Organization, Learning and Cooperation*

Jason Barr[†] and Francesco Saraceno[‡]

November 2003

Abstract

We model the organization of the firm as a type of artificial neural network in a duopoly framework. The firm plays a repeated Prisoner's Dilemma type game, but also must learn to map environmental signals to demand parameters. We study the prospects for cooperation given the need for the firm to learn the environment and its rival's output. We show how a firm's profit and cooperation rates are affected by its size, its rival's size and willingness to cooperate and environmental complexity.

JEL Classification: C63, C72, D21, D83, L13

Keywords: Artificial Neural Networks, Cooperation, Firm Learning, Simulations

*An earlier version of this paper was been presented at the 2003 Annual Meeting of the Society for the Advancement of Behavioral Economics, Lake Tahoe, Nevada, July 29-31.

[†]Rutgers University, Newark. Corresponding Author: Dept. of Economics, Rutgers University, Hill Hall, Newark, NJ 07102. email: jmbarr@andromeda.rutgers.edu.

[‡]Observatoire Français des Conjonctures Economiques. email: francesco.saraceno@sciences-po.fr

1 Introduction

Information processing and decision making by firms are not typically done by one person. Rather decisions are made by a group of people either in a committee or hierarchical structure. Bounded rationality and/or computational costs preclude the possibility of any one agent collecting, processing and deciding about information relevant to the firm and its profitability. Thus many agents are employed to process this information so that the firm can make informed decisions. But processing information and decision making are costly activities. Large firms, for example, employ hundreds even thousands of 'managers' who do not produce or sell anything, but rather process information and make decisions (Radner, 1993).

Building on previous work (Barr and Saraceno, forthcoming), we model the firm as a type of artificial neural network (ANN) that must make output decisions in a Cournot duopoly framework. The use of ANNs as a model of firm organization allows us to make explicit the nature and cost and benefits of processing information. Agents within the firm are required to evaluate data and communicate this evaluation to others who then make final decisions. As discussed in Barr and Saraceno (2002, forthcoming), the benefit to the firm of increased resources devoted to information processing (IP) is better knowledge of the environment and hence increased returns. But the costs include the costs of paying agents and the time and costs involved with processing and communicating information.

In Barr and Saraceno (forthcoming) we asked: How does learning affect a firm's ability to produce along the best response function, i.e., the optimal output a firm should produce given the decision of its rival? We showed that ANNs competing in a duopoly setting can learn to converge to the Nash equilibrium output, which is changing each period due to changing environmental states. Further we showed environmental complexity affects both organizational size and profitability. In that setting, firms choose an output level given the observation of the environmental state and, after that, the price is set, the market clears, and firms then compare their output choices to the best response they should have played given the output choice of their opponent.

The ANN embodies the knowledge of the firm and implicitly contains a history of what the rival played in the past. Over time, firms learn to forecast both the effect of changing environmental states on demand and the rival's output; both firms converge to the Nash equilibrium over time. In that framework, we did not consider the possibility of collusion in the sense that firms could possibly learn to produce less than Cournot output thus gaining larger profits.

Here, we are interested in studying the prospects for cooperation with two additional factors, as compared to traditional cooperation models (see Axelrod, 1984): (1) the need for firms to map environmental characteristics to changing product demand, and (2) many agents within a firm are needed to learn not only the environment but what exactly the rival is signalling about its output strategy, i.e., whether it will cooperate or defect. We argue that a firm's knowledge not only of the environment but also its ability to see what its rival is doing will ultimately affect its profitability. If, for example, a firm learns that its rival has a low desire to cooperate, all else equal, it too should change its output to defect most of the time. But if a firm learns that its rival is willing to cooperate then mutual gains are possible. The difficulty of learning affects a firm's ability to cooperate, since in highly complex environments, firms will have problems distinguishing what part of the variation in the market clearing price is coming from the variation in the environment versus the variation in the rival's strategy.

In this paper, we have two neural networks competing in a repeated duopoly setting. Each period, the firm (network) views an environmental state vector (an N-length vector of $0's$ and $1's$) which it uses to estimate two variables: the intercept parameter of the demand curve, and its rival's output. The firm uses the estimate of its rival's output to decide if its rival is going to defect or cooperate. If the estimate of its rival's output is greater than the shared-monopoly output (plus some margin), the firm plays the (estimated) Cournot best response, otherwise it plays the (estimated) shared-monopoly response. In this sense, both firms are playing a kind of tit-for-tat strategy: 'I defect if I think you will defect, otherwise I will cooperate.'

After each firm chooses an output, it observes the market clearing price, the rival's output and the true demand intercept. The firm uses this information to calculate the error it made in its estimation of the intercept and its rival's output, and then uses this error to improve its performance in the subsequent periods. Thus the knowledge of both the environmental characteristics and the rival's past plays lies within the network itself, rather than any one agent; the network serves as a economical storage device: 'given what I have learned in the past, I now can map environmental information to a rival's output and to the demand intercept.'

In this paper, we measure **environmental complexity** by the number of environmental bits (signals) a firm views each period. Each period a new environmental vector is randomly chosen with probability of $1/2^N$, where N is the number of bits in the vector. This is tantamount to what might be referred to as a 'pure generalization' process.[1] In our paper, for N large enough, this has the effect of having an environment that is always changing each period.

In this paper we ask:

- What is the relationship between network size, learning and profitability given that firms are learning both the environment and their opponent's behavior?
- How does environmental complexity affect performance, cooperation and profits?
- What is the relationship between a firm's willingness to be 'nice,' i.e., it willingness not to defect given that it estimates its rival will defect, profits and cooperation.
- What are the best response functions?
- What is the average firm size in equilibrium versus environmental complexity.

To anticipate some of our results, we will be drawing two sets of conclusions. The first is that environmental complexity and firm dimension interact in a complex way: performance is a function of the not only the difficulty of the IP task, but also the rival's size and willingness to cooperate. Second, we will show that cooperation rates are also affected by environmental complexity (more complex environments hamper complexity) and firm size (larger firms tend to be more aggressive).

The rest of the paper is as follows. In the next section we briefly review the literature to which our paper relates. The following (section 3) outlines

[1] In the computer science literature, the type of ANNs that we employ–the Backward Propagation Network–normally 'trains' on a fixed data set and then is presented new data for forecasting (Croall and Mason, 1992). After training the network can 'generalize' in the sense that it can make forecasts using new, unseen data. Here, the networks train and generalize simultaneously.

the standard Cournot model that we work with, which is an extension of the RPD game. Next, section 4 discusses the particular game that the neural networks play and the characterization of the economic environment. In section 5 we discuss the workings of the particular ANN that we use. Then sections 6 though 8 present the results of our simulation experiments. Finally, in section 9 we present concluding remarks and possible research extensions.

2 Related literature

There is a rich literature on the issue of cooperation and defection in Prisoner's Dilemma type games, of which the Cournot game is a variant. In this general framework, two players must make a decision over two outcomes: whether to cooperate or defect. The result (payoff) of the decision, however, is affected by the rival's decision. If decisions have to be made repeatedly and there is little prospect that the game will end in a short time then mutual sustained cooperation is possible, even if there is no direct communication. In this vein, firms can signal their willingness to cooperate over time by choosing low output over high output. If the other firm 'takes the bait' by also producing a low output, then their is mutual gain to be had (assuming a sufficiently high discount rate).

