Cournot Competition and Endogenous Firm Size*

Jason Barr† and Francesco Saraceno‡

Version 1 – Comments Welcome

May 31, 2004

Abstract

We model the firm as a type of artificial neural network that plays a repeated Cournot game. Each period the firm must learn to map environmental signals to both demand parameters and it's rival's output choice. In this paper, though, this Cournot game is in the 'background,' as we focus on the endogenous adjustment of network size. We investigate the long-run behavior of firm/network size as a function of profits, rival's size, and the type of adjustment rules used.

Kew words: Artificial neural networks, firm size, adjustment dynamics.

---

*This draft has been prepared for the 10th International Conference on Computing in Economics and Finance, July 8-10, 2004, University of Amsterdam, The Netherlands.

†Corresponding Author. Dept. of Economics, Rutgers University, Newark, NJ 07102, email: jmbarr@rutgers.edu.

‡Observatoire Français des Conjonctures Économiques, 69 Quai d'Orsay, Paris 75007, France. email: francesco.saraceno@sciences-po.fr

1 Introduction

In this paper we explore firm size dynamics, with the firm modeled as a type of artificial neural network (ANN). Two firms/networks compete at two different levels. The first level, which has been explored in detail in other work (Barr and Saraceno (BS), 2004; forthcoming), looks at Cournot competition between two neural networks. In this paper, this level of competition is essentially in the shadows, while the main form of strategic interaction is in regards to firm size. The firm, while playing the repeated Cournot game, has to make long run decisions about its size.; and size not only affects its own profits, but its rival's profits as well.

Our previous research showed that firm size is a crucial determinant of performance in an uncertain environment, but it was left exogenous. Here we reverse the perspective: we take for given the learning process and the dependence of firm profit on size and environmental complexity, and we endogenize firm size in order and investigate whether simple adjustment rules succeed in yielding the "optimal" (defined as the result of a best response dynamics) size. We explore the concept of firm size adjustment using simple adaptive adjustment rules. The computational requirements needed to discover the optimal firm size equilbria may be quite expensive; and thus we explore less costly methods for adjusting firm size.

We find that using these simple rules can often produce rather complex long-run behavior. We find that only under only very precise parameter values does the firm converge to the actual Nash equilibrium size. And that, in general, the firms using the simple rules actually have larger sizes but lower profits in the long run. Thus in many cases satisficing behavior may not produce optimal outcomes. But this deviation from optimality must be analyzed in terms of the relative costs and benefits: simple rules are less expensive to implement, but profits could be higher if the firm were close to equilibrium size. While we don't directly address this trade-off, we begin in this paper by exploring the outcomes for firm size when firms use simple rule of behavior.

1.1 Related Literature

Information Processing Organizations Our work relates to several different areas. As discussed in detail in other works (BS, 2002, forthcoming), our approach to using a neural network fits within the agent-based literature on information processing (IP) organizations (Chang and Harrington, forthcoming). In this field organizations are modeled as a collection or network of agents that are responsible for processing incoming data. IP networks and organizations arise because in modern economies no one agent can process all the data, as well as make decisions about it. The growth of the modern corporation has created the need for workers who are managers and information processors (Chandler, 1962; Radner, 1993).

Typical models are concerned with the relationship between the structure of the network and the corresponding performance or cost (DeCanio and Watkins, 1998; Radner, 1993). In this paper, the network is responsible for mapping incoming signals about the economic environment to both demand and a rival's output decision. Unlike other information processing models, we explicitly include strategic interaction: one firm's ability to learn the environment affects the other firm's pay-offs. Thus a firm must locate an optimal network size not only to maximize performance from learning the environment but also but also due to its rival's actions. In our case, the firm is able to learn over time as it repeatedly gains experience in observing and making decisions about environmental signals.

Evolutionary Theory and Routines A second area of literature that relates to our work is that of the evolutionary and firm decision making models of Nelson and Winter (1982), Simon (1982) and Cyret and March (1963). In this area, the firm is also boundedly rational, but the focus is not on information processing per se. Rather, the firm is engaged in a myriad of activities from production, sales and marketing, R&D, business strategy, etc. As the firm engages in its business activities it gains a set of capabilities that cannot be easily replicated by other firms. The patterns of behavior that it collectively masters are know as its 'routines' (Nelson and Winter,1982). Routines are often comprised of rules-of-thumb behavior: continue to do something if it is working, change if not.

In this vein, firms in our paper employ simple adjustment rules when choosing a firm size. In a world where there is a lot of information to process and where discovering optimal solutions are often computationally expensive firms will seek relatively easier rules for behavior, ones that produce satisfactory responses at relatively low cost (Simon, 1982).

Complexity Theory and Firm Behavior The next area that overlaps with our work is that of complexity theory (Puu, 2003). We show that simple adjustment rules, based on satisficing behavior, can actually generate very complex long-run dynamics. In fact, in a certain region of the parameter space, we find that firm size patterns can exhibit chaotic behavior in the long run. Though, based on profitability, we have reasons to exclude certain regions of the adjustment parameter space from being economically plausible. Similar work in this area includes Currie and Metcalf (2001) who show that prices can have chaotic behavior when firms employ relatively simple pricing, production and investment routines. Puu (2003) shows how a monopoly's output choices can be chaotic depending on the profit function, and the parameters used in a simple output adjustment rule. Kopel (1996) explore complex dynamics in the Cournot frame work to show how output patterns are dependent on the functional form of the profit and best response functions.

The rest of the paper is organized as follows. Section 2 discusses the motivation for using a neural network as a model of the firm. Then, in section 3, we give a brief discussion of the set up of the model. A more detailed treatment is given in the Appendix. Next, section 4 gives the benchmark cases of network size equilibria. Section 5 discusses the heart of the paper– the firm size adjustment algorithms and the results of the algorithms. Finally, section 6 gives some concluding remarks.

