacharia et al., 1999; Meijer and Verwaart, 2005; Diekman and Przepiorka, 2005). Few studies are based on an explicit inference of competence or intentions, and even fewer studies explicitly model the decision heuristics used. Exceptions here also are Pahl-Wostl and Ebenhöh (2004) and, with great psychological sophistication, Marsella et al. (2004). A key question is whether agents have 'a theory of mind' on the basis of which they attribute competencies and intentions to others.

While most studies model trust as adaptive, in the sense that it develops as a function of private or public experience, there is very little study, as far as I know, of adaptiveness of the importance attached to trust relative to profit, and of the adaptiveness of one's own trustworthiness or inclination to defect.

Our model, to be described in the next section, shares different features with different models. It focuses on intentional trust (not competence trust), in terms of loyalty or defection, based on private experience (no reputation effects). Trust is adapted on the basis of observed defection, but only with simple reinforcement, without theory of mind and explicit decision heuristics. Next to trust it includes own trustworthiness, i.e. inclination to defect. Trustworthiness and the importance attached to trust are both adaptive, as a function of experience.

The central purpose of our model is theoretical: to investigate whether the claim of transaction cost economics that trust cannot survive under competition (Williamson, 1993) is correct. Under what conditions, if at all, are trust and trustworthiness viable in markets where the performance criterion is purely profit? The analysis is conducted in the context of transaction relations between multiple buyers and suppliers, which is the classical setting for the analysis of transaction costs.

1.3. Sketch of our model

We employ a model that is developed, with revisions and refinements, from an earlier model of Klos and Nooteboom (2001), and we conduct a series of new experiments. In this model, agents make and break transaction relations, i.e. are loyal or not, on the basis of preferences, with both trust and potential profit as dimensions of utility, with differential weights attached to them. Profit is a function of product differentiation, economy of scale from specialization, and learning by cooperation in ongoing relations. Economy of scale yields an incentive for buyers to go for a supplier who supplies to multiple buyers, which yields a bias towards opportunism, in breaking relations with smaller suppliers. However, as argued in TCE, there is a loss involved in such switching, for activities that are based on relation-specific investments. In the present model, such investments are needed for specialty products, which carry higher profit margins. The percentage of specialty products and corresponding specific investments is a parameter of the model that can be set. The specialty part, which is relation-specific, yields higher profit, and is also subject to learning by cooperation, as a function of an ongoing relationship. The latter advantage is lost when the relationship breaks, and thus yields switching costs, which yields an incentive for loyalty, in longer relationships that yield the advantage of learning by collaborating. In sum, the model incorporates the usual features of TCE: opportunism by defection, specific investments, economy of scale for non-specific investments, and switching costs. However, the model adds learning by cooperation, and it includes the possibility of trust as a dimension of utility, next to potential profit. Here, trustful relationships may also have intrinsic value, next to instrumental value in the form of assurance against opportunism, as has been widely recognized in the trust literature.

Agents are adaptive in three ways. First, in the preference function, the relative weights of potential profit and trust are adaptive, as a function of realized profit. In this way, agents can learn to attach more or less weight to trust, relative to potential profit. Second, agents adapt their trust in a partner as a function of his past experience (Brousseau 2000) concerning his loyalty, exhibited by his continuation of relation, i.e. absence of defection. As a partner refrains from defection, this is interpreted as trustworthiness, and trust increases incrementally, but with decreasing returns. Trust drops discontinuously when the expectation of loyalty is broken and defection occurs. This reflects the idea, often referred to in the trust literature, that 'trust comes on foot and leaves on horseback'. Third, agents also adapt their own trustworthiness, modeled as a threshold of exit from a relation, on the basis of realized profit. Thus, agents can learn to become more or less loyal.

An essential feature of the model is that the adaptation of both the weight attached to trust and the threshold of defection occurs on the basis of realized profit. This biases the model in favour of Williamson's (1993) claim that in order to survive in markets firms must make the highest possible profit, and cannot afford to accept less profit for the sake of trust. In the model, trust and trustworthiness can only emerge when they are favourable for realized profit.

The model allows us to explore under what conditions, in terms of parameter settings, trust and loyalty increase, or are stable, i.e. when they are conducive to profit, and hence viable in markets. If it turns out that trust is viable in that sense, we have a strong claim, and trust can play an even bigger role when realized profit is not the only driver of adaptation.

Starting values of agent-related parameters, such as trust, threshold of defection, and weight attached to trust, can be set for each agent separately. These reflect institutional conditions, such as prevailing ethics of conduct, performance of the legal system for reliable contracting, and intermediaries that are conducive to trust (Shapiro, 1987; Fukuyama, 1995; Putnam, 2000; Nooteboom, 2002). Other, non agent-related parameters, such as the percentage of product differentiation and specific assets, strength of economy of scale, strength of learning by cooperation, speed with which trust increases with duration of a relation, number of buyers, number of suppliers, and number of time steps in a run, are fixed per run, equally for all agents.

Holland (1996: 38) proposed a categorization of complex systems in terms of *properties* of aggregation, non-linearity, flows, and diversity, and *mechanisms* of 'tagging', by which systems or components identify themselves, as a basis for connections between them, internal models of the system's environment, and building blocks. In particular, he characterized adaptive agents in terms of (1) a *performance system* of detectors that receive messages from the environment, and effectors in response to them, (2) *credit assessment* by which rules get rewarded or reinforced for adaptation, and (3) *rule discovery*. In our model, concerning properties there is aggregation of production of multiple suppliers for a single buyer, non-linearity in economies of scale and learning by collaboration, and in feedbacks from experience to decision rules, flows of intermediate products and money between suppliers and buyers, and diversity of agent characteristics. Concerning processes, there are tags identifying suppliers and buyers. Tags of buyers indicate the extent of product differentiation, whether they are looking for potential suppliers, and their past loyalty. Tags of suppliers indicate the number of their current customers, and their past loyalty. Agents have internal models of the level and the value of the loyalty of their partners, their own loyalty (threshold of defection), and profitability as a function of product differentiation, economy of scale, and learning by collaboration. Building blocks are agents, preference formation, and matching between buyers and suppliers on the basis of preferences. In the performance system of actors, there are detectors of bids or requests for bidding, loyalty of partners, profitability, as a function of product differentiation, economy of scale and learning by interaction, and realized profit. There are effectors in inviting, making and accepting bids on the basis of preferences, which entail intentions to switch partners if the bid is made to others than the current partner, and in production and supply. There is credit assignment in adapting the weight attached to partner loyalty relative to profitability, and thresholds of own defection, as a function of realized profit. While in the model there is adaptation in decision rules, on the basis of credit assignment and reinforcement, as indicated, there is no discovery of new rules.

The model is specified in more detail in the next section.

2. Details of the model

2.1 Preference and matching

We focus on the risk that a partner will defect and thereby cause switching costs. In our model trust may be interpreted as a subjective probability that expectations will be fulfilled (Gambetta, 1988), which here entails realization of potential profit. Thus, expected profit (E) would be: E = profitability·trust. However, in the model, agents are allowed to attach more or less weight to trust relative to potential profit (α), on the basis of a generalized preference score:

$$\text{score}_{ij} = \text{profitability}_{ij}^{\alpha_i} \cdot \text{trust}_{ij}^{1-\alpha_i} \qquad (1)$$

where: score_{ij} is the score *i* assigns to *j*, $\text{profitability}_{ij}$ is the profit *i* can potentially make 'through' *j*, trust_{ij} is *i*'s trust in *j* and $\alpha_i \in [0, 1]$ is the weight *i* attaches to profitability relative to trust, i.e. the 'profit-elasticity' of the score. α is adaptive, as a function of realized profit. This functional specification, a 'Cobb-Douglas' function adopted from the literature on production functions, entails that profitability and trust are complements (they both contribute to the preference score) as well as substitutes (less profitability can be compensated with more trust).