In terms of the repeated Prisoner's Dilemma (RPD), a standard game theoretic result is that of the 'folk theorem,' which says that if agents are patient enough, then there are an infinite set of Nash Equilibrium outcomes that have higher pay-offs than the 'defect every period' strategy (the so-called min-max payoff) (Fudenberg and Tirole, 1991).

Recently, models of the RPD have also been concerned with bounded rationality and the evolution of cooperation. Rubinstein (1986) and Cho (1994) model agents as boundedly rational automata-type machines. Rubinstein's machine is a finite automata and Cho's is a simple perceptron. These papers show the types of equilibria that can arise. If, for example, there is a bound placed on level of complexity in Rubinstein's machine then only a finite number of equilibrium outcomes can be generated. Cho's machine is able to 'recover' the perfect folk theorem using a neural network that maintains an upper bound on the complexity of equilibrium strategies.

While these papers focus on the nature of the machines and the nature of the equilibria outcomes, other papers focus on the evolution of cooperation (Axelrod, 1997; Miller, 1996; Ho, 1996). For example, Miller demonstrates how cooperation can evolve overtime if automata machines adapt using a genetic algorithm (Holland, 1975) that allows the strategic environment to change. Further he studies co-evolution of strategies under imperfect information. Axelrod (1997) also models the RPD with a genetic algorithm, but fixes the population of possible strategies.

Our paper fits within the literature on adaptive machines that play a repeated Prisoner's Dilemma type game but is different in the following ways. First, we are interested in studying an RPD using an agent-based model of the firm. Our interest is in asking the questions: How are profits and cooperation affected by agent-based learning? and what are the optimal number of agents needed to learn both the external economic environment and the rival's output decision over time? For simplicity we hold the firms' strategies constant (as a type of Tit for Tat strategy) and focus on the relationship between firm complexity (network size), environmental complexity (the quantity of information), profitability and cooperation. Thus our objective is to focus not on the evolution of strategies or the types of equilibria outcomes but rather the learning process that firms need to do in order to improve performance and profits.

In addition to RPD games, cooperation in Cournot models have been widely discussed (see Tirole (1988) for a review of these models). Similar to the RPD, collusion is possible if firms are sufficiently patient and the threat of punishment exists (Verboven, 1997). Cyret and DeGroot (1973) show cooperation is possible if firms maximize joint profits; further they can come to cooperate over time by a process of Bayesian learning. Vriend (2000) presents an adaptive model of a Cournot game, where agents evolve according to a genetic algorithm. He shows how equilibrium market outcomes can be different depending if agents perform individual rather than social learning. Our model is similar to these papers in the sense that firms are adaptive, but unlike these models firms's output decisions evolve based on not only then rival's behavior, but also given the nature of the environment and on the inherent 'niceness' of firms themselves.

Our work also fits within the literature on agent-based models of the firm (Radner, 1993; DeCanio and Watkins, 1998; Carley, 1996). These models, borrowing heavily from computer science, represent the firm as a network of information processing agents (nodes). In general these papers study which types of networks minimize the costs of processing and communicating information. Our model is also agent-based in the sense that we assume that output decisions by the firm are made by a network of information processing agents. However, our work is different in two respects: in general, and unlike other agent-based models, we directly model the relationship between the external environmental variables, firm learning and performance; secondly, we explicitly provide an agent-based model of Cournot competition and cooperation, which to our knowledge has not been done before.

Our agent-based approach models the firm as a type of artificial neural network, with the nodes representing managers. ANNs are common in computer science and psychology, where they have been used for pattern recognition and modeling of the brain (Croall and Mason, 1992; Skapura, 1996). In economics, neural networks have been employed less frequently. One application has been to use ANNs as non-linear estimation equations (Kuan and White, 1992). Because of the stochastic and non-linear nature of ANNs we employ a simulation-based approach to studying the relationship between firm performance, competition and size.

3 The duopoly framework

This section will give a textbook summary of standard Cournot theory in a static and repeated framework. Then, in the next section we will introduce uncertainty and show how we model firms as ANNs.

Let's say we have a market with two firms. Each period they face the demand function:

$$p_t = \alpha_t - \beta \left(q_{1t} + q_{2t} \right).$$

In section 4 α will be a variable, and we will assume that firms must estimate its value from period to period; but for the moment, to review the Cournot game, we take it as constant and known to the firms. Here and for the rest of the paper, we also assume that the slope is constant (and normalized to 1). Profits for each firm are

$$\pi_j = \left[\alpha - (q_1 + q_2) \right] q_j - c_j, \; j = 1, 2$$

where c_j is costs, such as the cost of network, and is set to zero for convenience and without loss of generality.[2]

[2]While we do not deny the importance of the cost of carrying a network of a given size, we do not include this cost in the paper for simplicity since the qualitative results would not change.

Under standard Cournot assumptions, the best response function is given by:

$$q_j^{br} = \frac{1}{2}\left[\alpha - q_{-j}\right],$$

with a Nash Equilibrium of

$$q^{ne} = \frac{\alpha}{3}, \;\; \pi_j^{ne} = \frac{1}{9}\alpha^2.$$

This is a typical prisoner dilemma's game. If the two firms could coordinate their output decisions and act as a monopoly their joint profit from production would be

$$\pi^m = \left[\alpha - Q\right] Q$$

with an optimal output of

$$q^m = \frac{\alpha}{4}.$$

Profit would then be

$$\pi_j^m = \frac{1}{8}\alpha^2 > \pi_j^{ne} = \frac{1}{9}\alpha^2$$

Unfortunately in a single shot game the cooperation outcome is not an equilibrium;[3] if one firm knew that its rival would play half of the monopoly output, then it could defect by playing its Cournot best response and achieve a higher payoff:

$$\begin{aligned} q_j^d &= \frac{1}{2}\left[\alpha - q^m\right] = \frac{3}{8}\alpha \\ \pi_j^d &= \frac{9}{64}\alpha^2 > \pi_j^m \end{aligned}$$

On the other hand, as is well known, if the game is repeated other equilibrium outcomes can emerge.

