Adaptive Learning in Evolving Task Allocation Networks

Tomas Klos
Delft University of Technology
P.O. Box 5031
2600 GA Delft
The Netherlands
t.b.klos@tudelft.nl

Bart Nooteboom
University of Tilburg
P.O. Box 90153
5000 LE Tilburg
The Netherlands
b.nooteboom@uvt.nl

ABSTRACT

In this paper, we study multi-agent economic systems using a recent approach to economic modeling called Agent-based Computational Economics (ACE): the application of the Complex Adaptive Systems (CAS) paradigm to economics. In this paper, we apply the CAS paradigm to the study of an industrial goods market, where firms need to decide between making and buying components.

Computer simulations using our model explain different emerging distributions of economic activity among organizational forms (market and hierarchy) in terms of the search problem facing the agents, and in terms of the negative consequences of the agents' search behavior on their perceived trustworthiness in the eyes of their potential partners. A further impediment to reaching optimal allocations we observe is that agents learn to protect themselves and their current allocation by being loyal and by focusing on their trust in their partner, rather than their partner's profit generating potential.

Categories and Subject Descriptors

I.2.11 [**ARTIFICIAL INTELLIGENCE**]: Distributed Artificial Intelligence—*Intelligent agents, Multiagent systems*; J.4 [**SOCIAL AND BEHAVIORAL SCIENCES**]: Economics

General Terms

Economics, Management, Experimentation

Keywords

complex adaptive systems, reinforcement learning, agent-based computational economics, transaction cost economics, trust, evolving trade networks

1. INTRODUCTION

We study task allocation in multi-agent systems. Task allocation has become a major research topic over the past years [22]. We are particularly interested in allocation of tasks among firms on industrial, interfirm markets. These are traditionally studied using transaction cost economics (TCE). However, as has been widely acknowledged, TCE in particular does not include dynamics of learning, adaptation and innovation in its analytical framework [13]. More generally, in the words of Holland and Miller, "economic analysis has largely avoided questions about the way in which economic agents make choices when confronted by a perpetually novel and evolving world" [10, p. 365]. Holland and Miller propose exploiting opportunities offered by recent advances in computer-based modeling techniques and machine learning, to counter the problems faced by standard tools and formal models. Their approach goes under the name of Artificial Adaptive Agents (AAAs) comprising Complex Adaptive Systems. The complex adaptive systems paradigm has by now been fruitfully applied for studying many different types of systems of interacting autonomous elements.

A Complex Adaptive System (CAS) [9] "is a complex system containing adaptive agents, networked so that the environment of each adaptive agent includes other agents in the system" [10, p. 365]. The application of this paradigm to economics has come to be known as Agent-based Comptutational Economics (ACE) [19, 4],[1] which is the computational study of economies modeled as evolving systems of autonomous interacting agents. In terms of ACE, "decentralized market economies are complex adaptive systems, consisting of large numbers of adaptive agents involved in parallel local interactions"[21, p. 55].

In the current paper, we employ an ACE perspective to the study of inter-agent relations on *intermediate goods markets*. As mentioned above, these are typically studied using Transaction Cost Economics (TCE) [2, 23]. TCE takes the 'transaction' as its basic unit of analysis, and analyzes which structural forms should be used for organizing such transactions, under different circumstances. The view of the *firm*, in [2], as an alternative means to the *market* for organizing transactions, rather than as a production function to be optimized, was a breakthrough, although it wasn't recognized as such until later.[2]

So TCE analyzes transactions [23, p. 1]: "A transaction occurs when a good or service is transferred across a technologically separable interface. One stage of activity terminates and another begins." If activities are thought of as nodes, and transactions as directed edges between nodes (specifying how the outputs of certain activities are inputs to others) then TCE is essentially concerned with the partitioning of nodes into subgroups (firms): edges between nodes within the same firm are organized using unified, hierarchical, firm governance, while edges between nodes in different firms are organized using market governance. Simplistically, TCE deals with the question of which nodes (and the transactions

[1]For a wide variety of materials related to ACE, see Leigh Tesfatsion's ACE website at `http://www.econ.iastate.edu/tesfatsi/ace.htm`.

[2]The "discovery and clarification of the significance of transaction costs and property rights for the institutional structure and functioning of the economy" eventually earned Ronald Coase the 1991 Nobel prize in Economic Sciences.

Cite as: Adaptive Learning in Evolving Task Allocation Networks, Tomas Klos, Bart Nooteboom, *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. 465–472

between them) should be organized *within firms* (the 'make' alternative), or *on the market*, i.e. across firm boundaries (the 'buy' alternative).

A number of criticisms of TCE have been raised [13, 16], similar to the more general criticisms of economics that spawned the emergence of the ACE approach. In general, TCE has been acknowledged to disregard the role of learning, adaptation and innnovation, including trust (see [7] for a more complete discussion). Furthermore, as Ronald Coase, the founding father of TCE admits [3], "[t]he analysis cannot be confined to what happens within a single firm. The costs of co-ordination within a firm and the level of transaction costs that it faces are affected by its ability to purchase inputs from other firms, and their ability to supply these inputs depends in part on their costs of co-ordination and the level of transaction costs that they face which are similarly affected by what these are in still other firms. What we are dealing with is a complex interrelated structure."

The CAS paradigm of course, is ideally suited for dealing with such a complex interrelated structure. Applying the CAS paradigm to the TCE domain in what we call Agent-based Computational Transaction Cost Economics (ACTCE), we let the distribution of economic activity across different organizational forms emerge from processes of interaction between autonomous boundedly rational agents, as they adapt future decisions to past experiences (cf. [4]).