2 Neural Networks as Models of the Firm

In previous work (BS, 2002; 2003) we argued that information processing is a crucial feature of modern corporations, and that efficiency in performing this task may be crucial in defining success and failure. We further argued that when focussing on this aspect of firm behavior, computational learning theory may give useful insights and modelling techniques. In this perspective, it is useful to view the firm as a *learning algorithm,* consisting of agents that follow a series of rules and procedures organized in both a parallel and serial manner. Firms learn and improve their performance by repeating their actions and recognizing patterns (i.e., learning by doing). As the firm processes information, it learns its particular environment and becomes proficient at recognizing new and related information.

Among the many possible learning machines, we focussed on Artificial Neural Networks as models of the firm, because of the intuitive mapping between their parallel processing structure and firm organization. Neural networks, as other learning machines, generalize from experience to unseen problems (as long as they are not too different), i.e., they recognize patterns. Firms (and in general economic agents) do the same: the know-how acquired is used in tackling new problems, at least as long it does not prove completely inadequate.

What is specific of ANNs is the parallel and decentralized processing. ANNs are composed of multiple units processing very simple tasks (on-off) in parallel. The result of this multiplicity of simple jobs may be a very complex task (as weather forecasting). In the same way firms are often composed of different units working autonomously on very specific tasks, and coordinated by a management that merges the results of these simple operations in order to design complex strategies.

Furthermore, the firm, like learning algorithms, faces a trade-off linked to the complexity of its organization. Small firms are likely to attain a rather imprecise understanding of the environment they face; but on the other hand they act pretty quickly and are able to design decent strategies with small amounts of experience. Larger and more complex firms, on the other hand, produce more sophisticated analyses, but they need time and experience to do so. Thus, the complexity of the environment is a crucial element in determining the optimal design of the learning machine. Likewise, the optimal firm structures may only be determined in relation with the environment, and it is likely to change with it. Unlike computer science, however, in economics the search for an optimal structure occurs given a competitive landscape, which imposes time and money constraints on the firm.

In our previous work we showed, by means of simulations, that the trade-off between speed and accuracy generates a hump shaped profit curve, in firm size (BS, 2002). We also showed that as complexity of the environment increases the firm size that maximizes profit also increases. These results reappeared when we applied the model to Cournot competition. We were also able to show that environmental complexity affects the propensity to cooperate of firms (BS, 2004). These results will constitute the background of the present paper. Here, we will leave the Cournot competition and the learning process in the background, and investigate how network size changes endogenously.

3 Setup of the Model

The setup of the neural network model is taken from BS (2004) and BS (forthcoming). Two firms compete in quantities facing a linear demand function whose intercept is unobserved. Firms observe a set of environmental variables that are related to demand, and have to learn the mapping between the two. But firms have to learn two variables: the demand intercept, and their rival's output choice. Appendix A gives the details of the model. Our previous work shows that in general they are capable of learning how to map environmental factors to demand, which allows the firms to converge to the Nash equilibrium. We further showed that the main determinants of firm profitability are, on one hand, firm sizes (i.e. the number of processing units of the two firms, m_1 and m_2); and on the other environmental complexity, that we modeled as being the number of inputs of the network, i.e., the number of environmental factors affecting demand (n).

A few robust features of the learning process may be recalled here (the reader is referred to BS (2004), and BS (forthcoming) for details about the model and the results). *(1)* More complex environments yield, ceteris paribus, lower profit for both competitors. *(2)* In general, firm profit is hump shaped with respect to own size. A feature that reflects a trade-off between speed and accuracy in the learning process (BS, 2002). *(3)* The peak of the hump-shaped profit function, i.e. the optimal firm size, shifts to the right when complexity increases. In other words, optimal size increases in more complex environments.

These facts may be captured by a polynomial in the three variables (in our previous work other variables affected profit, but here we hold them constant):

$$\pi_i = f(m_1, m_2, n) \tag{1}$$

To obtain a specific numerical form for equation 1, we simulated the Cournot learning process with different randomly drawn firm sizes ($m_1, m_2 \in [2, 20]$) and complexity ($n \in [5, 50]$), recording each time the profit of the two firms. With such a data set we ran a regression, that is reported in table 1 in the appendix (notice that the setup is symmetric, so that either firm could be used).

The specific polynomial relating profits to size and complexity, that will serve as the basis for our firm size dynamics, is given in equation (2).

$$
\begin{aligned}
\pi_1 = {} & 287 + 7.53m_1 - 0.81m_1^2 + 0.031m_1^3 - 0.0005m_1^4 + \\
& - 1.166m_2 + \\
& - 2.954n - 0.015n^2 + 0.00029n^3 + \\
& - 0.34(m_1m_2) + 0.001(m_1m_2)^2 - 0.000002(m_1m_2)^3 + \\
& + 0.079m_1n + 0.088m_2n - 0.000035(m_2n)^2 + 0.001(m_1m_2n)
\end{aligned} \tag{2}
$$

These coefficients are plotted to give better insight. In Figure 1 we can see how own and rival's size affects a firm's profits, for a particular environmental complexity ($n = 10$). In general, profit is humped-shape in own size, and in this case is declining in rival's size. That is, as the rival increases size firm 1's profits are lower.