3.1 Repeated prisoner dilemmas and the folk theorem

In a repeated framework, firms face the decision each period wether to 'cooperate' or 'defect'. Then, whether the horizon is finite or not will yield completely different results. Suppose that the game is repeated infinite times,

[3] Here we define cooperation as splitting the monopoly quantity and defection is each playing best response

and that firm j's payoff is given by the discounted sum of profits:

$$\Pi_j = \sum_{t=1}^{\infty} \delta^{t-1} \pi_{jt}$$

where δ is the discount rate.[4] In this case many other strategies, involving cooperation can constitute an equilibrium like the pure best response uncooperative one. Take for example the **grim trigger** strategy: cooperate until the opponent defect, and in case it defects revert to best response forever. Then the opponent faces the following strategy: Either defect, and obtain

$$\pi^d + \sum_{t=2}^{\infty} \delta^{t-1} \pi^{ne} = \pi^d + \pi^{ne} \frac{\delta}{1-\delta}$$

or cooperate and obtain

$$\sum_{t=1}^{\infty} \delta^{i-1} \pi^m = \frac{\pi^m}{1-\delta}$$

Hence the cooperative outcome will be assured for

$$\frac{\pi^m}{1-\delta} > \pi^d + \pi^{ne} \frac{\delta}{1-\delta}$$

In our case, this means that

$$\begin{aligned} \frac{1}{8}\alpha^2 \frac{1}{1-\delta} &> \frac{9}{64}\alpha^2 + \frac{1}{9}\alpha^2 \frac{\delta}{1-\delta} \\ &\Rightarrow \delta > \frac{9}{17} \end{aligned}$$

In words, if the discount rate is large enough, future losses will more than compensate the short run gain from defection, and cooperation will be the outcome of a grim trigger strategy. But the same could be showed for the so called **tit-for-tat** strategy, consisting in beginning with cooperation, and playing at each period what the opponent played in the preceding one. In fact

[4] Notice that δ can either be interpreted as the discount factor of an infinitely repeated game, or as the probability that the game is repeated after each round when the game length is undefined. analytically the two 'stories' boil down to the same thing.

it can be shown that if the discount rate is high enough, almost any strategy involving cooperation can be an equilibrium strategy. This is in fact what the folk theorem tells us: Any feasible expected payoff can be sustained in an equilibrium as long as each player can expect an equilibrium payoff larger than the uncooperative one. In that case, no player will have an incentive to deviate.

With a backward induction argument, on the other hand, it can be shown that cooperation is not sustainable if the game is repeated a finite number of times. Very simply put, if the game is over at time T, the lack of threat of future retaliation will give the non cooperative outcome at time T. But then, at time $T-1$ the same will hold, because both players will know that in the future they will not cooperate. This will also be true at time $T-2$ and so on, until time 1. Hence with finite horizons, there is no possibility of cooperation in a standard Cournot game.

An important extension of the preceding framework, that relates with the present paper, is the consideration of the effects of uncertainty on the sustainability of cooperative equilibria.[5] This is typically modelled as a price uncertainty: The constant term of a linear demand function shifts according to a given probability distribution. The obvious effect of such a feature of the model is that deviations from the collusive price and profit are not directly attributable to the competitor's unwillingness to cooperate, but may stem from shifts in the demand function. The punishment scheme designed by a firm to force cooperation has as a consequence to be more complex than in the case of certainty. This typically involves a trade-off, whose outcome depends on the particular model adopted: If punishing is too harsh, the firm loses possible advantages from collusion with its partner; but if it is too light, then the opponent may be tempted to adopt a noncooperative stance, and cheat. To sustain cooperation, hence, firms have to punish their opponents only if prices and profits deviate "too much" from the cooperative level. We'll adopt a similar perspective in what follows, with a crucial difference. We add the possibility for firms to reduce uncertainty by means of learning; our focus in fact is on this learning process, and on how it affects the willingness to cooperate.

The general result of these models is that the costs of monitoring and punishing increase with the number of participants, and with the noisiness of the environment. As a consequence, collusion becomes harder to sustain.

[5] The seminal work, on this issue is the often cited paper by Green and Porter (1984).

This section has shortly reviewed some textbook results on repeated prisoner dilemma games, that constitute the basis of standard Cournot duopoly analysis. As we already hinted in the introduction, we modify the basic game in two respects. The first is that we introduce learning, in the sense that the demand curve shifts from period to period depending on a series of observable environmental variables ($\alpha_t = f(\mathsf{x}_t)$), and that firms have to learn what the functional form f is in order to learn its position. Furthermore the firms also have to learn their opponents behavior. The second is that this learning is costly and depends on the organizational structure of the firm, that we will model as a Neural Network.

4 A model of firm learning in a repeated Cournot game

4.1 Strategies

In this paper we have each network employ a relatively simple strategy: a type of Tit-for-Tat (TFT). The standard TFT says that a firm should begin by cooperating and then play the same outcome as the rival's prior move. Given our framework, this strategy allows for the possibility of cooperation once firms begin to learn the external environment, and learn to separate the variance in price that is due to environmental change versus their rival's output decision.[6]

The TFT strategy that firms employ in this paper is a slight variation. Since firms estimate both the demand parameter and their rival's output quantity each period, they have to use this information to decide whether to defect or not. More specifically the firm chooses an output each period based on the following rule ($j = 1, 2,$):

$$q_j = \left\{ \begin{array}{c} \frac{1}{2}\left(\hat{\alpha}_j - \hat{q}^j_{-j}\right) \text{ if } \left(\hat{q}^j_{-j} - \frac{\hat{\alpha}_j}{4}\right) > \rho_j \\ \hat{\alpha}_j/4 \text{ otherwise} \end{array} \right\}, \tag{1}$$

[6] TFT, in general, behaves according to the four rules of thumb discussed by Axelrod (1984) for strategies that are likely to promote cooperation in a setting where boundedly rational agents are the players: (1) Be nice: never be the first to defect. (2) Be forgiving: be willing to return to cooperating even if your opponent defects.(3) Be simple: the easier it is to discover a pattern in a rival's output the easier it is to learn to cooperate. (4) Don't be envious: don't ask how well you are doing compared to your rival, but rather how much better you can do, given your rival's actions.

where q_j is firm $j's$ output, $\hat{q}^j_{-j}$ is firm $j's$ estimate of its rival's output, and $\hat{\alpha}_j$ is firm $j's$ estimate of α. Equation (1) says that if the firm estimates its rival to be a cheater: $\hat{q}^j_{-j} > \hat{\alpha}_j/4 + \rho_j$, i.e., that the rival is expected to deviate from forecast monopoly profit, then it plays the optimal forecasted Cournot output; that is, it defects as well.

The threshold value $\rho_j \geq 0$ represents the firm's 'willingness to be nice.' For relatively small values, e.g., $\rho_j = 0$, firm j will play defect relatively more often; for values $\rho_j \geq \ddot{\rho}$, the firm will be so nice that it will never defect. Notice that in making this decision the firm has two possible sources of error: the first is the environment, and the second is the opponent's quantity; this is why it will allow a deviation ρ_j from the monopoly output before reverting to the non cooperative quantity.

4.2 The economic environment

4.2.1 A shifting demand curve

Here we represent the external environment as a vector of binary digits $\mathsf{x} \in \{0,1\}^N$. The relationship between the environment and the intercept is given by

$$\alpha(\mathsf{x}) = \frac{1}{2^N} \sum_{k=1}^{N} x_k 2^{N-k},$$

x_i is the k^{th} element of x. This functional relationship converts a binary digit vector into its decimal equivalent.[7] We can think of this in the following manner: the vector x contains signals (information) from the environment, which are arranged in order of increasing importance. $\alpha(\mathsf{x})$ can be thought of as a weighted sum of the environmental signals. Each period, the firm views an environmental vector x and uses this information to estimate the value of $\alpha(\mathsf{x})$.

4.2.2 Environmental change

Each period an environmental vector is randomly chosen with probability $1/2^N$. For example, for $N = 10$, each vector has a probability of 0.000977 of being selected. In our simulations it is very unlikely for a for an environmental vector to be viewed by the firm more than once; the learning process

[7] The value of α is normalized to be between 0 and 1 by dividing by $1/2^{N}$.

highlights the pattern recognition features of our neural network model of the firm.