An agent in a CAS is defined as being '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" [10, p. 365]. When viewing an economic domain as a complex adaptive system, the problem facing boundedly rational agents as they try to survive and make or even maximize their profit, is one of induction [1]. We therefore model this process using Q-learning (see [18]), a form of reinforcement learning, which itself is based on psychologist Edward Thorndike's Law of Effect. Note the similarity with the description of a reinforcement learning agent in [18, p. 7], whose "sole objective is to maximize the total reward it receives in the long run." These agents, buyers and suppliers on an industrial market for intermediate goods, autonomously decide to make and break connections with partners, while adaptively learning to trust, be loyal or opportunistic, emphasize profit or trust, etc.

The next section (2) describes our model of adaptive agents in complex interrelated multi-agent systems of inter-firm relations. Then, Section 3 presents an analysis of the parameter space of our model, explaining how various parameters influence each other, and where interesting regions of the parameter space are located. This analysis forms the input for our computational experiments, results from which are described in Section 4. Section 5 concludes the paper.

2. THE MODEL: MATCHING ADAPTIVE BUYERS AND SUPPLIERS

We model interactions between buyers and suppliers on an industrial market, i.e. a market for an intermediate good; a component they use to produce a final good which they sell on a final goods (consumer) market.[3] The buyers may buy the component from a supplier ('buy') or produce it for themselves ('make'). In any case, the buyers use the component to produce a final good which they sell to consumers on a final goods market.

[3] We will use the terms 'buyer' and 'supplier' for the agents on the industrial market, and the terms 'seller' and 'consumer' for the agents on the final goods market. A buyer on the industrial market is a seller on the final goods market.

Essentially, we model the market as a connected system of agents, an evolving multi-agent trade network [12]. Evolving, because connections are established and broken based on the preferences of the individual agents involved. In this context, we can not rely on economic theory's standard anonymized randomized matching device. A way of performing a matching based on heterogenous agents' idiosyncratic preferences is provided by the Deferred Choice and Refusal (DCR) algorithm [20], a modification of the Gale-Shapley Algorithm [5]. The DCR algorithm matches agents on two sides of a market to each other, based on the preferences each agent has over all agents on the other side of the market. In our model, the agents set preferences by calculating each other's 'scores:' expected payoffs obtained from transacting with each other (see Sect. 2.2). Agents adaptively search for connections which maximize their expected profit.

Buyers on the components market have the choice between making and buying. We interpret 'making' as the buyer supplying to himself. In the context of a matching algorithm, this means the buyer is matched to himself or to a supplier (see Figure 1). The

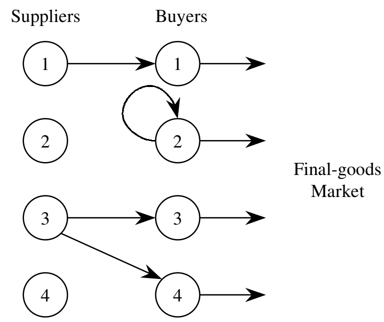


Figure 1: Buyers are assigned to suppliers or to themselves.

maximum number of suppliers (including himself) the buyer can be matched to at any one time, the buyer's *offer quotum* $q_o = 1$. The supplier has a maximum number of matches, her acceptance quotum $q_a \geq 1$.

This section proceeds as follows. First we describe how agents are matched to one another (section 2.1): this results in buyers either making or buying components. In the next section, we describe computer simulation experiments, where this matching occurs repeatedly, in each of a sequence of time steps. Matching is done using a matching algorithm, which produces a matching based on different agents' preferences for other agents. Section 2.2 shows how preferences are established based on 'scores,' expected values of profitability multiplied with trust, with an additional loyalty-threshold added to current partners' scores. Different weights are attached to profitability versus trust. Finally, Section 2.3 shows how the agents are *adaptive*: based on profits generated, they *learn* which weights to use for profitability versus trust and for their loyalty threshold.

2.1 Matching: Making and Breaking Edges

In terms of the DCR algorithm, our specification entails a further refinement: since buyers have the option of making rather than buying, a buyer does not *need* to be matched, while a supplier will rather be matched to a buyer than not be matched at all. This consideration is implemented by letting each buyer also calculate his own score for himself, and by letting all suppliers not scoring higher than his own score be 'unacceptable' for the buyer: he will rather be matched to (and make for) himself than buy from an unacceptable supplier. This effectively endogenizes the buyers' make-or-buy decision in the matching algorithm.

The matching algorithm assumes each agent calculates scores of all agents on the other side of the market (see Section 2.2). In addition, each buyer calculates his own score as a supplier to himself, and removes all suppliers not scoring higher than this own score from his list of acceptable suppliers. The matching algorithm then proceeds as follows.

1. Each buyer sends a maximum of q_o requests to the most preferred suppliers on his list of acceptable suppliers, if that list is long enough, and to the number of acceptable suppliers otherwise. Suppliers to whom requests have been sent, are removed from the list of acceptable suppliers.

2. Suppliers rank-order all requests received from buyers by those buyers' scores. They provisionally accept the requests from the most highly ranked q_a buyers, and reject the remainder, if any.

3. Each buyer who was rejected, tries to fill his quotum by sending additional requests to his most preferred remaining acceptable suppliers. Again, suppliers to whom requests are sent, are removed from the list of acceptable suppliers.