At each time step, all buyers and suppliers establish a strict preference ranking over all their alternatives. Random draws are used to settle the ranking of alternatives with equal scores. The matching of partners is modeled as follows. On the basis of preferences buyers are assigned to suppliers or to themselves, respectively. When a buyer is assigned to himself this means that he makes rather than buys. In addition to a preference ranking, each agent has a ‘minimum tolerance level’ that determines which partners are acceptable. Each agent also has a quota for a maximum number of matches it can be involved in at any one time. A buyer’s minimum acceptance level of suppliers is the score that the buyer would attach to himself. Since it is reasonable that he completely trusts himself, trust is set at its maximum of 1, and the role of trust in the score is ignored: $\alpha = 1$. The algorithm used for matching is a modification of Tesfatsion's (1997) deferred choice and refusal (DCR) algorithm and it proceeds in a finite number of steps, as follows:

1. Each buyer sends a maximum of o_i requests to its most preferred, acceptable suppliers. Because the buyers typically have different preference rankings, the various suppliers will receive different numbers of requests.
2. Each supplier ‘provisionally accepts’ a maximum of a_j requests from its most preferred buyers and rejects the rest (if any).
3. Each buyer that was rejected in any step fills its quota o_i in the next step by sending requests to next most preferred, acceptable suppliers that it has not yet sent a request to.
4. Each supplier again provisionally accepts the requests from up to a maximum of a_j most preferred buyers from among newly received and previously provisionally accepted requests and rejects the rest. As long as one or more buyers have been rejected, the algorithm goes back to step 3.

The algorithm stops if no buyer sends a request that is rejected. All provisionally accepted requests are then definitely accepted.

It may, and often does, happen that suppliers have no buyer to supply to. The question then is whether in the absence of profit they should fail to survive. In the model, to maintain the number of buyers and suppliers, for reasons of experimentation, they do not disappear, and are available in the next round. Implicitly, the assumption therefore is that either they have alternative means of subsistence, or they do drop out, but are replaced at the same rate by new entrants that inherit their characteristics.

2.2 Trust and trustworthiness

Trust, taken as inferred absence of opportunism, is modelled as observed loyalty, i.e. observed absence of defection. Following Gulati (1995), we assume that trust increases with the duration of a relation. As a relation lasts longer, i.e. there is no defection, one starts to take the partner's behaviour for granted, and to assume the same behaviour (i.e. commitment, rather than breaking the relation) for the future. Thus, agent *i*'s trust in another agent *j* depends on what that trust was at the start of their current relationship and on the past duration of that relationship:

$$t_i^j = t_{\text{init},i}^j + (1 - t_{\text{init},i}^j)\left(1 - \frac{1}{fx + 1 - f}\right), \qquad (2)$$

where t_i^j = agent *i*'s trust in agent *j*,

$t_{\text{init},i}^j$ = agent *i*'s initial trust in agent *j*,

x = the past duration of the current relation between agents *i* and *j, and*

f = trustFactor.

This function is taken simply because it yields a curve that increases with decreasing returns, as a function of duration x, with 100% trust as the limit, and the speed of increase determined by the parameter f.

In addition, there is a base level of trust, which reflects an institutional feature of a society. It may be associated with the expected proportion of non-opportunistic people, or as some standard of elementary loyalty that is assumed to prevail. If an agent *j*, involved in a relation with an agent *i,* breaks their relation, then this is interpreted as opportunistic behaviour and *i*'s trust in *j* decreases; in effect, *i*'s trust drops by a percentage of the distance between the current level and the base level of trust; it stays there as *i*'s new initial trust in *j*, $t^{j}_{\mathrm{init},i}$ until the next time *i* and *j* are matched, after which it starts to increase again for as long as the relation lasts without interruption.

The other side of the coin is, of course, one's own trustworthiness. This is modelled as a threshold τ for defection. One defects only if the advantage over one's current partner exceeds that threshold. It reflects that trustworthiness has its limits, and that trust should recognize this and not become blind (Pettit, 1995; Nooteboom, 2002). The threshold is adaptive, as a function of realized profit.

2.3 Costs and profits

An important assumption is that profit generated in any buyer-supplier relation is equally shared. In other words, the model does not include hold-up as a function of possible asymmetric dependence. Thus, we allow for defection but not for the threat of defection with the purpose of increasing one's share in jointly produced added value.

Profit has the following elements. First, buyers may increase returns by selling more differentiated products. Second, suppliers may reduce costs by generating production efficiencies. There are two sources of production efficiency: economy of scale from a supplier producing for multiple buyers, and learning by cooperation in ongoing buyer-supplier relations. Economy of scale can be reaped only in production of standardized products, with general-purpose assets, and learning by cooperation can only de achieved in production that is specific for a given buyer, with buyer-specific assets.

This yields a link with the fundamental concept, in TCE, of 'transaction specific investments'. We assume a connection between the differentiation of a buyer's product and the specificity of the assets required to produce it. In fact, we assume that the percentage of specific products is equal to the percentage of specific assets. This is expressed in a variable $d_i \in [0, 1]$. It determines both the profit the buyer will make when selling his products and the degree to which assets are specific, which determines opportunities for economy of scale and learning by cooperation. This parameter is part of the 'state of the world', in this case the market, and applies to all agents, in a given run of the model.

Economy of scale is achieved when a supplier produces for multiple buyers. To the extent that assets are specific, for differentiated products, they cannot be used for production for other buyers. To the extent that products are general purpose, i.e. production is not differentiated, assets can be switched to produce for other buyers. In sum, economy of scale, in production for multiple buyers, can only be achieved for the non-differentiated, non-specific part of production, and economy by learning by cooperation can only be achieved for the other, specific part.

Both the scale and learning effects are modelled as follows:

$$y = \max\left(0, 1 - \frac{1}{fx + 1 - f}\right) \qquad (3)$$

where:
for the scale effect, f= sf = scaleFactor, x is general-purpose assets of supplier j summed over all his buyers. Here, y denotes scale efficiency achieved by supplier j.
for the learning effect, f = lf = learnFactor; x is the number of consecutive matches between supplier j and buyer i. Here, y denotes efficiency achieved in the collaboration between supplier i and buyer j.

Formula (3) expresses decreasing returns for both scale and experience effects. The scale effect is specified in such a way that a supplier can be more scale-efficient than a buyer producing for himself only if the scale at which he produces is larger than the maximum scale at which a buyer might produce for himself. For the learning effect, a supplier's buyer-specific efficiency is 0 in their first transaction, and only starts to increase if the number of transactions is larger than 1. If a relation breaks, the supplier's efficiency due to his experience with the buyer drops to zero. The resulting specification of profit is specified in Appendix A.

2.3 Adaptation

An agent is adaptive if 'the actions of the agent in its environment can be assigned a value (performance, utility, payoff, fitness, or the like); and the agent behaves in such a way as to improve this value over time' (Holland and Miller,1991, p. 365). This entails 'credit assignment' (Holland, 1996): certain decision rules are reinforced or weakened according to the (dis)credit they are given for success (failure). Here, success is measured by realised profit, which is attributed to preference formation (see formula 1). More specifically, agents adapt the values for $\alpha \in [0, 1]$ (weight attached to profit relative to trust) and $\tau [0, 0.5]$ (threshold of defection) from one time step to the next, which may lead to changes in the scores they assign to different agents. There is some randomness in the adaptation of α and τ, thus allowing for error and experimentation, according to the following procedure.

At each time step, each agent assigns a 'strength' to each possible value of α and τ. This expresses the agent's confidence in the success of using that particular value. The various strengths always add up to constants C_α and C_τ, respectively. At the start of each timestep, the selection of values for α and τ is stochastic, with selection probabilities equal to relative strengths, i.e. strengths divided by C_α and C_τ, respectively. The strengths of the values that were chosen for α and τ at the start of a particular timestep are updated at the end of that timestep, on the basis of the agent's performance during that timestep, in terms of realized profit: the agent adds the profit obtained during the timestep to the strengths of the values that were used for α or τ. After this, all strengths are renormalized to sum to C_α and C_τ again (Arthur, 1993). The idea is that the strength of values that have led to high performance (profit) increases, yielding a higher probability that those values will be selected again. This is a simple model of 'reinforcement learning' (Arthur, 1991; Arthur, 1993; Kirman and Vriend, 2000; Lane, 1993).