4.2.3 Complexity

To measure the complexity of the information processing problem, we define environmental complexity as the number of bits in the vector, N, which, in the simulations below, ranges from a minimum of 5 bits to a maximum of 50.

4.3 The steps

Here we outline the behavior of the firm each period (time subscripts dropped for notational convenience):

1. Each firm observes an environmental state vector x.
2. Based on that each firm estimates a value of the intercept parameter, $\hat{\alpha}_j$. Further the firm also estimates its rival's choice of output, $\hat{q}^j_{-j}$, where $\hat{q}^j_{-j}$ is firm $j's$ guess of firm $-j's$ output.
3. Based on the values the firm estimated in step 2, it makes an output decision using the TFT-type rule given by equation (1).
4. It then observes the true value of and α, and q_{-j}, and uses this information to determine its errors using the following rules:[8]

$$\varepsilon_{1j} = (\hat{\alpha}_j - \alpha)^2 \tag{2}$$

$$\varepsilon_{2j} = \left\{ \begin{array}{c} \left(\hat{q}^j_{-j} - q_{-j}\right)^2 \text{if } \left(q_{-j} - \frac{\alpha_j}{4}\right) > \rho_j \\ \left(\hat{q}^j_{-j} - \alpha/4\right)^2 \text{ otherwise} \end{array} \right\} \tag{3}$$

5. Based on these errors the firm updates the weight values in its network for improved performance in the next period. This process is outlined in the next section.

[8] In Barr and Saraceno (forthcoming) we demonstrate the how profit and error are inversely related, and that maximizing profit is the same as minimizing the error. For the sake of brevity, we do not show a similar result in this paper since the error rule is relatively more complex, but the same result applies in this case and is confirmed by regression analysis (results not shown).

5 The firm as artificial neural network

The neural network is comprised of three 'layers': the environmental data (i.e., the environmental state vectors), a hidden/managerial layer, and an output/decision layer. The 'nodes' in the managerial and decision layers represent the information processing behavior of agents. Each agent takes a weighted sum of the information it views and applies a type of squashing function to produces an output/signal: a value between zero and one. In the managerial layer, this output is then passed/communicated to the decision layer. Each output node takes a weighted sum of the signals from the managerial layer to produce a particular decision: an estimated intercept value or an estimated rival's output. A graph of the network is shown in Figure 1. This type of ANN is referred to as a backward propagation network (BPN).

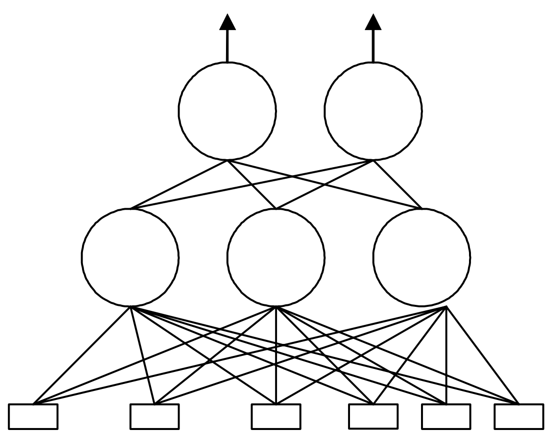

Figure 1: A graph of a neural network.

More specifically, the environmental data (information) layer is a binary vector $\mathsf{x} \in \mathsf{X}$ of length N. Each manager (node) in the management (hidden) layer takes a weighted sum of the data from the data layer. That is, the i^{th} agent in the management layer of firm j calculates

$$in_{ij}^{h} = \mathsf{w}_{ij}^{h}\mathsf{x} \equiv \left[w_{1i}^{h}x_1 + \ldots + w_{Ni}^{h}x_N\right],\ i = 1, \ldots, M_j;\ j = 1, 2.$$

Thus the set of 'inputs' to the M_j managers of the management layer is

$$\mathsf{in}_j^h = \left(in_{1j}^h, \ldots, in_{ij}^h, \ldots, in_{M_j}^h\right) = \left(\mathsf{w}_{1j}^h\mathsf{x}, ..., \mathsf{w}_{ij}^h\mathsf{x}, ..., \mathsf{w}_{M_jj}^h\mathsf{x}\right).$$

Each manager then transforms the inputs via a squashing (voting) function to produce an output, out_{ij}^h. Here we use one of the most common squashing functions, the sigmoid function: $out_{ij}^h = g(in_{ij}^h) = 1/\left(1 + \exp\left(-in_{ij}^h\right)\right)$. Large negative values are squashed to zero, high positive ones are squashed to one, and values close to zero are 'squashed' to values close to 0.5.[9]

The vector of processed outputs from the management layer is

$$\mathsf{out}_j^h = \left(out_{1j}^h, \ldots, out_{ij}^h, \ldots, out_{M_jj}^h\right) = \left(g(in_{1j}^h), ..., g(in_{ij}^h), ..., g(in_{M_jj}^h)\right).$$

The inputs to the output layer are weighted sums of all the outputs from the hidden layer:

$$in_{\iota j}^o = \mathsf{w}_{\iota j}^o \mathsf{out}_j^h \equiv \left(w_{1\iota j}^o out_{1j}^h + ... + w_{M_j\,\iota\,j}^o out_{M_j\,j}^h\right), \ \iota = 1, 2.$$

All weights in both layers can take on any real value. Finally, the outputs of the network–the estimate of demand intercept $\hat{\alpha}_j$ and $\hat{q}_{-j}^j$–are determined by transforming $in_{\iota j}^o$ via the sigmoid function, $\left\{\hat{\alpha}_j = g\left(in_{1j}^o\right), \hat{q}_{-j}^j = g\left(in_{2j}^o\right)\right\}$. We can summarize the behavior of network with two outputs as

$$\begin{aligned} \hat{\alpha}_j &= g\left[\sum_{i=1}^{M_j} w_{i1j}^o g\left(\mathsf{w}_{ij}^h\mathsf{x}\right)\right], \\ \hat{q}_{-j}^j &= g\left[\sum_{i=1}^{M_j} w_{i2j}^o g\left(\mathsf{w}_{ij}^h\mathsf{x}\right)\right]. \end{aligned}$$

5.1 The learning algorithm

The above process describes the input-output nature of the neural network. However, the distinguishing feature of the network is its ability to learn. After the market price is set and the market clears, the firms learn the true values of the intercept and rival's output. They then use this information

[9] The sigmoid function can represent the votes of managers since the sigmoid is a continous version of the Heaviside function.

to help improve their performance in successive periods. They calculate the error they made and then update the network weights using a gradient decent method that changes the weights in the opposite direction of the gradient of the error with respect to the weight values. We begin with a completely untrained network by selecting random weight values (i.e., we assume the network begins with no prior knowledge of the environment).