4. Each supplier who receives requests adds those to her list of provisionally accepted requests, rank-orders the total list, and provisionally accepts the top q_a buyers if she hasn't yet provisionally accepted them. The remainder, if any, is rejected.

The algorithm stops if no further requests are sent by buyers; all provisionally accepted requests are now accepted. One can easily see that this algorithm always terminates, and that it gives a matching which is optimal from the point of view of the agents who make the requests. In our application, it is plausible that the buyers are those agents.

2.2 Scores: Profitability, Trust and Loyalty

We assume that agents want to maximize their profit, but that they are boundedly rational and therefore unable to collect and process all information required for making optimal decisions—if establishing optimality is even tractable in the first place. TCE assumes that any suboptimal behavior (including trust, see [7]) is eradicated by market pressure, so now we can assess to what extent this is actually the case. Since the trade networks constantly evolves, and agents are continually confronted with a novel world, maximizing expected profit is the best they can hope to do.

Expected profit breaks down into two components. The profit that can potentially be made in a relation between two agents can be determined by them (see below). However, as also assumed by TCE [23], the agents' bounded rationality, and the fact that typically, transaction specific investments are put in place which make the agents dependent on each other, make for less than certain execution of contracts. In order to estimate expected profit, the agent needs a probability value that potential profit will actually be realized. For this, we assume that agents build up and maintain trust assessments on other agents. Trust has been shown to play a crucial role in interfirm relations [8, 14].

An agent j's score to another agent i is then the *potential profit* to be made in their mutual relation, multiplied with i's *trust* in agent j. Trust is then interpreted as the probability that agreements will be fulfilled (see Section 2.2.2), leading to a multiplicative specification. In order to allow agents to attach differential weights to trust versus profitability, we change the score calculation from a simple expected value calculation to a Cobb-Douglas functional form with constant returns to 'scale:'

$$\text{score}_{ij} = \text{potential profit}_{ij}^{\alpha_i} \cdot \text{trust}_{ij}^{1-\alpha_i} + p_{ij} \cdot \tau_i, \qquad (1)$$

where $\alpha_{\min} \leq \alpha_i \leq \alpha_{\max}$ is the weight agent i attaches to making a profit in a relation with agent j, relative to agent j's trustworthiness, and $\tau_{\min} \leq \tau_i \leq \tau_{\max}$ is agent i's *loyalty*. If agent j is agent i's current partner, then $p_{ij} = 1$ (otherwise $p_{ij} = 0$) and agent i adds his loyalty τ_i to agent j's score to express that other agents scores have to be at least τ_i higher than agent j's score in order for agent i to prefer them to agent j.

2.2.1 *Potential Profit*

Product Differentiation.

On the final goods market, products may be heterogenous or *differentiated*, meaning that the products of individual sellers are treated as being to some extent unique: different sellers' goods are imperfect substitutes for each other, giving sellers a degree of *market power*, i.e. the ability to set their price independent of their competitors and thus make a profit. We model this using an exogenous differentiation parameter, $0 \leq d \leq 1$, which is the same for all sellers on the market. Product differentiation is what buyers (on the industrial market—sellers on the final goods market) have to offer to their suppliers (on the industrial market): the way buyers contribute to the profit potential inherent in a relation between a buyer and a supplier. This consideration is implemented by allowing the buyer to sell their products on the final goods market at a price $1+d$: costs of production are ≤ 1, so d is the markup allowed to the buyers. A supplier has savings to offer: savings due to economies of scale and economies of learning. In general, 1 unit of assets is required to produce 1 unit of the good, but the supplier may economize on that. The buyer is able to *increase returns* because of his position on the final goods market, whereas the supplier is able to *reduce costs* (as compared to the buyer) because of his specialization as a supplier in terms of being able to generate economies of scale and of learning. (Details of our specification are available in a more detailed technical report to which we will add a reference should the current paper be accepted.)

Asset Specificity.

In this context, the most important of a transaction's characteristics (see Sect. 1) is the *specificity* of the assets invested in it. To the extent that assets are specific to a transaction, they can not be used for another transaction. Since a heterogenous, differentiated product implies that it is different from competitors' products, we assume equivalence of the differentiation of a product d and the specificity k of the assets required to produce it: $d = k$. Since a buyer's product's differentiation $d \in [0, 1]$, the assets required to produce that product are *specific* to its production to the extent of d, and *unspecific* ('general purpose') to the extent of $1 - d$. TCE predicts that as a transaction requires more specifically invested assets, choosing the market to organize it carries less advantages, as a supplier will more and more be producing exclusively for the buyer: increasing differentiation will thus be expected to lead to more making relative to buying.

If *unspecific* assets are invested in by a supplier, she may accumulate them in the production for multiple buyers, and attain economies in their increased scale. Economies of scale imply decreasing average costs per unit produced, as the number of products produced increases. On the other hand, continuous uninterrupted use of *specific* assets will, over time, generate economies due to learning-by-doing. In an inter-organizational context, firms learn from each other [17, 24], but it is reasonable to consider learning just in the deployment of assets which the supplier invests in *specifically* for the relation with the particular buyer—this is the $k = d$ fraction of the assets. Zollo et al. [25], for example, define 'interorganizational routines' as "stable patterns of interaction among two firms developed and refined in the course of repeated collaborations.".