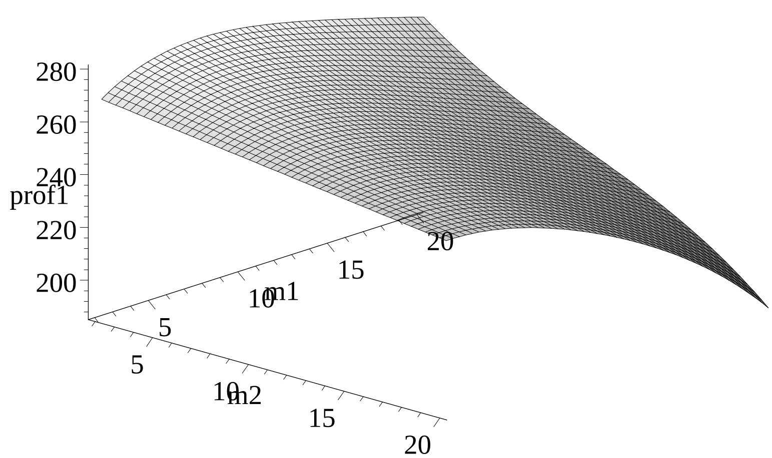

Figure 1: A firm's profit versus network size and rival's network size, $n = 10$.

Figure 2 shows the hump shape of profit with respect to own size. For a particular complexity value, three curves are reported, corresponding to small, medium and large opponent's size.

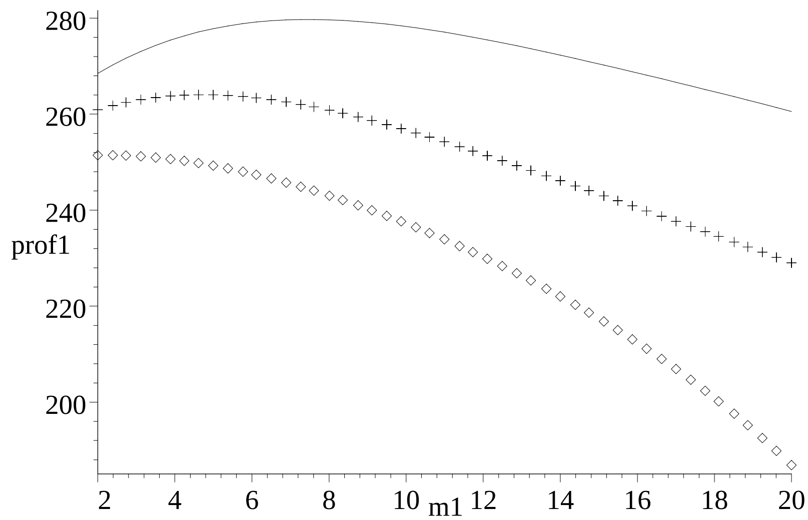

Figure 2: A firm's profit function vs own size ($m_2 = 2$, solid line; $m_2 = 10$, crosses; $m_2 = 10$, diamonds). $n = 10$.

4 Network Size Equilibria

In this section, as a benchmark, we discuss the firms' best response function and equilibria that arise from the game.

4.1 The Best Response Function

Given the functional form for profit, as shown in equation (2), we can derive the best response function by setting the derivative of profit with respect to size equal to zero, i.e., given each choice of rivals network size, there exists (at least one) m_i^*, such that

$$\frac{\partial \pi_i}{\partial m_i} = f' \left(m_i^*, m_{-i}, n\right) = 0 \tag{3}$$

Notice however that the 'best response' function is polynomial in m_i and as a result, there is generally more than one solution, and often some of the solutions are complex numbers. Thus to solve for the best response function, i.e., $f' \left(m_i^*, m_{-i}, n\right) = 0$, we perform the following steps: for each choice of rival's network size $m_{-i} \in [2, 20]$ we solve equation (3), and we discard all but the *positive and real solutions* (in our simulations there is always exactly one positive, real value solution for each m_{-i}). This then gives us a new data set that gives a map between m_{-i} and m_i^*. Figure 3 shows the best response mapping for equation (2). In fact, in spite of the complexity of the profit function, the best response is quasi-linear. Notice further the relationship between the bet responses and environmental complexity: increasing complexity shifts the best response function upward.

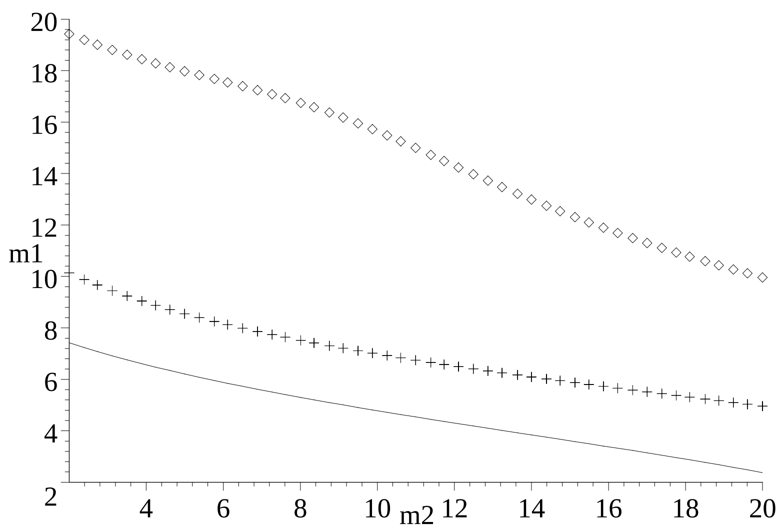

Figure 3: Firm 1 best response functions ($n = 10$, solid line; $n = 25$, crosses; $n = 50$, diamonds)

4.2 Equilibrium

Given the best response functions for each point, we can solve for the equilibrium point. We find that there is always only one equilibrium, as it was to be expected given the quasi-linearity of the best response functions. As we can see in figure 4, the equilibrium firm size is about 5.8 nodes for each firm.[1]

For each complexity level, we are able then to solve for the equilibrium value, i.e., the intersection of each firm's best response function. In Figure