2.4 The algorithm

The algorithm of the simulation is presented by the flowchart in Figure 1. This figure shows how the main loop is executed in a sequence of discrete time steps, called a 'run'. Each simulation may be repeated several times as multiple runs, to even out the influence of random draws in the adaptation process. At the beginning of a simulation, starting values are set for certain model parameters. The user is prompted to supply the number of buyers and suppliers, as well as the number of runs, and the number of timesteps in each run. At the start of each run, all agents are initialized, e.g. with starting values for trust, and selection probabilities for α and τ. In each timestep, before the matching, each agent chooses values for α and τ, calculates scores and sets preferences. Then the matching algorithm is applied. In the matching, agents may start a relation, continue a relation and break a relation. A relation is broken if, during the matching, a buyer does not send any more requests to the supplier, or he does, but the supplier rejects them.

Figure 1 about here

After the matching, suppliers that are matched to buyers produce and deliver for their buyers, while suppliers that are not matched do nothing. Buyers that are not matched to suppliers produce for themselves ('self-matched', in 'make' rather than 'buy'). Afterward, all buyers sell their products on the final-goods market. Profit is shared equally with their supplier, if they have one. Finally, all agents use that profit to update their preference rankings (via α and τ), used as input for the matching algorithm in the next timestep. Across timesteps realized profits are accumulated for all buyers and suppliers, and all the relevant parameters are tracked.

Note that it is conceivable, given the logic of matching, that a supplier breaks with a buyer in his aim to go for a more attractive one, then lose the bidding for that buyer, and be left empty-handed. Then, it would be more reasonable for the supplier to first verify his goal attainment before breaking his existing relationship. However, in a large set of simulations, across a wide area of parameter space, this happened only once, at a very high level of opportunism, and it may not be unrealistic that sometimes such error is made, in an over-eagerness to switch to a more attractive partner.

3. Experiments

3.1 Expectations

According to TCE efficient outcomes are attained. Here, we expect:

Proposition 1: In interaction between multiple agents efficient outcomes are seldom obtained

According to TCE, high asset specificity leads to more make rather than buy. We expect the same result in our extended framework, according to the same argumentation, as was also found also in the initial experiments by Klos and Nooteboom (2001). When trust is low, specific assets yield hazards due to switching costs, and with few generalized assets there is little economy of scale to be achieved in outsourcing to producers who supply to multiple buyers. However, going beyond the initial experiments in Klos and Nooteboom (2001), we investigate the possibility of high trust allowing for make even at high levels of specificity.

Proposition 2: When trust is low, higher asset specificity/differentiated products yields less outsourcing.

We further expect:

Proposition 3: The more trust, the more collaboration in 'buy', rather than 'make'.

Proposition 4: The lower the weight attached to profit relative to trust (α), the more collaboration (buy rather than make), and the more loyalty (less switching).

Proposition 5: The higher the threshold of defection (τ), the more collaboration (buy rather than make), and the more loyalty (less switching).

Counter to TCE we expect:

Proposition 6: Even in markets, where profit guides adaptation, high trust (low α; high τ) may be sustainable.

In particular, we expect:

Proposition7: Trust is viable when there is a strong effect of learning by cooperation relative to economy of scale, which yields relatively high switching costs, and favours loyalty, in ongoing relations.

Recall that if during the matching between buyers and suppliers a buyer decides to 'buy' rather than 'make', he can follow two different strategies. One is an opportunistic scale strategy, where the buyer seeks a profit increase on the basis of economy of scale, by trying to find a supplier who serves more than one buyer. This entails much switching and less emphasis on loyalty and trust. The other strategy is the learning by cooperation strategy, seeking an increase of profit in ongoing relations. This entails less switching and more emphasis on loyalty and trust. Thus, in manipulating the strength of the scale effect relative to the effect of learning by cooperation, we can bias the model towards opportunism or loyalty. This interacts with the degree of asset specificity/specialization, since economy of scale applies only to general-purpose assets, and learning by cooperation only to specific assets. Note that there is an overall bias towards the opportunistic scale strategy, in that economy of scale is immediate, thus yielding a more immediate return in profits, while learning by cooperation takes time to build up. Thus, we are again stacking the odds in favour of the TCE theory that we criticize. However, this does seem to be a realistic feature, supporting the intuition that trust is more viable in a long-term perspective.

3.2 Model parameters

Given the number of variables and the range of potential values for each, the total parameter space, with all possible combinations of values, is huge, and some limitation must be imposed to limit the number of simulations.

One limitation was imposed for technical reasons: Each buyer's offer quota (maximum number of suppliers used) was fixed at 1. A larger number would require the distribution of the total volume of the buyer's production across the multiple suppliers. In the present setting, having more suppliers makes little sense, since it reduces the corresponding volumes of production, lowering economy of scale, without any compensating advantage. The usual reason for a buyer to engage in 'multiple supply' is to improve bargaining position in the division of added value. However, in the present model there is no room for conflict over that. Division of jointly produced added value is assumed to be equal. In this model, opportunism lies entirely in defection to another partner, not in the expropriation of proceeds.

The present model incorporates some refinements and revisions of an earlier model published by Klos and Nooteboom (2001), but the main difference concerns the experiments conducted with it. In their article, Klos and Nooteboom focused on the basic logic and composition of the model, on the calibration of parameters across the full range of possible values, and the model was used:

'.. only for a first round of experiments, aimed at illustrating how this type of model might work and at testing elementary hypotheses to assess the plausibility of outcomes. As expected, and predicted by TCE, product differentiation favours 'make' relative to 'buy'. This is because the transaction cost of switching is larger and there is less potential for economy of scale in large volumes of production for multiple buyers by specialized suppliers. As also expected, when there is more 'make', due to high product differentiation, the weight attached to trust .. declines faster. The experiments also illustrate that profits are not always optimal, even after adaptation settles down in a stable outcome …. This reflects the path dependence that arises in the making and breaking of relations: paths to optimal results can get blocked.' (Klos and Nooteboom, 2001, pp. 523-524).

In other words, previously the model was not yet used to explore under what conditions trust would be viable, if adaptation is driven by realized profit, which is the central question in the present paper. Here, we conduct the following experiments. First, we investigate whether even at high levels of specificity/product differentiation there may be less outsourcing, if initial trust is high. Second, we explore the central question when high initial trust and trustworthiness are sustainable, and do not unravel, and when, if ever, low initial trust and trustworthiness increase.

In the present experiments, we do not aim to develop a representative map of the occurrence of loyalty or defection across the entire space of possible parameter values. That might be the purpose of another set of simulations. Here, we make use of some of the results of previous experiments to zoom in on parts of the space of parameter values that are relevant or salient for our present purpose. The main guiding principle for this was that we are especially interested in how certain key variables affect the prevalence of loyalty or defection. Those key variables are: the extent of product differentiation, trade-off between economies of scale and economies of learning, and 'types of societies' in terms of starting values of parameters related to loyalty (weight attached to partner loyalty in the calculation of preference, and threshold of own loyalty). We excluded values for other parameters that ruled out, from the start, either loyalty or defection, for all values of the variables of interest. This yielded the following choices.

While the threshold of defection τ, determining the extent of loyalty, can conceivably rise up to 1, a maximum of 0.5 was set because earlier simulations showed that otherwise relations would get locked into initial situations, with little switching. Note that this lowering of the possible extent of loyalty biases the model in favour of opportunism, and hence against our claim that trust may be viable in markets, and thus favours TCE.

Each supplier's acceptance quota (maximum number of customers) was set to 3. In previous experiments with each supplier j's acceptance quota set to the total number of buyers, the system quickly settled in a state where all buyers buy from a single supplier. Then, economies of scale swamp advantages from economies of learning by collaboration. To give a chance to advantages from collaboration, the potential gains from economy of scale had to be constrained. This limits the validity of results to industries where a maximum of three customers is usual. This appears to apply to a range of industries, such as: cars, consumer electronics, computers, ships, telephones, large buildings and large infrastructural projects, etc.

The resulting list of parameters and variables used in the simulation is given in Table 1.