2.2.2 *Trust, Trustworthiness and Loyalty*

Trust has been demonstrated to play an important role in interfirm relations [8, 14]. Trust is commonly interpreted as the probability that agreements will be fulfilled and that no harm will be done even though it can be [6, 11]. Following [8], we assume trust between partners to increase with longer sequences of transactions (cf. the notion of 'habituation' in [15]) and, conversely, an agent's trust in his partner to break down after the partner has broken off the sequence. Trustworthiness is then the absence of opportunism, another central concept in TCE. As a relation (sequence of transactions) lasts longer, without defection in the sense of the partner breaking off the relation, one starts to take the partner's behavior for granted, and to assume or expect the same behavior for the future.

For the precise model of trust we employ the following specification:

$$\text{trust}_i^j = \text{trust}_{init,i}^j + (1 - \text{trust}_{init,i}^j)\left(1 - \frac{1}{1 - f_t + f_t x}\right), \quad (2)$$

where trust_i^j is agent i's trust in agent j, $\text{trust}_{init,i}^j$ is agent i's initial trust in agent j, x is the past duration of the *current* relation between i and j, and f_t is a parameter allowing us to control for the strength of the trust effect. The baselevel of trust reflects the notion that one of the foundations of trust is basic, ex ante trust as an institutional feature of a society—a standard level of elementary decency that is assumed to prevail. On top of the baselevel, one can develop partner-specific trust on the basis of mutual experience. In case a relation is broken off by one of the partners j, the other agent i's trust in j decreases: it drops by half the distance between the current and the baselevel of trust. This new value becomes trust_{init}, and i's trust in j stays at that level until they are matched again. This reflects the notion that defection by a highly trusted partner is punished more severely than by an untrusted, or unknown partner in whom trust has yet to be established, and more generally, that trust (like reputation) is harder to establish than to break down.

2.3 Adaptively Learning α and τ

For agents' adaptive learning, we employ reinforcement learning (RL) methods [18], as explained in Sect. 1. In the RL paradigm, an agent's *policy* determines its actions given the state of the environment. In our model, the policy tells the agent, in each round of the simulation, which α- and τ-values to use—these values are the actions the agent can take. A *reward function* maps each state of the environment (or each combination of state and action, if actions are tailored to states) to a reward, indicating the desirability of that state. In our model, the reward is the profit the agent makes in a round of the simulation, depending on the agent's chosen action (α- and τ-values), and on the other agents' actions. Finally, a *Value function* tells the agent what the long run accumulated reward of α- and τ-values are.[4] Whereas rewards are immediate, they can be used to estimate the long-run Value of each action.

In our model, the agents' adaptive learning pertains to the values they use for α and τ. For both α and τ, a number of possible values is entertained by each agent. In each round of the simulation, the agents start by selecting a value to use for both α and τ, giving preference to values with high estimated Value. They use these α- and τ-values to calculate scores (see Eq. 1) and establish their preference ranking. Then, the matching algorithm assigns buyers to suppliers or to themselves. Next, all suppliers who are matched to buyers invest in assets and produce for those buyers, possibly generating economies of scale and/or learning-by-doing in the process. The price at which suppliers deliver to buyers is set in such a way, that profits are shared equally between the agents involved: the suppliers have now made their profit for this round. Buyers who are not matched to a supplier produce for themselves, after which all buyers sell their (differentiated) product on the final goods market and make their profit. Finally, both buyers and suppliers use the profit made in the current round as the reward with which to update their estimate of the Value of the particular values they used for α and τ in the current round.

3. ANALYSIS OF PARAMETER REGIME

It is instructive to analyze the influence of some of the parameters of the model, because such knowledge can be helpful in restraining the possibilities for combining different values of simulation parameters, and in interpreting the results from the simulations.

3.1 Initial Trust

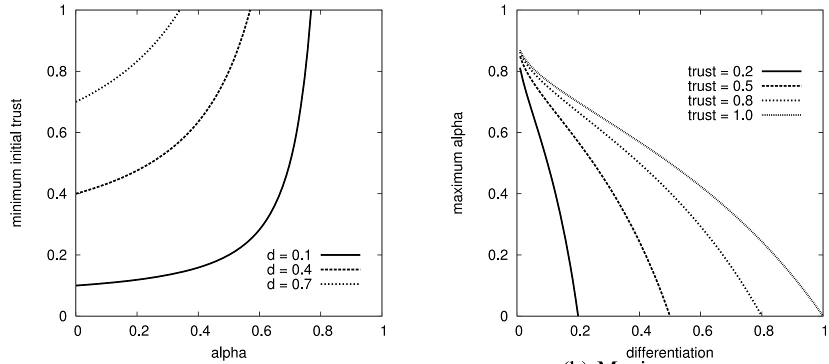
An important parameter dictating the relative attractiveness of the various organizational forms (market or hierarchy) is the differentiation of the buyers' products. Depending on its value, a supplier is able to generate savings through economies of scale or learning. The problem at the start of the simulation is that neither are yet available, so the profit a buyer can make in a relation with a supplier is just $d/2$. The profit a buyer can make by producing for himself is d, leading to the buyer's initial score for himself of d (using $\alpha = 1$ since the buyer's trust in himself is not considered relevant). This means that, for a supplier to have a higher initial score for a buyer than the buyer's initial score for himself, we need that

$$\begin{aligned} \left(\frac{d}{2}\right)^{\alpha} \text{trust}^{1-\alpha} &> d \\ \text{trust} &> d\left(\frac{1}{2}\right)^{\frac{-\alpha}{1-\alpha}}. \end{aligned} \quad (3)$$

As the differentiation of the buyer's product increases, the difference between the profit he can make by producing for himself and the profit he can make by outsourcing also increases. In order for the buyer to still consider outsourcing attractive, he has to have more (initial) trust in suppliers, and should also put more weight on trust vs. profitability (i.e., low α).