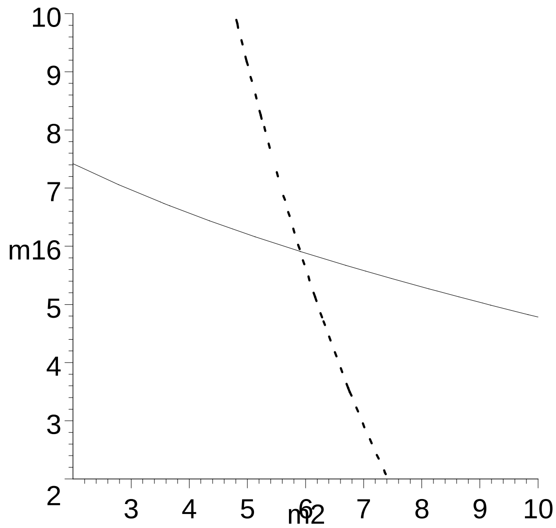

Figure 4: Size equilibrium is determined at the intersection of the two best responses: $m_1 = br(m_2)$ (solid line) and $m_2 = br(m_1)$ (dashed line). $n = 10$.

5 we show the firm size equilibria and profits as a function of complexity. As can be seen, firm size is increasing in complexity, while firm profits are decreasing.

5 Adjustment Dynamics

As discussed above, firms face a large amount of information to process. In standard economic theory, firms are assumed to understand how the world functions. Meaning, they are assumed to know their cost functions, profit functions, and the effect of a rival's decisions on profits. But in a complex world, with boundedly rational agents, the cost to discover such knowledge is relatively high. As a result, firms learn as they engage in their activities and use this knowledge to help guide them.

[1]For simplicity we ignore the integer issue in regards to firm size and assume that firm size can take on any real value greater than zero.

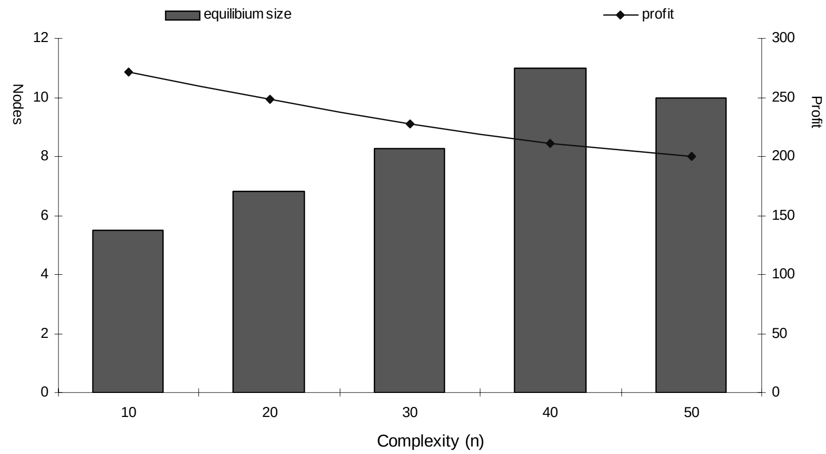

Figure 5: Network size equilbria and profits for each complexity level.

The knowledge that they have is acquired, in general, from following relatively simple rules of behavior and general approximation schemes (Nelson and Winter, 1982; Simon, 1982). That is to say, firms are not necessarily concerned with what makes the world work the way it does, but rather are concerned with learning the effect of general action-reaction type behaviors. For example, two firms producing oil, are presumably not concerned with actual mathematical Cournot functions, so much as they are concerned with learning: 'if my rival produces so much oil and I don't, what range of prices can I expect to charge for my oil?' In other words, the existence of a equilibrium, for example does not lead to the simple assumption that firms will know what this equilibrium is and then select the correct response. But rather, in a world where information processing and knowledge acquirement are expensive, firms may find it easier to behave according to simple rules of actions, e.g., hire more workers if my profit is positive or higher more workers if my profits are growing, etc.

For this reason, in this section, we explore simple dynamics for firm size; dynamics that basically do not assume any greater knowledge on the part of firms then simply being able to observe the effect of changing agents (nodes) on profits. The best response function in section 4.1 is quite complex, and if is assumed that the firm knows it; more specifically, it is assumed to know the expect maximal profit obtainable for the entire range of a rival's choice of network size.[2] (Note also that the best response functions graphed above are numerical approximations) But in a world in which production and market conditions constantly change, past information may quickly become irrelevant. Meaning that even if a firm has perfect knowledge of its best response function at a certain point in time, that function may quickly become outdated. That means that, even when in possession of the computational capabilities that are necessary to compute the best response, a firm may not find it convenient to actually do it.

Using the profit function generated in section 3, we explore adjustment dynamics for firms using rule-of-thumb type adjustment rules. As the level of complexity in determining the equilibrium is quite complex, we assume the following general adjustment dynamics:

$$m_{i,t} = (1-\rho)m_{i,t-1} + \beta\pi_{i,t-1} + \alpha I_i \left[(m_{-i,t-1} - m_{i,t-1})(\pi_{-i,t-1} - \pi_{i,t-1})\right] \quad (4)$$

where $\rho \in [0,1]$ is an exogenous "quit rate" parameter; β represents the sensitivity of firm size to profits; it captures the "trend" growth of the firm, that is linked to profit. The parameter α captures the "imitation" factor behind size adjustment; I_i is an indicator function taking the value of 1 if the opponent's profit is larger than the firm's own one, and a value of 0 if the contrary holds:

$$I_i = \begin{cases} 1 \Leftrightarrow (\pi_{-i,t-1} - \pi_{i,t-1}) > 0 \\ 0 \Leftrightarrow (\pi_{-i,t-1} - \pi_{i,t-1}) \leq 0 \end{cases}$$

In case $I_i = 1$, then size will be adjusted in the direction of the opponents'. Thus, the firm will adjust towards the opponent's size, whenever it observes a better performance of the latter. Notice that if $\rho m_{i,t-1} > \beta\pi_{i,t-1} + I_i \left[\alpha(m_{-i,t-1} - m_{i,t-1})(\pi_{-i,t-1} - \pi_{i,t-1})\right]$, then the firm will reduce its size, *regardless of the imitation effect sign,* whereas if the contrary holds, it will increase it.