An environmental state vector is realized and the networks processes it, as described above, to obtain outputs $\left\{\hat{\alpha}_j,\ \hat{q}_{-j}^j\right\}$. These outputs are compared to true values to get an error for each one, according to equations (2) and (3). Total error is then calculated:

$$\xi_j = \varepsilon_{1j} + \varepsilon_{2j}$$

This information is then propagated backwards as the weights are adjusted according to the learning algorithm, that aims at minimizing the total error, ξ_j. Define $\hat{y}_j = \left\{\hat{\alpha}_j,\ \hat{q}_{-j}^j\right\}$ and $y_j = \{\alpha,\ q_{-j}\}$. The gradient of ξ_j with respect to the output-layer weights is

$$\frac{\partial \xi_j}{\partial w_{i\iota j}^o} = -2\left(y_{\iota j} - g\left(in_{\iota j}^o\right)\right) g'\left(in_{\iota j}^o\right) out_{ij}^h,$$

where $i = 1, ..., M_j;\ \iota = 1, 2;\ j = 1, 2$. For the sigmoid function, $g'\left(in_{\iota j}^o\right) = \hat{y}_{\iota j}(1 - \hat{y}_{\iota j})$.

Similarly, we can find the gradient of the error surface with respect to the hidden layer weights:

$$\frac{\partial \xi_j}{\partial w_{ikj}^h} = -2g'(in_{ikj}^h)x_k\left[\left(y_{\iota j} - g\left(in_{\iota j}^o\right)\right) g'\left(in_{\iota j}^o\right) w_{i\iota j}^o\right],$$

where $i = 1, ..., M_j;\ \iota = 1, 2;\ j = 1, 2;\ k = 1, ..., N$.

Once the gradients are calculated, the weights are adjusted a small amount in the opposite (negative) direction of the gradient. We introduce a proportionality constant η, the learning-rate parameter, to smooth the updating process. If we define $\delta_{\iota j}^o = .5(y_{\iota j} - \hat{y}_{\iota j})g'(in_{\iota j}^o)$, we have the weight adjustment for the output layer as

$$w_{i\iota j}^o(t+1) = w_{i\iota j}^o(t) + \eta\delta_{\iota j}^o out_{ij}^h.$$

Similarly, for the hidden layer,

$$w_{ijk}^h(t+1) = w_{ijk}^h(t) + \eta\delta_{i\iota j}^h x_k,$$

where $\delta_{i\iota j}^h = g'(in_{ij}^h)\delta_{\iota j}^o w_{i\iota j}^o$. When the updating of weights is finished, the firm views the next input pattern and repeats the weight-update process.

6 Organization, learning and cooperation: a simulation experiment

In this section we present the results of a simulation experiment. The steps of the experiment were outlined in sections 4 and 5 above. We are mainly interested in two different issues: The first is whether the two firms learn, i.e., if in the long run they are able to map signals from the environment to demand conditions and the opponent's. The second is how the environmental complexity that firms face, especially in the first stages of their interaction, affects their decision to cooperate or defect as well as their profitability. For each set of parameter values, we rerun the simulation 50 times and take average values in order to smooth out fluctuations.

In this section we show some particular runs of the model that are representative of its features. The robustness of these results will then be verified in section 7 by means of an econometric investigation over the parameter space. Here we fix many of the variables and look only at one particular outcome.

6.1 Error

The first question is: Can firms learn both the relationship between the environment and demand and learn to forecast the rival's output decision, given the complexity level of the environment? Below, in Figure 2, we show the results of one firm for two networks each with 8 nodes in the hidden layer, $T = 250$ iterations, $N = 10$, and $\rho_j = 0.05$ $(j = 1, 2)$.[10] We show the separate errors to see that both converge to zero over the long run. We can see that the network is able to improve its forecasts over time. Interestingly, the estimation of the rival's output has a lower error than the estimation of the intercept values.

6.2 Profit and cooperation

Once we made sure the networks learn, we next see how profit and cooperation evolve. Below, in Figure 3, we can see how profits increase over time. We show profits for two different 'niceness' values, $\rho_j = 0$ and $\rho_j = 0.05$ $(j = 1, 2)$ (i.e., both firms play the game with the same niceness parameter).

[10] Since firms are adaptive we do not include a discount parameter in the simulations.

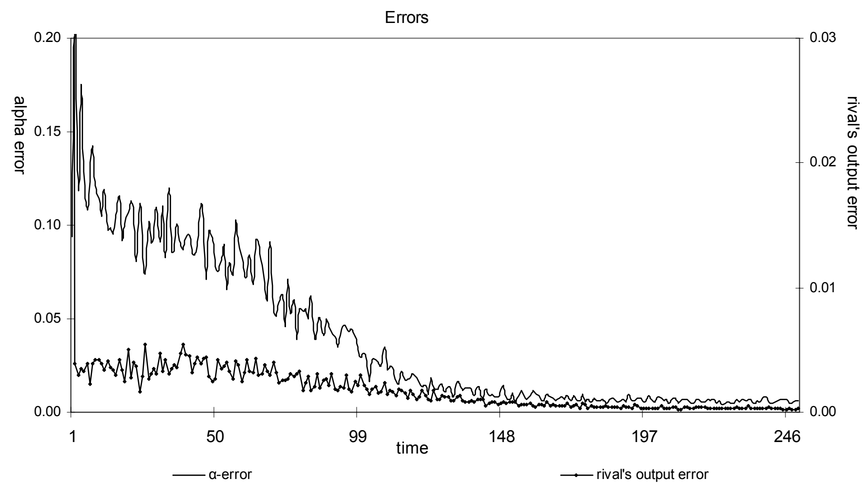

Figure 2: Firm 1 average errors of network over time.

We present the results only for firm 1 since the two firms are symmetric. In the case of $\rho_j = 0$ firms are less willing to cooperate, all else equal. This results in lower profits for the firm. When $\rho_i = 0.05$ profits are relatively higher. Notice, however, that the relative increase in profits due to increased cooperation is not that large. This is due to the fact that since the error is much greater for α, added cooperation only increases profits by a little bit.

(Cesco if you don't like this new graph then the only think i can suggest is that we replace the graph with only moving averages)

In Figure 4 we show how average cooperation rates for firm 1 change over time for two different 'niceness' values. Here we define cooperations rates as follows. Let

$$c_{1t} = \left\{ \begin{array}{c} 1 \text{ if firm 1 plays } C \text{ in period } t \\ 0 \text{ otherwise.} \end{array} \right\}.$$

Then over 50 runs we take the average cooperation rate for each period as $\bar{c}_{1t} = \frac{1}{50}\sum_{r=1}^{50} c_{1rt}$ and this is what we plot versus time in figure 4. Notice that if $\rho_j = 0$, the cooperation rates steadily decrease over time. This is due to the fact that the firm continues to make errors in estimating its rival's output and thus is more likely to defect over time. While if the firm has a sufficiently high niceness parameter this helps sustain cooperation in the face

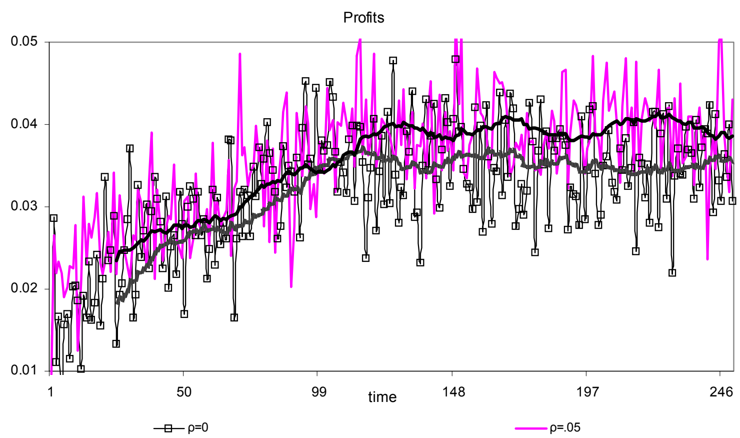

Figure 3: Firm 1 profits over time for two 'niceness' values. Moving Averages (25) also provided.

of error.