These conditions are shown in Figures 2(a) and 2(b), respectively. Specifically (see Eq. 3), Figure 2(a) shows the minimum initial trust for different values of d, over the range of possible values for α. Conversely, the maximum value a buyer may use for α that would still allow a positive score-difference for a supplier, can

[4]To avoid terminological confusion between the different values that α and τ can take (the actions) on the one hand, and the long run *Value* of actions (values for α or τ) on the other hand, we call them α- and τ-values and (capitalized) Values, respectively.



(a) Minimum trust. (b) Maximum α

Figure 2: The conditions for suppliers' positive score-differentials.

be obtained from Eq. 3:

$$
\begin{aligned}
d\left(\frac{1}{2}\right)^{\frac{-\alpha}{1-\alpha}} &< \text{trust} \\
\alpha &< \frac{\ln\left(\frac{\text{trust}}{d}\right)}{\ln\left(\frac{\text{trust}}{d}\right) - \ln\frac{1}{2}}. \qquad (4)
\end{aligned}
$$

This is shown graphically in Figure 2(b) for different values of trust, over the range of possible values for d. If trust $= 1$, then the range of values for α from which the buyer can choose at the start of the simulation and still potentially give the supplier a positive score-difference compared to himself, becomes smaller when d increases. In the simulation, we will run experiments for different values of d, so these figures show how high the initial trust should be set for each value of d, to get some interesting dynamics going.

3.2 Supplier's Acceptance Quotum

The savings that can be generated due to economies of scale are limited by the supplier's acceptance quotum: only when the *total* general purpose assets accumulated by a supplier in her production for multiple buyers exceeds 1, are savings possible. This implies that we need $q_a > 1/(1-d)$. In the simulations, we will only use $d < 0.75$, because higher values require unrealistically high initial trust. Accordingly, we will use $q_a = 4$.

4. NUMERICAL EXPERIMENTS

4.1 Experimental Setup

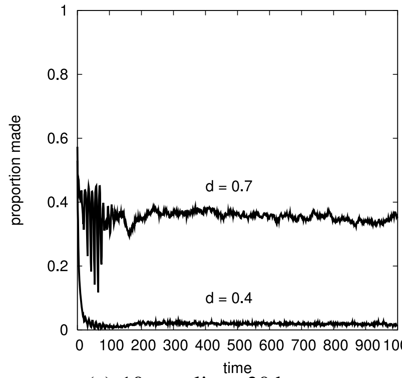
Each experimental run lasted for 1000 timesteps, and was replicated 100 times. Results are presented as averages across runs, buyers, suppliers, etc. as indicated. Each of the agents (both buyers and suppliers) was given a total of 11 evenly distributed α- and τ-values between 0 and 1 inclusive (.0, .1, .2, ..., 1.0) for α. The Value estimates of each of these were initialized 'optimistically' (see [18, p. 39–41]), and subsequently estimated as the 'weighted average' of previous rewards, with stepsize of 0.2. (Optimistic initialization leads to the oscillatory phenomena observable in the initial learning phase of the simulations.) Action selection for both α and τ was done using an ϵ-greedy method with $\epsilon = .1$.

We set basetrust to 0.05, so that, with $d \in \{.4, .7\}$, basetrust is always the lowest score available. The value of trustfactor, f_t in Eq. 2, is 1, same as scale- and learning-parameters f_s and f_l. In each experiment, We set initial trust to $d + 0.1$, so that, for each value of d we simulate, there are values for α within its range which allow positive initial score-differentials for suppliers (see Figure 2(a)).

In the experiments, we simulated different market settings by varying the degree of product differentiation on the buyers' final goods market $d \in \{.4, .7\}$. Setting $q_a = 4$ guarantees that outsourcing is potentially profitable with both values of $d \in \{.4, .7\}$ ($q_a > 1/(1-d)$).

4.2 Emergent Economic Organization

The resulting proportion of intermediate goods that the buyers make rather than buy is shown in Fig. 3.[5] When differentiation is



(a) 10 suppliers, 30 buyers

Figure 3: Proportion made in different experiments.

relatively low, various factors contribute to the relative attractiveness of outsourcing versus making. The suppliers have savings due to economies of scale to offer instantly, and there is also a wider range of values for α which support a possible score-advantage for suppliers, from the buyers' point of view (see Figure 2). This contributes to the minimal proportion of making in situations when $d = 0.4$.

In any case, these results confirm the intuition from transaction cost economics that lower differentiation, and thus lower asset specificity of investments, should lead to more outsourcing. This observation serves as a validation of our model: we do not disagree

[5]The oscillations at the start of the simulations are an artefact of the combination of our methods for Value initialization and Value estimation discussed above.

with transaction cost economics' basic analysis which predicts this qualitative relation, so the fact that our model reproduces it, indicates that, to a certain extent, it is congruent with TCE.