We can think of the Cournot game as happening on a short term basis, while the adjustment dynamics occurs on longer periods of time. Basically, our questions is this section are: under what parameter choices using equation (4) will firms reach the equilibrium level presented in Figure 4 and, what kinds of behavior can we expect for firm size as a function of the parameter space?

[2] A common assumption in game theory is that firms know the equilbrium and that they know that their rival knows it, and they know that their rival knows they know it, ad infitinum (Fudenberg and Tirole, 1996).

5.1 Scenario 1: The Basic Case–The Isolationist Firm

Suppose that $\alpha = 0$. Then each firm will only look at its own performance when deciding whether to add or to remove nodes:

$$m_{i,t} = (1-\rho)m_{i,t-1} + \beta\pi_{i,t-1}(m_{i,t-1}, m_{-i,t-1}) \tag{5}$$

Of course, this does not mean that the firm's own dynamics is independent of the other, as in fact $\pi_{i,t-1}$ depends on both sizes at time $t-1$. Suppose further that size is below the Nash equilibrium value, m^*. Then, in order for the algorithm to converge, we need to have that all along the path

$$\pi_{i,t-1}(m_{i,t-1}, m_{-i,t-1}) > \frac{\rho}{\beta} m_{i,t-1},$$

and vice versa if $m_{i,0} > m^*$. The system will settle at the level given by the solution of

$$m_{i,t} = m_{i,t-1} = m^* \Rightarrow m^* = \frac{\beta}{\rho}\pi_i(m^*, m^*)$$

For the 'basic' analysis we assume that the two firms begin with 2 nodes each (holding $n = 10$). This implies that their profits will be the same, and the dynamics will be the same as well. We can then look at only one of them. Further, in this scenario, we place no integer restrictions or value restrictions on the nodes other than the number of nodes is set to 1, if the value becomes less than one. We explore the parameter space for β (holding ρ constant) and see how the number of nodes and profits evolve over the long run, i.e., we explore the limit points. Lastly, we investigate the relationship between long run firm size and complexity (holding β constant.)

5.2 Scenario 2: The Imitationist Firm

In this case the contrary of the isolationist firm holds: each firm does not simply care about its absolute situation, but rather about the comparison with the other. For simplicity suppose that the quit rate and own profit adjustment parameter are zero: $\beta = \rho = 0$. Thus,

$$m_{i,t} = m_{i,t-1} + I_i\left[\alpha(m_{-i,t-1} - m_{i,t-1})(\pi_{-i,t-1} - \pi_{i,t-1})\right].$$

In this case, we also investigate the effect of complexity on long run firm size.

5.3 Results

5.3.1 Scenario 1: The Basic Case

For $n = 10$, and two different values of ρ, the equilibrium is given by figure 6. Quite intuitively, the higher the dropout rate ρ, the lower the equilibrium size.

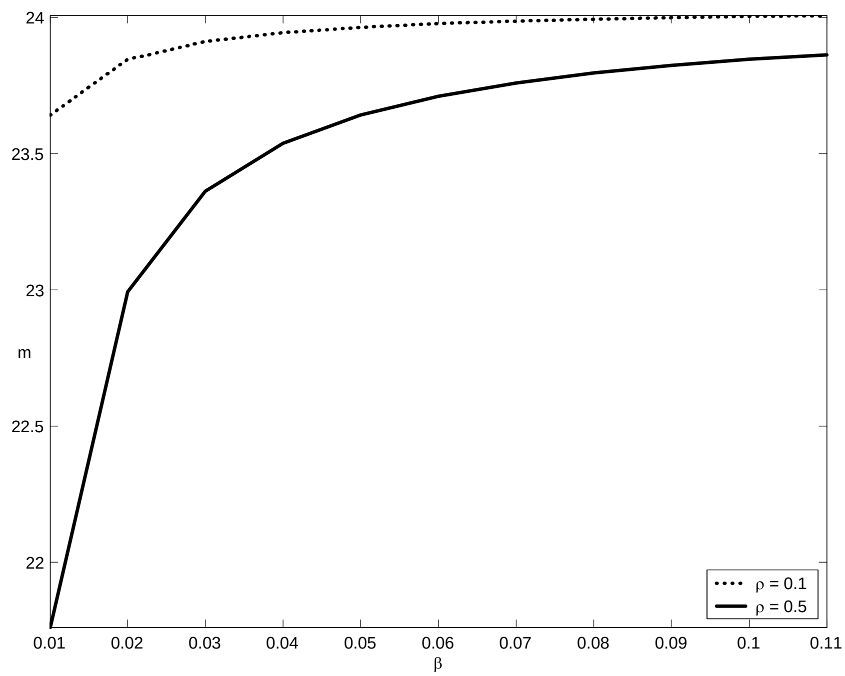

Figure 6: Equilibrium values of $m_1 = m_2$ as a function for β. Complexity is fixed at $n = 10$. Two values for ρ are plotted.

On the other hand, a larger β also implies a larger steady state value. The more reactive firms are to profits, the larger their long run value will be.

Notice that this steady state value is in general not reached by the firms following the isolationist rule. In fact, as shown in figure 7,only for low values of $\beta(= 0.01)$ the convergence is monotonic and to a value larger than the steady state level of figure 6. For larger values, firms sizes begin oscillating, the oscillation being wider the larger the value of β.