Table 1 about here

Each simulation run involves 12 buyers and 12 suppliers. Admittedly, this is fairly arbitrary, and may be varied in later simulations. With a maximum of 3 buyers for each supplier, this yields the possibility of 4 suppliers each supplying the maximum of 3 buyers, as the situation of maximum usage of economy of scale. Each run continues for 100 timesteps. This is chosen to yield convergence on reasonably stable outcomes. In order to reduce the influence of random draws, each run is repeated 25 times and results are averaged across all runs.

For the test of our propositions, we consider different values for the percentage of specific assets/differentiated products: d = 25, 45, and 65 %. We vary initial trust in the range 10, 50 and 90%, initial threshold for defection (τ) from 0 to 0.5, initial weight attached to profit relative to trust (α) from 0.0 to 1.0, and the fixed parameters of both the strength of economy of scale and learning by cooperation from 0.5 to 0.9.

First to give some feel of what is going on, we present and discuss time paths for selected individual relations, to illustrate the process of switching, or lack of it. These cases were selected not as representative, but as salient illustrations of what can happen in the simulation. Subsequently, we present the overall results, in comparison with our expectations, on the basis of averages, across runs as well as agents (all buyers, all suppliers).

4 Results

4.1 An individual trajectory for high trust and low α

Averages across agents and runs hide a great variety of trajectories for individual firms and relationships. Therefore, we first present some of those trajectories, to give a feel for what is going on. Since we are particularly interested in what happens at high trust, we select high trust cases. First, we take a case from a simulation with parameters: Initial Trust is 90%, the weight attached to profit relative to trust (α) is zero, the threshold of own defection (τ) is at its maximum of 0.5, the factors indicating the strength of learning and economy of scale (Learnfactor & Scalefactor) had an intermediate value of 0.5. In this case, which is biased towards trust and loyalty, agents are expected to favour loyalty and the learning by cooperation strategy.

Figure 2 about here

Figure 2 shows the actual profit for buyer 2. Recall that higher levels of specialization (d) yield higher profit margins. For the run with d=0.25 buyer 2 buys from supplier 6, for d=0.45 from supplier 3 and for d=0.65 from supplier 2. In his relation with supplier 6, for d = 0.25, his trust in that supplier increases up to almost the maximum of 1.0 (see Figure 3). Not shown is the result that the same applies to supplier's trust in buyer 2. For other values of d buyer 2 has no relation with supplier 6, so that his trust does not get any opportunity to rise. Supplier 6's profit is shown in Figure 4. For d = 0.25, he offers a big scale effect, producing for 3 buyers simultaneously: buyers 4 and 6 in addition to buyer 2. For d = 0.65, supplier 6 has 2 buyers, nr. 5 and 6, and he doesn't have any buyer for d=0.45. Not shown is the result that supplier 6's weight attached to profit relative to trust (α) and threshold of own defection (τ) remain about the same as their starting values. As shown in Figure 5, buyer 2 at first thinks he can increase profit by increasing the weight attached to profit relative to trust (α), but then learns that this does not work and reduces it again. Not shown is that his threshold of own defection (τ) remains at its initial value of 0.5.

Figure 3 about here

Figure 4 about here

Figure 5 about here

4.2 An individual trajectory for high trust and high α

Next, we take a case with the same parameter settings (e.g. high initial trust, maximum own loyalty), except that now the weight attached to profit relative to trust (α) switches from the lowest to the highest value: $\alpha = 1.0$.

Figure 6 about here

Figure 6 shows that buyer 5 makes rather than buys for average and high *d* during the full simulation period. For low *d* he makes until he finds supplier 1 who already produces for two other buyers, i.e. nr. 2 and 6, and then maintains a stable relation with supplier 1. It allows him to get a maximum attainable profit, from combining advantages of scale with those of learning by cooperation. As illustrated in Figure 7, he now learns to reduce the weight he attaches to profit (α), and to attach more weight to trust. Not included in the figures is the result that for both the buyer and supplier trust increases from 0.9 to approximately 1.0.

Figure 7 about here

Figure 8 shows how in this run supplier 1's profit increases as he builds up relationships with multiple buyers.

Figure 8 about here

Figure 9 shows that early on in the run supplier 1 lowered the value of his defection threshold τ. This reflects the fact that suppliers would rather have a buyer with more highly differentiated products (higher *d* than the current low value of 0.25), with higher profit margins and more potential for economies of cooperation. However, subsequently supplier 1 learned to raise the threshold τ again, tending to return to its original high of 0.5. Evidently, the lower τ did not yield the increase of profit he hoped for.

Figure 9 about here

4.3 Overall results

Now we turn to the more representative, overall results, in terms of averages across agents and runs. Earlier, we proposed that:

1. In interactions between multiple, adaptive agents, maximum efficiency is seldom achieved
2. In the absence of high levels of trust, outsourcing occurs only at low levels of asset specificity
3. High trust levels yield higher levels of outsourcing at all levels of asset specificity
4. Low weight attached to profit relative to trust yields more outsourcing and less switching
5. High threshold of defection yields more outsourcing and less switching
6. Under fairly general conditions, high trust levels are sustainable in markets
7. The choice between an opportunistic switching strategy and loyalty depends on the relative strength of scale effects and learning by cooperation

We present the results in the order of different starting values of trust, from high through intermediary to low trust, under the same conditions concerning the relative strength of scale effect and learning by cooperation, and under the same initial conditions concerning weight attached to trust and threshold of defection, all taken at intermediate levels of their possible range. The different starting levels of trust reflect different institutional settings, from high to low trust 'societies'. In particular, the question is whether high initial trust can be sustained, and whether perhaps distrust can evolve into trust. For a high level of initial trust we also investigate the effects of different starting values of weight attached to trust, threshold of defection, and relative strengths of economies of scale and learning by collaboration.

In the analysis, we consider maximum profit actually obtained, towards the end of a run, because of adaptation processes. We look at maximum profit relative to maximum attainable profit. This was done to enable comparison of effects of collaboration between different conditions of markets and technology (differentiated products, scale and learning effects) that affect maximum attainable profit in absolute terms. In other words, we compare the percentage of potential profit actually achieved. Detailed results are given in Appendix B. Here we discuss the main conclusions.

All propositions are borne out by the experiments, except Proposition 7, which turned out to be difficult to test unambiguously, due to interaction between different effects. We will return to this point later. Overall, relative profit declines more often than it increases, as we go from high to low trust. More precisely, under high trust maximum realized normalized profit is significantly higher than under low trust for intermediate and low levels of asset specificity (91 % and 80 % of potential profit vs. 87% and 63%), and marginally lower for high asset specificity (98% vs. 99% of potential profit). Overall, this confirms the central Proposition 6 that trust can well be viable in markets. Outsourcing is highest under highest initial trust. This is in agreement with Proposition 3. For high product differentiation/specific investments, yielding limited potential economy of scale, buyers employ suppliers only at high initial trust, and then have the maximum of 1 supplier per buyer. This is in agreement with Proposition 2. For details, see Tables B1, B3 and B4 in Appendix B.

While we had the general expectation that maximum efficiency would seldom be achieved, due to the complexities of interaction between multiple agents (Proposition 1), we did not specify why, precisely. A recurring effect that was not foreseen but makes sense is the following. When product differentiation/specific investments (*d*) is high, there is a high potential for learning by collaboration, but little potential for economy of scale from involving multiple suppliers per buyer. Achieving maximum benefit from scale requires a building up of multiple suppliers per buyer, which takes time, and maximum possible profit may not, and in fact never is, achieved, due to the competition for suppliers among buyers, who block each others' access to the maximum number. This effect is greater to the extent that the potential of employing multiple suppliers is greater, which arises when differentiation./specificity is lower. As a result, the percentage of maximum possible profit achieved is systematically lower for lower values of differentiation/specificity. This reflects the fact that in multiple interactions maximum efficiency is more or less difficult to achieve, which is neglected in TCE. This is in agreement with Proposition 1, and gives a specification of why, more precisely, maximum efficiency is not achieved. The difficulty of approaching optimal efficiency is also reflected in the fact that maximum actual profit is achieved only after many iterations.