Figures 4(a) and 4(b) show the buyers' indegree, as it evolves over time in the different experiments. A buyer's indegree is 0 if he makes and $q_o = 1$ if he buys. In Figure 4, we have split this latter situation up, according to the number of buyers that the buyer's supplier is supplying to. When $d = 0.4$ (in the case of Figure 4(a)), most of the buyers who buy, do so from the suppliers who have the maximum of $q_a = 4$ buyers, and that is also the vast majority of all suppliers (see Figure 4(c)): on average a little over 6 suppliers supply to an average of about 25 buyers.

The results become more interesting when $d = 0.7$ (Figure 4(b)). The suppliers' possibilities for generating savings due to economies of scale are now limited, so the allocation suffers accordingly. The search problem of the buyers for the optimal supplier is exacerbated by the smaller margin by which different suppliers' profit potentials differ from each other. A much higher fraction of the buyers now makes rather than buys. There is an interesting non-monotone effect in Figure 4(b): because there are not many buyers, it becomes harder for large groups to coalesce around individual suppliers. Finding a supplier who supplies to a single other buyer, or maybe 2, is reasonably easy, but very few buyers buy from suppliers with the maximum ($q_a = 4$) number of buyers. There is on average just 1 of those among the suppliers (cf. Figure 4(d)).

Overall, the highest efficiency is achieved when as many suppliers as possible supply to the highest possible number of buyers (q_a). In the case of overcapacity of supply, allocative efficiency clearly suffers from the agents' coordination problems. With tighter markets, the allocation is eventually closer to optimal. This shows, importantly, that TCE's predictions may in fact not be reached in realistic circumstances: given agents' limited rationality, and their lock-in due to trust and loyalty effects, certain optimal outcomes may be out of reach of the complex adaptive system's developmental paths.

4.3 Evolving Trade Networks

The repeated nature of the interaction process leads to different possibilities with respect to evolution of the topology of the network connecting buyers and their suppliers. Table 1 shows a simple taxonomy of the possibilities. Figure 5 shows the distribution of all outcomes across the matchings in each of the timesteps of the simulation. When $d = 0.4$, especially when the search problem is hard for the buyers (Figure 5(a)), as described above (they have to find each other in a group, buying from a given supplier, which is a harder problem when there are fewer buyers), there remains quite a lot of switching by buyers who go from one supplier to another in search of a solution to their coordination problem. Of those, the majority dump their current supplier and find a new one (outcome 5), which leads to some others being dumped by their supplier, and being forced (and able) to find a new one—there is oversupply in this setting. A rather small fraction switches from buying back to making, all of which explains why the number of suppliers with the highest number of buyers steadily increases over time, at the expense of the number of suppliers with a smaller number of buyers (see Figure 4(c)).

Table 1: Possible outcomes from one matching to the next, from the point of view of the buyer.

		is buying	
	is making	same supplier	different supplier
was making	1	2	
was buying	6: was dumped 7: has dumped	3	4: was dumped 5: has dumped

The results for $d = 0.7$ are more stable. On the one hand suppliers have less of a profit generating advantage relative to the buyers. On the other hand, economies due to learning play a bigger role (because differentiation, and hence, specificity of assets is higher), so over time, partners become more important for each other's profits, and they become more attached to each other, especially compared to others they have not had a long relation with yet.

4.4 Adaptively Learning α and τ

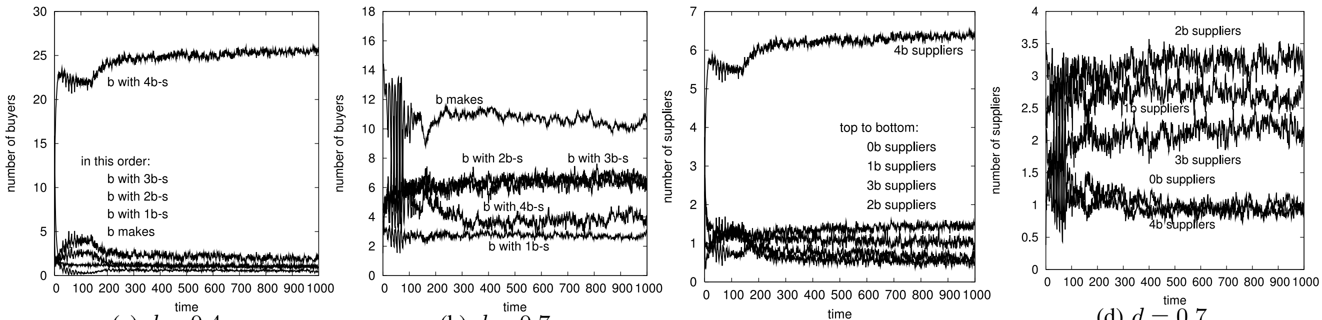
The events in the simulation are generated as a result of the matching in each time step. This, in turn, is determined by the scores the agents assign to each other, which, finally, depend on the values the agents use for α and τ, as described in Section 2. In the current section, we study the α, τ-strategies resulting from the agents' adaptive learning. Figures 6 and 7 show the values used for α and τ by the buyers and the suppliers, respectively, in their different categories.

Some classes of buyers (making, buying from suppliers with different number of buyers) and suppliers contain very few samples, across the 100 runs of the experiment. We only plotted, therefore, the values used for α and τ by agents in categories that had at least 20% of the total number of agents in them. This setup forced us to use points, rather than lines connecting them, because there are now many missing values.