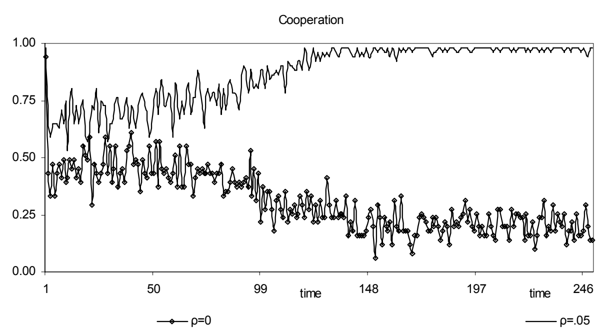

Figure 4: Firm 1 cooperation rates for different 'niceness' values.

6.3 Profits and cooperation versus complexity

Here we look at average profits and average cooperation rates as a function of increasing complexity ($M_1 = M_2 = 8$, $\rho_1 = \rho_2 = 0.05$). We see that the more information to process (i.e., the greater environmental complexity) the lower is profits and average cooperation. This finding implies that all else equal, the prospects for sustained cooperation is diminished in environments that are characterized by complex environments. Figure 5 shows this result.

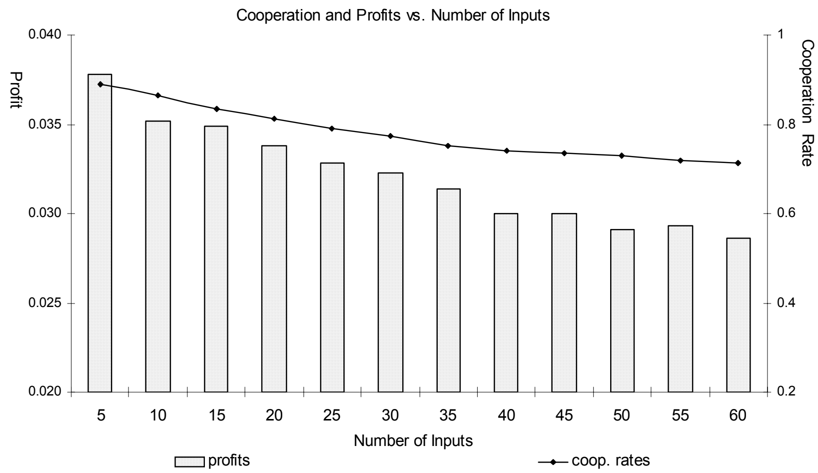

Figure 5: Firm 1 average profits and cooperation rates versus environmental complexity.

6.4 Cooperation versus 'niceness'

The above results are given for symmetric firms. Here we ask the question: What does changing the 'niceness' value for firm 1 do to cooperation rates and profits, holding firm two's niceness parameter constant? Figures 6 and 7 show the results. Here we see that if firm 1 increases its niceness but firm two holds its value constant (where $\rho_2 = 0$) then increased niceness does not pay. Notice, in figure 6, that from a value of about $\rho = 0.25$, firm 1 cooperates almost all the time. The threshold is large enough that defection by the opponent never triggers a non cooperative response. This is why, in section 7, we will focus on a subsample, with $\rho \in [0, 0.25]$.

While firm two also increases its cooperation rate it does not do so at the same rate as firm one. Interestingly firm one experiences a negligible drop in profits, while firm two's profits increase due to its 'defection advantage' compared to firm 1.

In summary, the results of this section show that neural networks competing in a duopoly setting can learn both the economic environment and

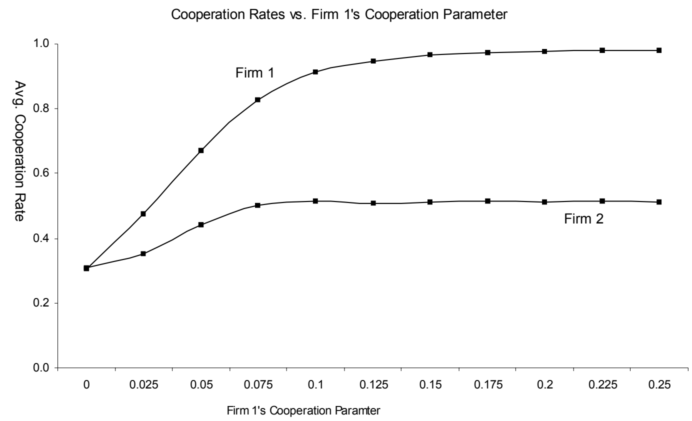

Figure 6: Average cooperation rates for firm 1 and firm 2 versus firm 1's niceness value.

the rival's output decisions. This learning results in increased profit over time, but the level of profit that the two firms can achieve is a function of their willingness to cooperate. The 'nicer' they are the more they can achieve mutual gains to cooperation. But if their niceness threshold is relatively low, they are less likely to cooperate. This is due to the fact that since the environment state changes each period, firms will have some error in estimating the demand function, thus they will always have difficulty estimating the rival's true output and this will make firms more likely to defect. Finally, increasing environmental complexity is associated with lower profits and lower cooperation rates.

7 Regression Results

In our model profits and cooperation choices depend on a number of variables, that interact at different levels (environment, firm complexity, interaction between firms). This is why, after using particular runs to show interesting results of our setup, we need to check for the robustness of our findings. This section will apply standard econometric techniques to a data set generated by our model by random draws of the most interesting parameters. Each observation (we had 4500 of them) consists of the parameters, and of the two dependent variables: Average profit over the run, and the degree of

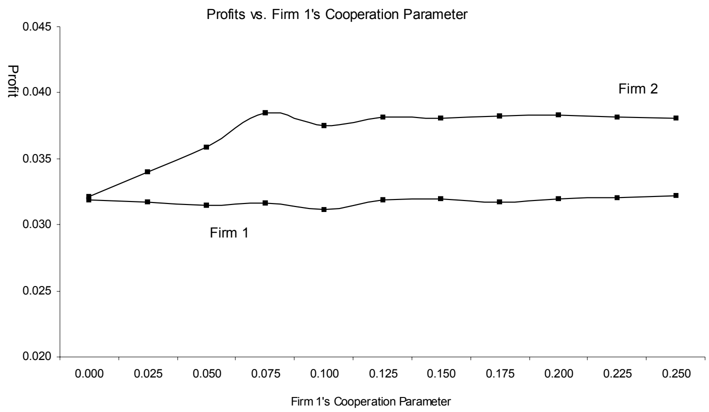

Figure 7: Average profits for firm 1 and firm 2 versus firm 1's niceness value.

cooperation:

$$\Pi_j = \frac{1}{T}\sum_{t=1}^{T}\pi_{jt};\ C_j = \frac{1}{T}\sum_{t=1}^{T}c_{jt}.$$

Table 1 reports the parameter ranges. A first look at the correlation coef-

Variable	M_1, M_2	N	T	ρ_1, ρ_2
Min	2	5	50	0
Max	25	60	500	.25

Table 1: Range of explanatory variables. Note: $\rho_j = \hat{\alpha}_j / r_j, j = 1, 2$.

ficients (Table 2) of our data set can give us some information.[11] {UP-DATE} This rough picture tells us that something interesting. First, as expected, more complex environments (higher number of inputs) yield, in general, lower profit. But, and on this we had no prior, they appear to affect the cooperation rate. Table 2 also tells us that if player 1 has a greater willingness to cooperate then it will have lower profits; an increased willingness of player 2 to cooperate will increase player

[11] Notice that we only focus on firm 1. Firm two is symmetric, so that analyzing it would yield the same results.