Given the purpose of this article, an important outcome is that the high levels of initial trust are sustained, and in fact increase, on average, from 90% to the maximum of 100%, which confirms Proposition 6. At high levels of trust, outsourcing takes place even at high levels of asset specificity. At a high level of differentiation/specificity, at a high level of initial trust buyers' profit is about equal to profit under average or low initial trust. The advantage of realizing economy of scale by outsourcing, on the basis of higher trust, is limited at high level of specificity of investments. For details see Table B1 in Appendix B.

At the high level of initial trust, we tested for differential effects of different strengths of scale effects relative to learning by collaboration (see Proposition 7). Stronger learning by cooperation increases profit more for high than for low levels of specific investments, as is to be expected. A stronger scale effect, however, also increases maximum attainable profit, which is not realized, so that the ratio between actual and potential profit declines. With these mixed and partly unforeseen results, it is difficult to test Proposition 7 unambiguously. We did find that a higher initial value for weight attached to profit relative to trust yielded less outsourcing, which is in agreement with Proposition 4. We also found that a higher initial value of the threshold of defection yielded more outsourcing and higher relative profit, in agreement with Proposition 5. For details, see Table B2 in Appendix B.

Under medium or low trust, high product differentiation favours make relative to buy because the switching cost is larger and there is less potential for economy of scale. As a result, there is less outsourcing (Proposition 2). We find outsourcing only at the medium level of trust and at the lowest level of differentiation/specificity (see Table B3 in Appendix 2).

One of the surprises from the analysis, though perfectly understandable in retrospect, is the following paradox. Profit from learning by cooperation is highest for the highest level of product differentiation (high *d*), but when trust is low buyers then prefer to make rather than buy, and thereby forego the opportunities for learning by cooperation. Also, when buyers focus on profitability rather than trust (high α), profit from economy of scale is instantaneous while learning by cooperation is slow, and the potential for economy of scale is low at high levels of differentiation. Thus, under low trust and low weight attached to it, buyers lock themselves out from the advantages of collaboration. When they outsource, it is mostly at low levels of differentiation, when learning by cooperation yields only modest returns, and then they learn to appreciate its accumulation. They wind up in outsourcing at high differentiation only ‘by mistake’, then learn to appreciate it, and once learning by doing gets under way, a focus on profit keeps them in the relationship. In time, they may learn to attach more weight to loyalty and trust.

Another surprising but understandable outcome is that, in partial contradiction with Proposition 4, it is not always the case that a high weight attached to profitability relative to trust favours opportunism. Once a buyer begins to profit from learning by cooperation, an emphasis on profit may also lead to loyalty, in an ongoing relationship, after learning that loyalty can pay.

Overall, the results can be summarized as follows. A strong effect of learning by cooperation, a high weight attached to trust, and high loyalty favour the learning by cooperation strategy for high levels of specific investments, while a high weight attached to profit and high loyalty favour the scale strategy for low and average levels of specific investments.

5. Discussion

Of course, simulation is not equivalent to empirical testing. The test is virtual rather than real. Furthermore, we investigated only part of the total possible parameter space. We have only shown that under certain assumptions emergent properties of interaction satisfy expectations concerning the viability of trust in markets The significance of this depends on how reasonable the assumptions in the model and the restrictions on parameter values are considered to be.

While overall results conform to expectations, surprises arose in the details of the dynamics of interaction. The overall outcome is that both trust and opportunism can be profitable, but they go for different strategies. This suggests that there may be different societies, going for different strategies, of switching or of loyalty, which settle down in their own self-sustaining systems.

The experiments and the model have a number of shortcomings. In the experiments, we investigated only part of total parameter space, to zoom in on affects of variables that are of central theoretical interest. To arrive at more general conclusions, experiments would have to open up to unexplored areas of parameter space. That may be a topic for future research.

One shortcoming of the model is that learning from experience is myopic. One learns about the loyalty of partners only from own experience, not from their conduct in relationships with others. In other words, there is no reputation mechanism. That design of the model was deliberate, to make a strong case against the thesis from TCE that trust is not viable in markets. With a reputation mechanism, opportunistic defection would pay even less. However, apart from the debate with TCE, exploration of the effects of a reputation mechanism might be worthwhile.

Learning is also myopic in that adaptation by ‘credit assignment’ to parameters (weight attached to profit relative to trust, threshold of defection) is given only on the basis of one’s own realized profit. One might also learn from making inferences from profit observed from other agents. It was reasonable, in a first approach, to assume that one does not have information on both realized profit and relevant decision variables of other agents, to make inferences from correlations between them, but some learning from the success of others may need to be included.

For example, we found that under conditions of low trust, low value attached to trust and low loyalty, buyers learn to appreciate the value of ongoing collaboration, and hence of loyalty and of trust, only when they wind up in outsourcing ‘by mistake’. On the one hand, this reflects a well-known principle of asymmetry from the trust literature: with an excess of trust one learns from betrayals of trust, while with a deficiency of trust, which keeps one from engaging in collaboration, one robs oneself of the opportunity to learn that people may be trustworthy. On the other hand, it might be reasonable to assume that buyers include in their preference formation not only what they experienced but also the potential of collaboration that they did not experience but might either hear from others or observe from their conduct of collaboration and their success.

Related to this, while the decision rules in the model include adaptation on the basis of experience, in three respects, i.e. update of observed loyalty, and adjustment of weight attached to profit relative to trust and of threshold of defection, and includes random experimentation concerning the latter two, the model does not include discovery of new decision rules.

Finally, while the model includes some learning by doing, in collaboration, this learning is exploitative rather than explorative. The model does not cover buyer-supplier relationships for the purpose of innovation. This choice was also deliberate, to make a strong case against the thesis from TCE that trust is not viable in markets, since TCE is strongest under static conditions. But apart from the debate with TCE models of collaboration for innovation could be very worthwhile. However, the set-up of such a model would be very different. It goes beyond the scope of the present paper to discuss possibilities.

Appendix A Specification of profit

The number of general-purpose assets that a supplier *j* needs in order to produce for a buyer *i*, is equal to $(1-d_i)(1-e_{s,j})$, where $e_{s,j}$ is an efficiency factor ($0< e_{s,j} <1$) of scale (s) in the production volume of supplier *j*. The number of buyer-specific assets that a supplier *j* needs, to produce for a buyer *i*, is equal to $d_i(1-e^i_{l,j})$, where $e^i_{l,j}$ is an efficiency factor ($0< e^i_{l,j}<1$) of learning by cooperation (l) in the relationship between buyer *i* and supplier *j*. Thus, the profit that can potentially be made in a transaction between a buyer *i* and a supplier *j* is:

$$p_i^j + p_j^i = (1+d_i) - (d_i(1-e^i_{l,j}) + (1-d_i)(1-e_{s,j})). \qquad \text{(A1)}$$

The first part of the formula specifies returns and the second part specifies costs. It is assumed that the agents involved share the profit equally.

Appendix B Details of results

B.1 High initial trust

First, we consider an initial situation of high, 90% trust across all agents. First, we take intermediate initial expected values for α (0.5) and τ (0.25). Next to the variation of degree of specificity (*d* = 0.25, 0.45, 0.65), we vary the strength of economy of scale (scale factor *sf*) and learning by cooperation (*lf*), as follows:

- both scale and learning have intermediate strength (*lf* = *sf* = 0.5, see formula (3))
- high learning (*lf* = 0.9), medium scale (*sf* = 0.5). This is expected to favour a learning by cooperation strategy, with high loyalty
- medium learning by cooperation (*lf* = 0.5), high scale (*sf* = 0.9). This is expected to favour a scale strategy, with less loyalty.

The results are given in Table 1.

Table 1 about here

Table 1 supplies maximum *normalized* profit actually achieved in the course of time, obtained by dividing the buyers' maximum profits by the maximum attainable (theoretical) profit they can potentially make. This is the profit a buyer makes when he has an infinite relation with a supplier who produces for the maximum of 3 buyers. Usually maximum actual profit is achieved at the last steps of simulation because of adaptation processes in relations between buyers and suppliers. At the start point normalized profit turns out to be about 52 % for high d and 61 % for low *d*. Maximum scale effect is achieved when *d* is low. Here, the maximum arises in a situation where 12 buyers together buy from only 4 suppliers (each, i.e. one third of all suppliers each producing for the maximum of three buyers). Because the optimal network configuration, where suppliers produce for 3 buyers, rarely emerges, buyers organize closer to the optimum when *d* is higher. Then, buyers are less sensitive to the optimal configuration of network between agents, having less scope for increased efficiency by getting into an arrangement of one supplier producing for him as well as two other buyers.