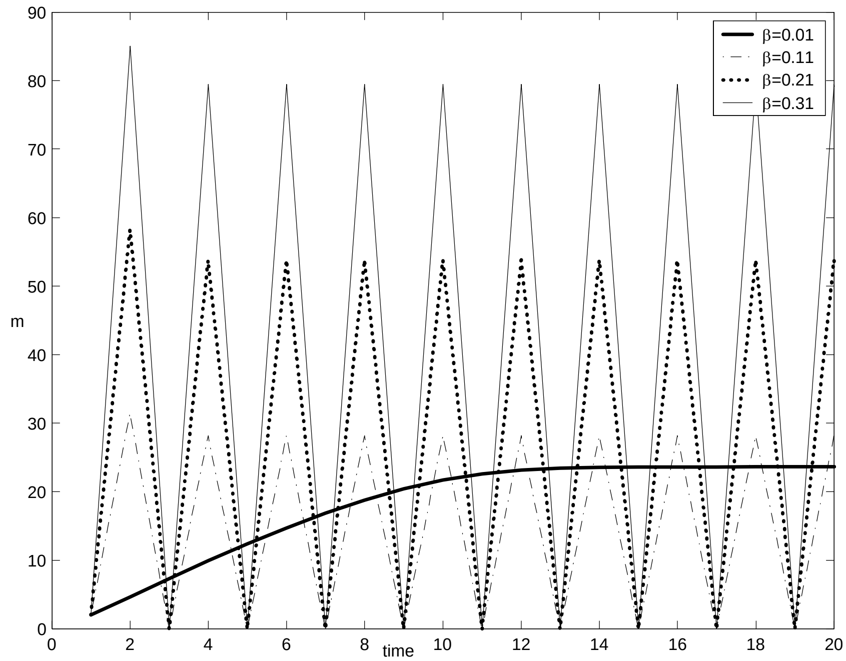

Figure 7: Isolationist firm size dynamics for different values of β.($\rho = 0.01$, and $n = 10$).

To explore the long run nature of the system given by equation (5), we look at the steady state value as a function of β. Interestingly, over a certain range of β we get complex dynamics, including chaos. Figure 8 gives the bifurcation graphs for m_j and figure 9 gives the average profits versus β.

The results show that for a firm to rely on a simply heuristic to choose a level of firm size can result in cyclical and even chaotic behavior over time. Looking at average profits, however, we see that profits are maximized for the firm for a small value of β. And as β increases, profits become increasing volatile and after some point turn negative. Thus relatively large values of β will not be feasible in the long run. The maximum profit occurs when $\beta \in [0.00005, 0.00045]$.

Interestingly, however, the two firms will achieve the Nash equilibrium firm size when at approximately $\beta = 0.004$.

With $\beta = 0.004$ and $\rho = .01$, we investigate the relationship between complexity and firm size and profits. In Figure 10 we present the results of

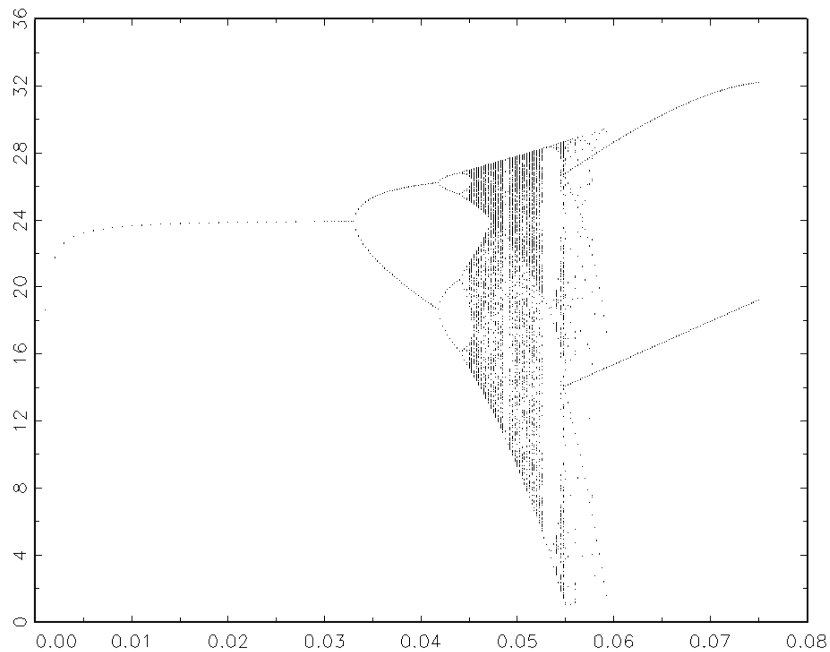

Figure 8: Bifurcation graph: Firm size versus β

this run.

Here we see a slight increase in both firm size and profits as we increase complexity. Increasing firm size with complexity is something we also see in the equilibria benchmark cases, but here we have increasing profits.

5.3.2 Scenario 2: The Imitationist

In this section we present the results from using the following adjustment equation:

$$m_{i,t} = m_{i,t-1} + I_i \left[\alpha(m_{-i,t-1} - m_{i,t-1})(\pi_{-i,t-1} - \pi_{i,t-1}) \right].$$

In this case, the firm will no change its size if it has a larger profit, and it will converge to the other if it has a smaller profit. Thus, we can say a number of things before feeding actual parameter values. First, at each period, only one firm moves. Second, at the final equilibrium, the two profits must be equal, which happens when firm sizes are equal, but not necessarily only in this