	Π_1	C_1	M_1	M_2	N	T	ρ_1	ρ_2
Π_1	1							
	(−)	()	()	()	()	()	()	()
C_1		1						
		(−)	()	()	()	()	()	()

Table 2: Correlation coefficients (t-stats below estimates; significant values in bold).

one's profits. Increasing the number of iterations has strong effects on profits, while the effects on cooperation are modest (even if significant).

Finally, organizational complexity seems to have a positive effect on own profits (and a negative one on the opponent's), but the low value of the coefficient, and our own priors suggest to wait before drawing a conclusion.

The next step is to look more thoroughly into these relationships. To do this we ran some regressions; the results for profit are reported in Table 3.

Variable	Coeff.	Variable	Coeff.	Variable	Coeff.
M_1	19.11 (8.69)	M_1^2	−2.34 (−13.21)	M_1^3	0.03 (7.69)
M_2	−7.91 (−3.63)	M_2^2	0.7 (3.95)	M_2^3	−0.01 (−2.44)
$M_1 \cdot M_2$	−0.33 (−14.37)	T	5.30 (39.69)	T^2	−0.01 (−29.18)
r_1	−7.22 (−23.25)	r_1^3	0.001 (8.12)	N	−7.23 (−22.45)
$M_1 \cdot T$	0.05 (37.25)	$M_1 \cdot N$	0.49 (43.26)	$M_2 \cdot r_1$	0.03 (2.75)
$T \cdot N$	−0.03 (−42.43)	$r_1 \cdot N$	0.04 (8.02)	*constant*	2614 (140.1)
$R^2 = 0.949$					
Number of obs: 4500					

Table 3: Regression results for the dependent variable $10,000 * \Pi_1$. Non significant variables have been omitted.

We can give a summary of the results of the regressions in regards to profits:

- The regression confirms the humped-shaped relationship between firm dimension and profits, cet. par. This result is robust as we've also found it in our other papers under different circumstances.
- The relatiohship between profit and environmental complexity is linearly negative.
- Cet. par. increasing time increases profits since it lowers the firm's error.
- The relationship between profit and willingness to cooperate is negative and non-linear. Holding constant firm 2's willingness to cooperate, if firm 1 increases its willingness to cooperate, it putts itself in the classic prisoner's dilemma situation: if the rival defects then cooperation will yield a lower pay-off.

Before proceeding, we can draw a first conclusion: Our model confirms previous findings (Barr and Saraceno; 2002, forthcoming) on the relationship between firm profitability, size and the environment it faces; specifically, environmental complexity negatively affects profits, and the trade-off relative to firm size, between speed and accuracy emerges in this setting as well, giving a hump shape relationship between firm size and profit. We also showed that more cooperative firms have higher profits.

We now turn to the other dependent variable, the one specific to this paper: the degree of cooperation; the results of the regression are reported in Table 4.

The regressions yield the following results:

- As expected, the willingness to cooperate by any of the two firms, increases the cooperation rate over the relevant range of ρ_1 and ρ_2.
- Also expected is the negative relationship between environmental complexity (N) and degree of cooperation. In a more complex environment, uncertainty makes it harder for a firm to detect defection, by the opponent, who has thus less incentive to cooperate.

Variable	Coeff.	Variable	Coeff.	Variable	Coeff.
M_1	-25.9 (−2.83)	M_1^2	1.79 (2.37)	M_1^3	-0.04 (−2.04)
M_2	51.9 (5.6)	M_2^2	-1.89 (−2.52)	M_2^3	0.032 (1.76)
T	2.30 (22.3)	r_1	-247.7 (−135)	r_1^2	3.10 (118)
r_2	-2.22 (−3.40)	N	-15.55 (−7.14)	N^2	0.12 (4.43)
$M_2 \cdot T$	0.03 (5.97)	$M_2 \cdot N$	-0.65 (−13.39)	$T \cdot r_1$	-0.09 (−37.7)
$r_1 \cdot r_2$	0.05 (2.51)	$r_1 \cdot N$	0.59 (26.14)	*constant*	9609 (139)
$R^2 = 0.959$					
Number of obs: 4500					

Table 4: Regression results for the dependent variable $10,000 * C_1$. Non significant variables have been omitted

- A more interesting picture relates the cooperation rate with firm dimension. We can see that larger firms tend to cooperate less, and to induce the opponents to cooperate more. Said differently, a firm is more likely to cooperate if it is small and if it faces a large competitor. This also tells us that the cooperation rates of the two firms will tend to be approximately equal if they are of similar sizes, but that they will diverge when size diverges.

To summarize, our regressions substantially confirm the robustness of the results of section 6: cooperation is hampered by more complex environments, and of course by lower "niceness," or willingness to cooperate. In addition, our regression analysis sheds light on the relationship between firms size and cooperation, drawing the conclusion that "big is aggressive."

8 Equilibria

Up till now, we've explored the relationship between complexity and niceness on firm profits and cooperation rates. In this section we focus on the likely outcomes of firm size in the Cournot competition framework. The two firms are strategic rivals in three different senses: a rival's choice of output, network size and niceness all affect the firm's pay-offs.

In this section we further explore this relationship by studying the rival's choice of firm size and cooperation on the firm's profit. First we explore network size equilibrium, the we explore a niceness equilibrium:

8.1 To be explore

In section: We saw that we have the function form for profit as:

$$E\Pi_1 = \alpha_0+\alpha_1 M_1+\alpha_2 M_1^2+\alpha_3 M_1^3-\beta_1 M_2+\beta_2 M_2^2-\beta_3 M_2^3+\beta_4 M_1 M_2+\gamma_1 M_1 T+\gamma_2 M_1 N-\gamma_3 M_2 \rho_1+c$$

Thus the best response functions are fairly complicated.....

8.2 Network Size Equilibria

Here we explore the concept of equilibrium with regards to network size, which we call a network size equilibrium (NSE). We define an NSE as a pair of $M's$ such that neither firm has an incentive to change the number of managers. That is to say, in an equilibrium, each network, given the number of agents (nodes) of its rival, finds that switching to another number of agents will decrease its average (total) profit. The equilibrium is a pair $\{M_1^*, M_2^*\}$ such that

$$\Pi_j\left(M_j^*, M_{-j}^*\right) \geq \Pi_j\left(M_j, M_{-j}^*\right), \qquad \forall M_j \quad j=1,2.$$

where

$$\Pi_j = \frac{1}{T}\sum_{t=1}^{T}\pi_{jt}\left(M_j, M_{-j}\right).$$

The regression results in section 7 showed that a rival's network size affects the profitability of the firm. Thus if a rival changes its network size, it will alter the firm's profitability and it too may want to alter its firm size to improve its profitability. Further we see that cooperation rates are affected by a rival's choice of network size.