For all levels of asset specificity (*d*), in each run at least one supplier produced for the maximum of 3 buyers, on average across runs 10 % of suppliers did this, 15% of suppliers produced for 2 buyers, 40 % for 1 buyer, and 35% for 0 buyers. The results indicate that in this high-trust society buyers follow the strategy of learning by cooperation and loyalty for all *d*, without switching between suppliers, even for the low value *d*=0.25, where only 25% of assets are subject to learning by cooperation.

Table 1 also shows the effect of different values for the strength of learning by cooperation (*lf*) and economy of scale (*sf*). Stronger learning by cooperation increases profit more for high than for low levels of specific investments, as is to be expected. A stronger scale effect, however, also increases maximum attainable profit, which is not realized, so that the ratio between actual and potential profit declines.

So far, we assumed intermediate levels for the initial weight attached to profit (α) and for the threshold of defection (τ). Now we analyze the effects of varying those values: $\alpha = 0.0$ and 1.0; $\tau = 0.0$ and 0.5. Learn and scale factors are fixed at the average level, i.e. 0.5. The results are given in Table 2. Here, we also supply the average number of suppliers per buyer, as an indicator of the extent of outsourcing.

Table 2 about here

When α=0, agents put their emphasis on trust and follow the strategy of learning by cooperation for all *d*. The distribution of suppliers between buyers in this case is the same as before (Table 2). Each buyer has ongoing transactions with the same supplier but when loyalty is equal to zero (τ=0) buyers sometimes break relations with suppliers for high d because then profit doesn't exceed the level of when they make. These buyers try to switch to other suppliers but they don't succeed because all agents are concentrated on trust built up in the past of their current relation. Opportunistic buyers then return to their initial partners and as a result they lose in profit slightly, for high *d*, because of switching costs. If loyalty is high (τ=0.5) there is no switching for any level of *d*, and agents try to generate as much profit as possible in stable relations by using the advantage of loyalty and trust, in learning by cooperation.

When α=1, agents focus on profitability rather than on trust, and buyers follow two strategies simultaneously: some of them buy from suppliers and others make themselves.

When τ=0.0 approximately half of buyers have suppliers for *d*=0.25 and these buyers follow the scale strategy, seeking a supplier who already serves two buyers, and trying to match with him. As a result, in this case 17% of all suppliers produce for three buyers. For *d*=0.45 and *d*=0.65 buyers prefer to make themselves, mostly because outsourcing is only preferred as relations with suppliers last longer and generate economies of learning, but this is unlikely to happen at zero loyalty. However, because of high initial trust buyers try to reach suppliers sometimes and then lose profit a little because of switching costs. If τ=0.5 the proportion of buyers who have suppliers increases for all d: 60 % of buyers have suppliers for *d*=0.25, 40% for *d*=0.45 and 30 % for *d*=0.65. However, the distribution of suppliers over buyers is different for all *d*. When *d*=0.25 approximately 20% of suppliers produce for three buyers and therefore profit is higher than in the case with τ=0.0. When *d*=0.45 about 12% of suppliers produce for three buyers and 5% of suppliers produce for one buyer and when *d*=0.65 suppliers produce only for one buyer and it is about 30% of them. Therefore, for low and average *d* more buyers follow the scale strategy because high loyalty allows them to keep stable relations with matched suppliers and generate higher profit than in the case with zero loyalty. For high *d* one part of buyers (70%) produce themselves and other part (30%) follow the strategy learning by cooperation because economies of learning are more important than scale effect.

B.2 Average and low initial trust

Now we turn to ‘societies’ with an average (50%) and lower level (10%) of initial trust. Learn and scale factors are again fixed at the average level, i.e. 0.5. The main outcome here is that buyers make for high and average levels of specific assets (*d*), and buy only for low levels. The results are specified in Tables 3 and 4.

Table 3 about here

At a medium level of initial trust, under low specific assets ($d = 0.25$) trust increases from an average to the highest level. This may seem surprising, since then the effect of learning by cooperation is lowest, so that the rewards of a trust strategy seem lowest. The explanation is as follows. Under average trust, suppliers are more attractive than buyers consider themselves to be only for low *d*, because potential losses in case of switching are smaller for low *d* than for high *d*. For high levels of specificity, buyers never enter into relations with outside suppliers, and thus never profit from collaboration and forego opportunities for the build-up of trust. For low specificity, the risk of outsourcing is less, and outsourcing occurs even if trust is not high. Then, advantages of learning by doing, even though limited by low *d*, set in, and advantages of loyalty are experienced, yielding an increase of trust.

Compared with the corresponding case in the high trust world (first column, Table 1), normalized profits are the same for high and low values of *d*, but lower for intermediate values. The network configuration of suppliers and buyers for low *d* is the same as in the case of high initial trust: 10% of suppliers produce for 3 buyers, 15% of suppliers produce for 2 buyers, 40 % for 1 buyer, 35% for 0 buyers. Buyers follow the learning by cooperation strategy in ongoing relations without switching.

In the case of low initial trust, see Table 4, buyers produce themselves (have no suppliers) even for a low level of specific assets. The result is a drop of normalized profits for low *d*, compared to the medium and high trust cases. All opportunities for learning by cooperation in collaboration are foregone.

Table 4 about here

References

Arthur, W. Brian: 1991, ‘Designing economic agents that act like human agents: A behavioural approach to bounded rationality’, American Economic Review 81/2, 353-359.

Arthur, W. Brian: 1992, ‘Designing economic agents that behave like human agents’, Journal of Evolutionary Economics 3/1, 1-22.

Breban, S. 2002. A coalition formation mechanism based on inter-agent trust relationships, AAMAS conference, July 15-19, Bologna.

Brichoux, D. and P.E. Johnson 2002. The power of commitment in cooperative social action, Journal of Artificial Societies and Social Simulation, 5/3. http://jass.soc.surrey.ac.uk/5/3/1.html

Brousseau, E. 2000. Confiance ou contrat, confiance et contrat, in: F Aubert et J.P. Sylvestre (eds.). Confiance et rationalité, INRA editions, Paris, pp. 65-80.

Castelfranchi, C. and R. Falcone 1999. Social trust: A cognitive approach, CAiSE conference, Heidelberg.

Diekmann A. and W. Przepiorka 2005. The evolution of trust and reputation: Results from simulation experiments, unpublished paper, Dept. of Humanities, Social and Political Sciences, Swiss Federal Institute of Technology, Zürich.

Fukuyama, F. 1995. Trust, the social virtues and the creation of prosperity, New York: Free Press.

Gambetta, D. (ed.): 1988, Trust: The Making and Breaking of Cooperative Relations, Basil Blackwell, Oxford.

Gans, G., M. Jarke, S. Kethers, G. Lakemeyer, L. Ellrich, C. Funken and M. Meister 2001. Towards (dis)trust-based simulations of agent networks, proceedings of the 4th workshop on deception, fraud and trust in agent societies, Montreal, 13-25.

Gulati, R.: 1995, ‘Does familiarity breed trust? The implications of repeated ties for contractual choice in alliances’, Academy of Management Journal 38/1, 85-112.

Holland, J.H. 1996[1995] Hidden order; How adaptation builds complexity, Cambridge US: Perseus Books.

Holland, J.H. and J.H. Miller: 1991, ‘Artificial adaptive agents in economic theory’, American Economic Review 81/2, 365-370.

Kirman, A.P. and N.J. Vriend: 2001, ‘Evolving market structure: An ACE model of price dispersion and loyalty’, Journal of Economic Dynamics and Control 25/3and4, 459-502.

Klos, T.B. and B. Nooteboom: 2001, ‘Agent based computational transaction cost Economics’, Journal of Economic Dynamics and Control 25, 503-526.

Lane, David A.:1993, 'Artificial worlds and economics, part II', Journal of Evolutionary Economics 3/3, 177-197.