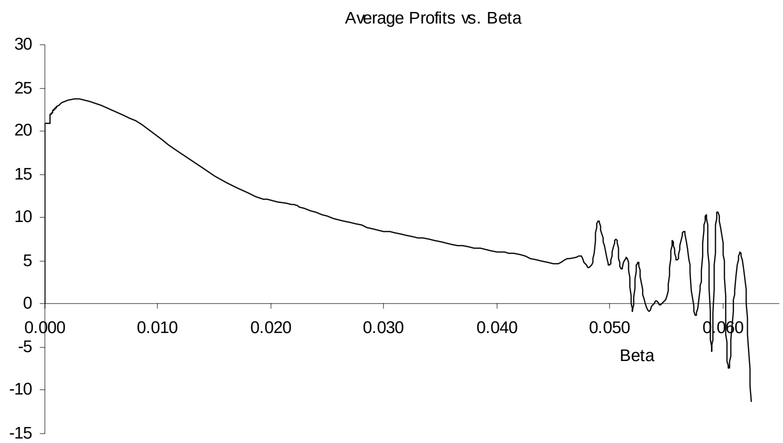

Figure 9: Average profits as a function of β

situation. For $n = 10$, suppose one firm begins small ($m_1 = 2$), and the other large ($m_2 = 15$). The resulting dynamics depend on α, as shown in figure 11 It may be seen that increasing values of α add noise, up to the point where firm sizes fluctuate. On the other hand, a too low value of $\alpha(= 0.01)$ does not allow sufficient growth of the small firm to the point in which it attains a profitable size and forces an adjustment of the competitor. For intermediate values, on the other hand, the two firms converge to similar sizes. Increasing complexity, somewhat changes the situation (figure 12, where $n = 50$). On one side, larger firms are more profitable, so that smaller firms are pushed to change more even for low levels of α. On the other hand, though, the oscillatory region begins earlier, i.e. for lower values of α, with respect to the case of simpler environments.

Lastly we look at the effect of complexity on long run firm size in this scenario. Figure 13 presents the results

Interestingly we see that in this scenario, there is a bifurcation point at about $n = 38$. At that point average profits also become negative (results not shown). Thus we can conclude that the imitationist scenario, given our parameter values is not sustainable in highly complex environments.

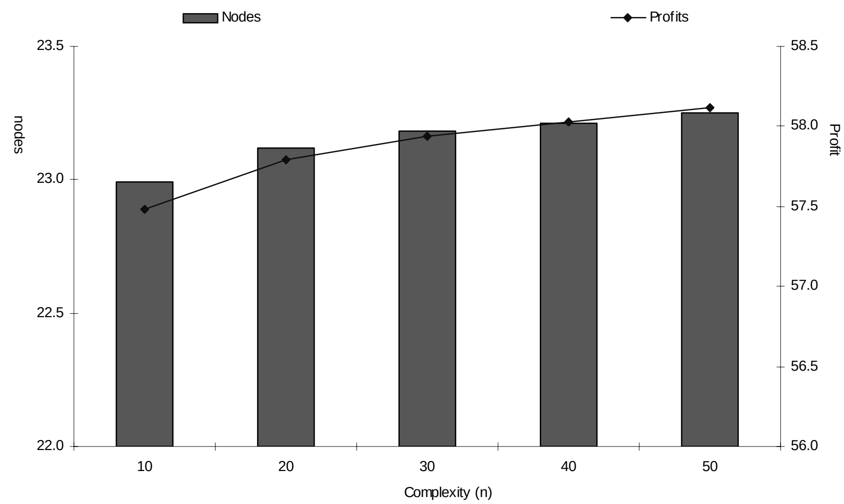

Figure 10: Long Run Firm Size versus Complexity. $\beta = 0.004$, $\rho = 0.01$

6 Conclusion

This paper has presented a model of the firm as an artificial neural network. In prior papers (BS, 2004; forthcoming) we have explored the relationship between firm size, environmental complexity and rival's size in a Cournot game with uncertainty. There, sizes were exogenous parameters; in this paper we explore the endogenous relationship between two firms long-run sizes. Via regression analysis, we relate a firm's profit yielded by the Cournot game, with its own size, its rival's size and the environmental complexity; we find it to be a fourth-order polynomial. Using this profit function, we then generate the firm's best response functions in sizes and related equilibria as a function of environmental complexity. These results are used a benchmark for the core of the paper, which looks at long-run firm size based on two types of dynamic, adaptive rules: the 'isolationist' and the 'imitationist.' If a firm uses an isolationist rule it updates its firm size each period as a function of its last period's size and profits; ignoring it's rival's actions. The imitationist rule has the firm adjusting its size based on last period's size and comparison with the rivals' profits.

Using the relatively simple rules, interestingly, generates some complex behavior. First we find that only under only very precise parameter values does the firm converge to the Nash equilibrium size given by the best re-

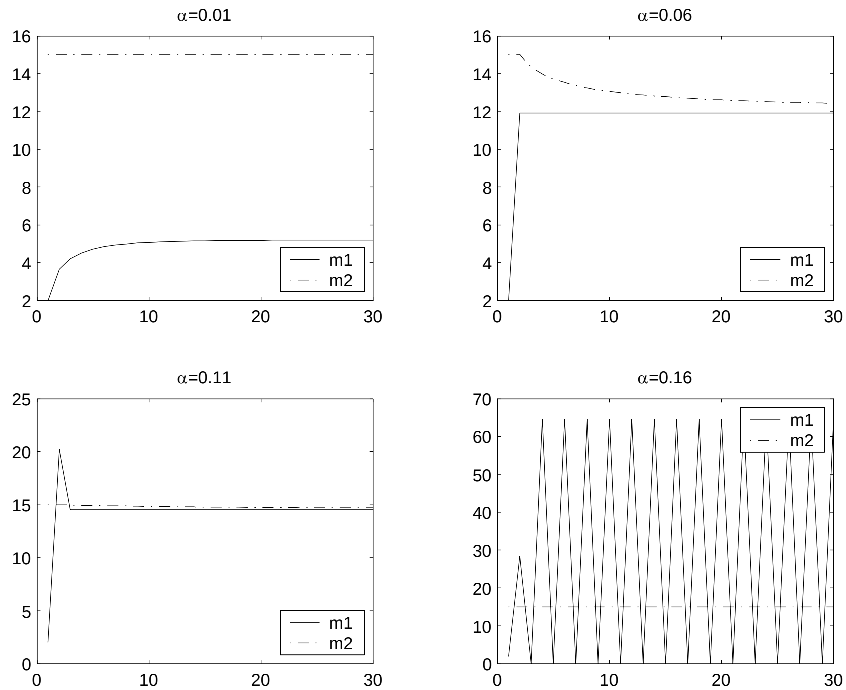

Figure 11: Firm dynamics for different values of α. $n = 10$, $\rho = 0$.