In this section we ask: What is the relationship between environmental complexity, network size and cooperation rates at equilibrium? We focus on the equilibria that exist after T periods, and do not have endogenous dynamics with the number of managers, i.e., we do not examine firms changing the number of managers during the learning process. Rather we conduct a kind of comparative statics exercise, whereby we look at the NSE that arise for given environmental conditions.

In this experiment, we have networks of different sizes compete for $T = 250$ iterations (for each pair of $M's$ and input numbers we take 50 runs and take averages to smooth out fluctuations in each run). That is to say, firm one and two compete against each other for each size, $M_1, M_2 \in \{2, ..., 20\}$, and $\rho_j = .05$. We repeat this competition for $N = 5, 10, ..., 50$.

For each complexity value we calculated the NSEs that emerge (for each number of inputs there were always more than one equilibrium) and then added the total number of managers of the two firms for each equilibrium to obtain an 'equilibrium industry size.' This gave us a data set of 53 NSEs (for an average of 5.3 per complexity level). Figure 8 shows the results, where we plot average 'industry size' and average cooperation rates (the average of the cooperate rates for the two firms) for each complexity level. As we can see average industry size is increasing in complexity, while average cooperation is decreasing.

8.3 Niceness and Firm Size

9 Conclusion

This paper has presented a model of firm learning and cooperation. We investigate the prospects for cooperation given an agent-based model of the firm which must learn to map environmental signals to changing demand and its

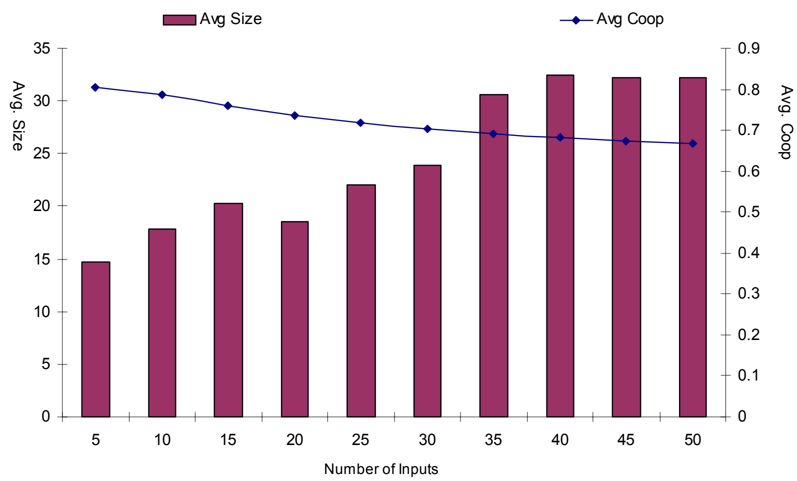

Figure 8: Avg. 'industry' size and avg. cooperation rates versus environmental complexity

rival's output decision. We demonstrate that increased environmental complexity is associated with lower rates of cooperation and lower profitability. We show that in complex environments frequent cooperation is more sustainable when firms are more willing to be 'nice' in the sense that firms are less likely to defect if they estimate their rival will defect. Further we show that firm size has an effect on both profits and cooperation. Increasing firm size, holding the rival's firm constant, increases profits up to a point and then profits decrease. This trade-off is due to the fact that adding agents improves accuracy greater than the loss of speed in learning, up to a point. Increasing firm size, ceteris paribus, has an interesting effect on cooperation: we see a decline in the firm's cooperation as it increases its size; yet we see that it has a positive effect on the willingness of the rival to cooperate. Finally, we have shown that given our definition of network size equilibrium, increasing complexity is associated with larger average firm size and lower average cooperation rates in equilibrium.

This paper leads to several possible research extensions. First we can explore the prospects for cooperation given that firms play different strategies than just the Tit-For-Tat type employed here. Also we can explore the evolution of cooperation given that firms can switch strategies over time.

Further we can see how the composition of the network affects learning and cooperation. For example, we can study how the networks behave when they have more than one hidden layer or when information to agents is restricted, in order to model specialization.

References

[1] Axelrod, R. (1984). *The Evolution of Cooperation*. Basic Books: New York.

[2] Axelrod, R. (1997). *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*, Princeton University Press: Princeton.

[3] Barr, J. and F. Saraceno (forthcoming). Cournot competition, organization and learning. *Journal of Economic Dynamics and Control.*

[4] Barr, J. and F. Saraceno (2002). A computational theory of the firm. *Journal of Economic Behavior and Organization*, 49, 345-361.

[5] Carley, K. M. (1996). A Comparison of artificial and human organizations. *Journal of Economic Behavior and Organization*, 31, 175-191.

[6] Cho, I.-K. (1994). Bounded rationality, neural networks and folk theorem in repeated games with discounting. *Economic Theory*, 4, 935-957.

[7] Croall, I. F. and J. P. Mason (1992). *Industrial Applications of Neural Networks: Project ANNIE Handbook*, Springer-Verlag: New York.

[8] Cyert, R.M. and M.H. DeGroot (1973). An analysis of cooperation and learning in a duopoly context. *American Economic Review*, 63, 24-37.

[9] DeCanio, S. J. and W. E. Watkins (1998). Information processing and organizational structure. *Journal of Economic Behavior and Organization*, 36, 275-294.

[10] Fudenberg, D. and J. Tirole (1991). *Game Theory*, The MIT Press: Cambridge.

[11] Green, Edward and Robert Porter (1984). Noncooperative Collusion under Imperfect Price Competition. *Econometrica*, 52:87-100.

[12] Ho, T-H. (1996). Finite automata play repeated prisoner's dilemma with information processing costs. *Journal of Economic Dynamics and Control*, 20, 173-207.

[13] Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*, University of Michigan Press: Ann Arbor.

[14] Kuan C.-M. and H. White (1994). Artificial neural networks: an econometric perspective. *Econometric Reviews*, 13, 1-91.

[15] Miller, J. H. (1996). The coevolution of automata in the repeated prisoner's dilemma. *Journal of Economic Behavior and Organization*, 29, 87-112.

[16] Radner, R. (1993). The organization of decentralized information processing. *Econometrica*, 61, 1109-1146.

[17] Rubinstein, A. R. (1986). Finite automata play the repeated prisoner's dilemma. *Journal of Economic Theory*, 39, 83-96.

[18] Skapura, D. M. (1996). *Building Neural Networks*, Addison-Wesley: New York.

[19] Tirole, J. (1988). *The Theory of Industrial Organization*, The MIT Press: Cambridge.

[20] Verboven, F. (1997). Collusive behavior with heterogeneous firms. *Journal of Economic Behavior and Organization*, 33, 122-139.

[21] Vriend, N. J. (2000). An illustration of the essential difference between individual and social learning, and its consequences for computational analyses. *Journal of Economic Dynamics and Control*, 24, 1–19.