Lewicki, R.J. and B.B. Bunker:1996, 'Developing and maintaining trust in work relationships', In R.M. Kramer and T.R. Tyler (eds), Trust in organizations: Frontiers of theory and research, Sage Publications, Thousand Oaks, pp. 114-139.

March, J.G. 1991. 'Exploration and exploitation in organizational learning', Organization Science, 2/1: 101-123.

Marsella, S.C., D.V. Pynadath and S.J. Read 2004. PsychSim: Agent-based modeling of social interactions and influence, proceedings International Conference on Cognitive Modelling, Mahwah, Earlbaum, 243-248.

Maximilien, E.M. and M.P. Singh. 2005. Agent-based trust model involving multiple qualities, AAMAS conference, July 25-29, Utrecht, Netherlands.

Meijer, S. and T. Vervaart 2005. Feasibility of multi-agent simulation for the trust and tracing game, in M. Ali and F. Esposito (eds.), Proceedings IEA/AIE, Heidelberg: Springer, 145-154.

Nooteboom, B.: 1999, Inter-firm alliances: Analysis and design, Routledge, London.

Nooteboom, B.: 2002, Trust: forms, foundations, functions, failures and figures, Edward Elgar, Cheltenham UK.

Pahl-Wostl, C. and E. Ebenhöh 2004. 'Heuristics to characterise human behaviour in agent based models', in C. Pahl-Wostl, S. Schmidt, A.E. Rizzoli and A.J. Jakeman (eds.), Complexity and integrated resources management, proceedings 2nd biennial conference IEMSS, 14-17 June, Osnabrück, 177-184.

Pettit, Ph.:1995, 'The virtual reality of homo economicus', The Monist 78/3, 308-329.

Putnam, R.D. 2000. Bowling alone; the collapse and revival of American community, New York: Simon and Schuster.

Reich, W. 2004. 'Reasoning about other agents: A plea for logic-based methods', Journal of Artificial Societies and Social Simulation, 7/4. http://jass.soc.surrey.ac.uk/7/4/4.html

Shapiro, S.P. 1987. 'The social control of impersonal trust', American Journal of Sociology, 93: 623-658.

Six, F.: 2005. The trouble about trust: The dynamics of interpersonal trust building, Cheltenham UK: Edward Elgar

Tesfatsion, L.S.:1997, 'A trade network game with endogenous partner selection', in H.M. Amman, B. Rustem and A.B. Whinston (eds), Computational approaches to economic problems, advances in computational economics, Vol. 6, Kluwer, Dordrecht, pp. 249-269.

Williamson, O.E.: 1975, Markets and Hierarchies: Analysis and Antitrust Implications, The Free Press, New York.

Williamson, O.E.: 1993, 'Calculativeness, trust, and economic organization', Journal of Law and Economics 36/1, 453-486.

Williamson, O.E. 1999. 'Strategy research: Governance and competence perspectives', Strategic Management Journal, 20: 1087-1108.

Younger, S. 2005. 'Reciprocity, sanctions, and the development of mutual obligation in egalitarian societies', Journal of Artificial Societies and Social Simulation, http://jass.soc.surrey.ac.uk/8/2/9.html

Zand, D.E.: 1972, 'Trust and managerial problem solving'. Administrative Science Quarterly 17/2, 227-239.

Zucker, L.G.: 1986, 'Production of trust: Institutional sources of economic structure, 1840-1920', in B.A. Staw and L.L. Cummings (eds), Research in Organizational Behavior, Vol. 8. JAI Press, Greenwich Conn., pp. 53-111.

Fig. 1. Flowchart of the simulation.

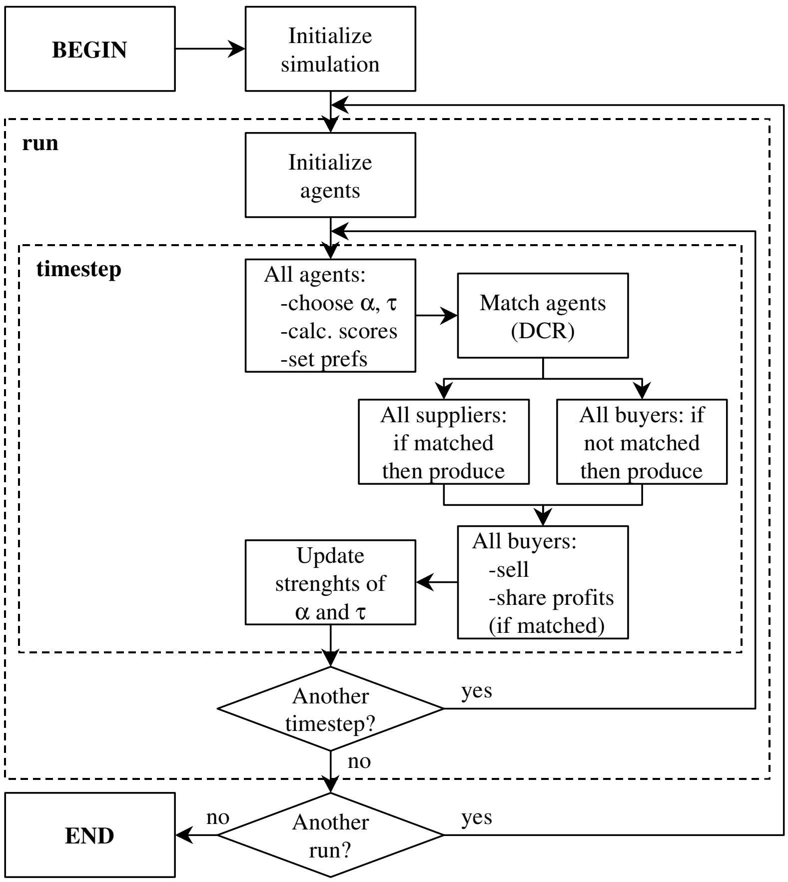

Fig. 2 Buyer 2's profit.

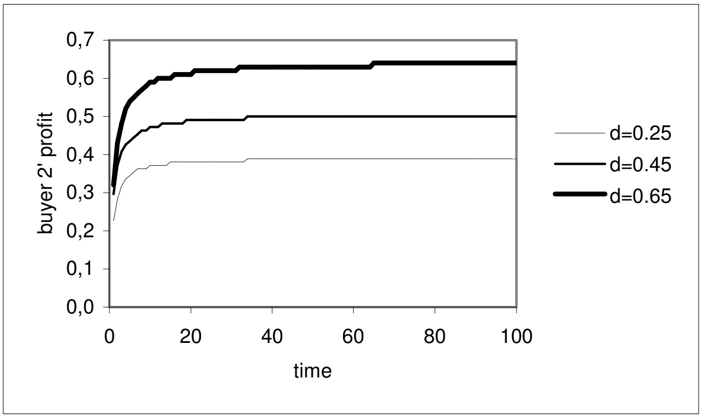

Fig.3 Buyer 2's trust in supplier 6.

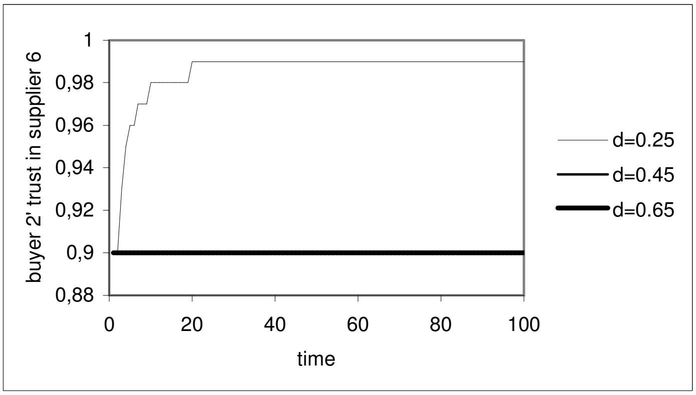

Fig.4 Supplier 6's profit.