Profits are used as feedback to learn about the suitability and appropriatenesss of using different values for α and τ. Optimal strategies in this respect one can not hope to design analytically, and we don't assume economic agents to be able to, so we assume agents learn, by trial and error, and via reinforcement, what to do. The main effect we see in Figure 6(a) is that the buyers, all of whom are buying from a supplier with $q_a = 4$ buyers, initially learn to use a high value for τ. (The other buyers form too small a fraction to warrant including their strategy in the graph, as discussed above.) As soon as these buyers are 'attached' to these suppliers, the value they use for τ becomes less important, so the distribution of Values estimates becomes flatter again. In terms of α, we notice a progressive decrease for the agents buying from 'fully loaded' suppliers when suppliers are in short supply. Paradoxically, for enhancing profitability, it pays to weigh profitability less than trust. When $d = 0.7$, the interesting result to note is that the agents learn to be less loyal to themselves, and more loyal to suppliers who supply to higher numbers of buyers. Switching away from oneself as a supplier should be easier than switching away from a supplier once one is that supplier's customer.

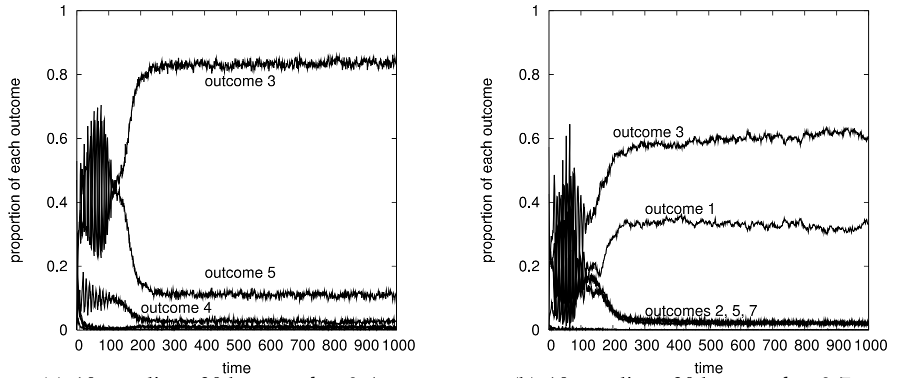
5. CONCLUSION

We have studied a multi-agent buyer-supplier system using an Agent-based Computational Economics model of interacting adaptive agents on an industrial market. The market is modeled as a trade network under continual reconstruction by the agents involved, making and breaking connections as they see fit. In each of a sequence of time steps, the agents are matched using a matching algorithm which takes the agents' preferences for potential partners as input, and returns a topology linked suppliers to their buyers, and buyers to their suppliers (when they buy) or themselves (when they make). Such preferences are functions of expected profit, itself a function of potential profit and trust (as a measure of the probability that the profit potential will indeed materialize), as well as loyalty. The agents adaptively learn to use different relative weights



(a) $d = 0.4$. (b) $d = 0.7$. (c) $d = 0.4$. (d) $d = 0.7$.

Figure 4: The buyers' indegree (4(a) and 4(b)) and the suppliers outdegree (4(c) and 4(d)).



(a) 10 suppliers, 30 buyers, $d = 0.4$. (b) 10 suppliers, 30 buyers, $d = 0.7$.

Figure 5: Dynamic distribution of the different outcomes from Table 1.

of profit vs. trust, and different levels of their loyalty to a current partner.

Having a computer implementation of our artificial market system allows us to investigate agents' adaptive strategies in a wide range of circumstances. Experiments are easily reproducible, and many different types of data, at different levels of analysis, may be extracted from the simulations. A view on the outcomes of the simulations allowed us to investigate in detail the switching behavior by the agents, in search of the optimal allocation.

In general, we are able to explain different emerging distributions of economic activity among organizational forms (market and hierarchy) in terms of the search problem facing the agents. Furthermore, we illustrate the impact of the negative consequences of the agents' search behavior itself on their perceived trustworthiness in the eyes of their potential partners. A further impediment to optimal outcomes we observe is that agents learn to protect themselves and their allocation by being loyal and by focusing on their trust in their partner, rather than their partner's profit generating potential. An important conclusion we draw from our analysis, and show in our simulations, then, is that the rational, optimal outcome predicted by economic theory may not actually be found by boundedly rational agents.

6. ACKNOWLEDGMENTS

We are grateful to the anonymous referees for their comments.

7. REFERENCES

[1] W. B. Arthur. Inductive reasoning and bounded rationality. *Am. Ec. Rev., Papers and Proceedings*, 84:406–411, 1994.

[2] R. H. Coase. The nature of the firm. *Economica NS*, 4(16):386–405, 1937.

[3] R. H. Coase. The new institutional economics. *Am. Ec. Rev.*, 88(2):72–74, 1998.

[4] J. M. Epstein and R. L. Axtell. *Growing Artificial Societies: Social Science from the Bottom Up*. Brookings Institution Press/MIT Press, 1996.

[5] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *Am. Math. Monthly*, 69(January):9–15, 1962.

[6] D. Gambetta. Can we trust trust? In D. Gambetta, editor, *Trust: Making and Breaking Cooperative Relations*, pages 213–237. Basil Blackwell, 1988.