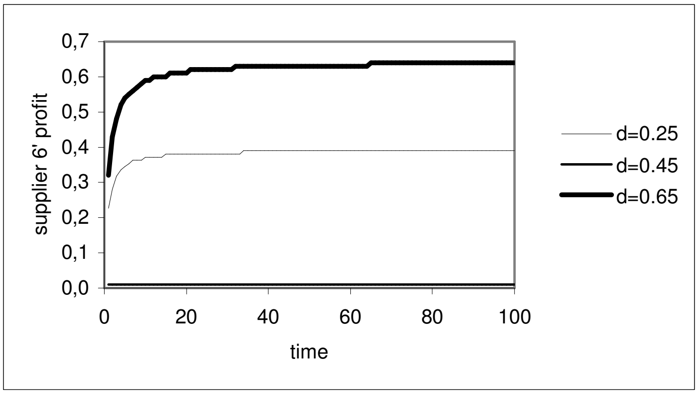

Fig.5 Buyer 2's α

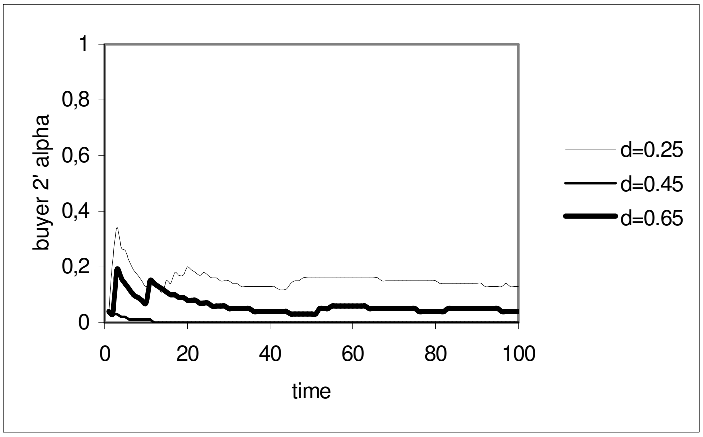

Fig.6 Buyer 5's profit.

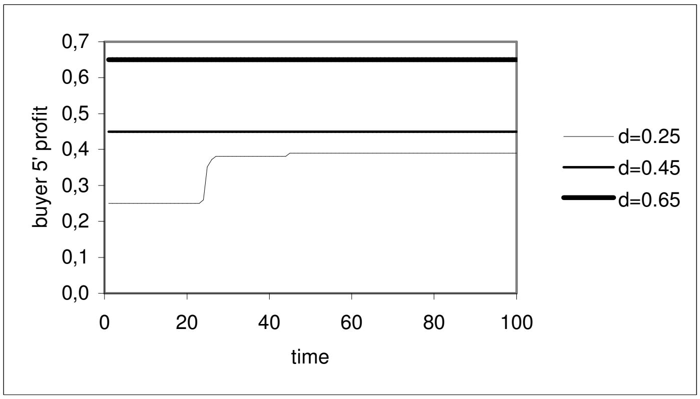

Fig.7 Buyer 5's α

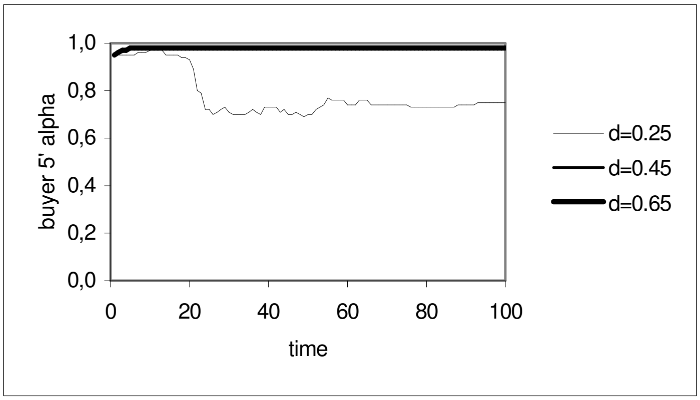

Fig.8. Supplier 1's profit.

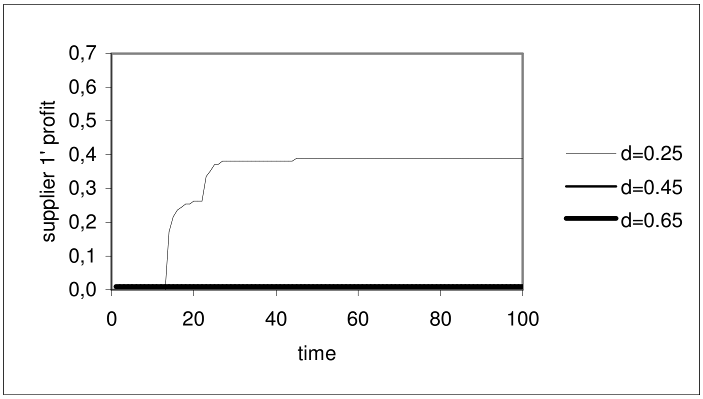

Fig.9. Supplier 1's τ

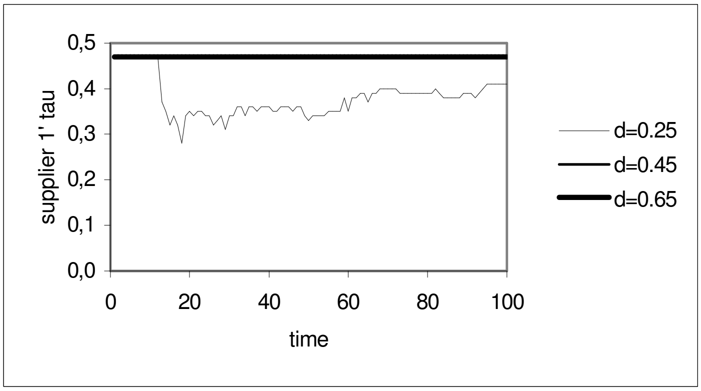

Table 1. Parameters and variables in the simulation.

	Param./var.	Value range	Value used
General	Number of buyers, I Number of suppliers, J Number of runs, R Number of timesteps, T Number of values for $\alpha \in [0, 1]$ Number of values for $\tau \in [0, 0.5]$ Base Trust, b Initial Trust, X Trust Factor	{1,2,…} {1,2,…} {1,2,…} {1,2,…} {2,3,…} {2,3,…} ⟨0,1] ⟨0,1] [0, 1]	12 12 25 100 5 5 0.3 0.9/0.5/0.1 0.5
Per buyer I	Differentiation, d_i Offer quota, o_i	[0, 1] {1,2,…, J}	{0.25, 0.45, 0.65} 1
Per supplier j	Acceptance quota, a_j Scale Factor, Y Learn Factor, Z	{1,2,…, I} [0, 1] [0, 1]	3 0.9/0.5 0.9/0.5

Table B1. Buyers' maximum normalized profits for high initial trust, at different learn and scale factors

D	#S.per buyer	Buyers max. normalized Profit			
		lf =0.5;sf=0.5	lf=0.9;sf=0.5	lf=0.5;sf=0.9	lf=0.9;sf=0.9
0.65	1	0.98	0.994	0.978	0.99
0.45	1	0.91	0.92	0.89	0.90
0.25	1	0.80	0.81	0.77	0.78

High initial trust is sustained, and in fact increases from 0.9 to the maximum of 1.0

Table B2. Buyers' maximum normalized profits for different α and τ

D	Buyers max. normalized profit & #suppliers per buyer							
	α=0.0; τ=0.0		α=0.0; τ=0.5		α=1.0; τ=0.0		α=1.0; τ=0.5	
0.65	0.96	1	0.99	1	0.96	0	0.99	0.3
0.45	0.91	1	0.91	1	0.85	0	0.92	0.4
0.25	0.80	1	0.80	1	0.82	0.5	0.84	0.6

Table B3. Buyers' maximum normalized profits for average initial trust

D	#Suppl. per buyer	Buyers max. normal. Profit
0.65	0	0.97 trust remains at 0.5
0.45	0	0.87 trust remains at 0.5
0.25	1	0.80 trust increases to 1.0

Table B4. Buyers' maximum normalized profits for low initial trust

D	#Suppl. per buyer	Buyers max. normal. Profit
0.65	0	0.99 trust remains at 0.1
0.45	0	0.87 trust remains at 0.1
0.25	0	0.63 trust remains at 0.1