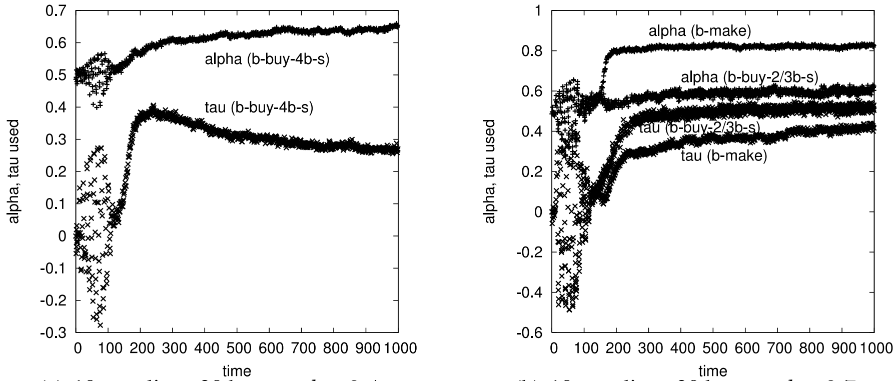
[7] A. Gorobets and B. Nooteboom. Adaptive build-up and breakdown of trust: An agent-based computational approach. *Journal of Management Governance*, 10(3):277–306, 2006.

[8] R. Gulati. Does familiarity breed trust? The implications of repeated ties for contractual choice in alliances. *Acad. Management Journal*, 38(1):85–112, 1995.

[9] J. H. Holland. Complex adaptive systems. *Daedalus*, 121(1):17–30, 1992.

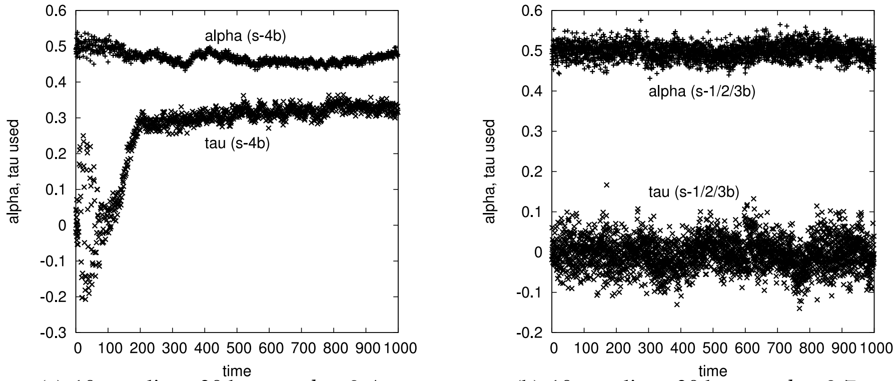
[10] J. H. Holland and J. H. Miller. Artificial adaptive agents in economic theory. *Am. Ec. Rev.*, 81(2):365–370, 1991.

[11] T. B. Klos and H. La Poutré. Decentralized reputation-based



(a) 10 suppliers, 30 buyers, $d = 0.4$. (b) 10 suppliers, 30 buyers, $d = 0.7$.

Figure 6: The average values used for α and τ by the buyers in each of the degree-categories, in the different experiments.



(a) 10 suppliers, 30 buyers, $d = 0.4$. (b) 10 suppliers, 30 buyers, $d = 0.7$.

Figure 7: The average values used for α and τ by the suppliers in each of the degree-categories, in the different experiments.

trust for assessing agent reliability under aggregate feedback. In *Trusting Agents for Trusting Electronic Societies*, volume 3577 of *LNAI*, pages 110–128. Springer, 2005.

[12] T. B. Klos and B. Nooteboom. Agent-based computational transaction cost economics. *Journal of Economic Dynamics & Control*, 25(3–4):503–526, 2001.

[13] B. Nooteboom. Towards a dynamic theory of transactions. *Journal of Evolutionary Economics*, 2(3):281–299, 1992.

[14] B. Nooteboom. *Trust: Forms, Foundations, Functions, Failures and Figures*. Edward Elgar, 2002.

[15] B. Nooteboom, H. Berger, and N. G. Noorderhaven. Effects of trust and governance on relational risk. *Acad. Management Journal*, 40(2):308–338, 1997.

[16] B. Nooteboom, T. B. Klos, and R. J. Jorna. Adaptive trust and co-operation: An agent-based simulation approach. In *Trust in Cyber-Societies*, volume 2246 of *LNAI*, pages 83–109. Springer, 2001.

[17] W. W. Powell. Learning from collaboration: Knowledge and networks in the biotechnology and pharmaceutical industries. *California Management Review*, 40(3):228–240, 1998.

[18] R. S. Sutton and A. G. Barto. *Reinforcement Learning*. MIT Press, 1998.

[19] L. S. Tesfatsion. Introduction to the special issue on Agent-based Computational Economics. *Journal of Economic Dynamics & Control*, 25(3–4):281–293, 2001.

[20] L. S. Tesfatsion. Structure, behavior, and market power in an evolutionary labor market with adaptive search. *Journal of Economic Dynamics & Control*, 25(3–4):419–457, 2001.

[21] L. S. Tesfatsion. Agent-based computational economics: Growing economies from the bottom up. *Artificial Life*, 8(1):55–82, 2002.

[22] M. M. De Weerdt, Y. Zhang, and T. B. Klos. Distributed task allocation in social networks. In *AAMAS*, 2007.

[23] O. E. Williamson. *The Economic Institutions of Capitalism*. Free Press, 1985.

[24] S. Wuyts, M. G. Colombo, S. Dutta, and B. Nooteboom. Empirical tests of optimal cognitive distance. *Journal of Economic Behavior & Organization*, 58(2):277–302, 2005.

[25] M. Zollo, J. J. Reuer, and H. Singh. Interorganizational routines and performance in strategic alliances. *Organization Science*, 13(6):701–713, 2002.