sponsed. And that in general, as a result of the dynamics, the firm's are larger, but also earn lower profits. Furthermore, we find that simple adjustment rules are quite 'parameter sensitive.' In fact, the long run firm size can be either stable, oscillitory or chaotic depending on the values of the adjustment parameters. If firm's use the imitationist rule and start at different sizes, then whether they converge to the same size or not is also a function of the adjustment parameters. Lastly, the relationship between long run firm size and complexity is shown for specific adjustment parameter values. In general firm size (and sometimes profit) is increasing in complexity, but sometimes in a highly non-linear way.

This paper is a first attempt to understand the implications that the interaction of two complex systems (neural networks) can have for simple–boundedly rational–adjustment rules. Since both the behavior of the networks, the corresponding profit functions and the adjustment rules are quite complex, much work still needs to be done on exploring the robustness and

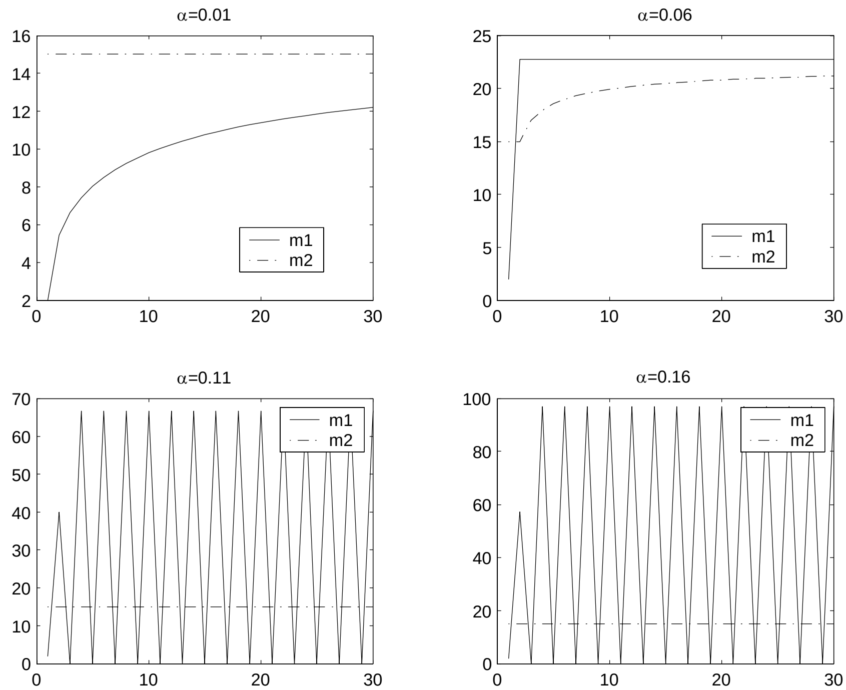

Figure 12: Firm dynamics for different values of α. $n = 50$, $\rho = 0$.

generalizability of our findings. Furthermore, future work will seek to link the findings with these models to stylized facts around firm growth and size (i.e., the literature on Gibrat's Law).

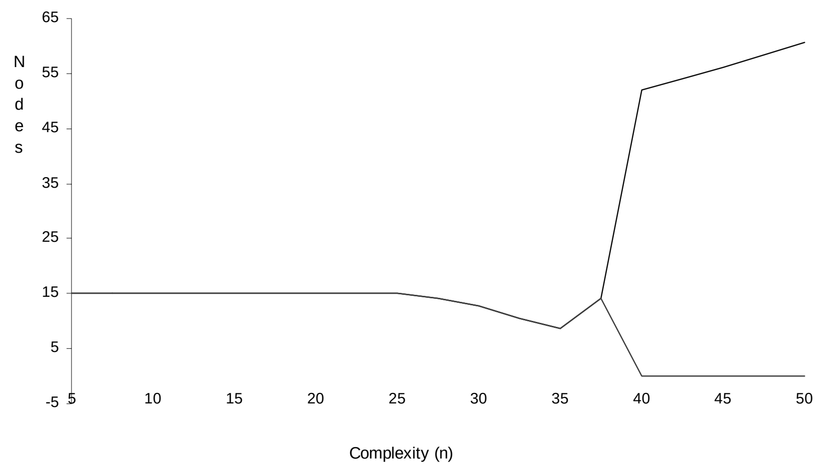

Figure 13: Long Run Firm size versus complexity. $\beta = 0$, $\rho = 0$.

Appendix

A Neural Networks

This appendix briefly describes the working of Artificial Neural Networks. For a more detailed treatment, the reader is referred to Skapura (1996). Neural networks are nonlinear function approximators that can map virtually any function. Their flexibility makes them powerful tools for pattern recognition, classification, and forecasting. The Backward Propagation Network (BPN), that we used in our simulations, is the most popular network architecture. It consists in a vector of $\mathbf{x} \in \mathbb{R}^N$ inputs, and a collection of processing nodes, $\mathbf{m} \in \mathbf{I}^M$, organized in layers.

For our purposes (and for the simulations) we focus on a network with a single layer. Inputs and nodes are connected by weights, $\mathbf{w}^h \in \mathbb{R}^{N\times M}$, that store the knowledge of the network. The nodes are also connected to an output vector $\mathbf{y} \in\mathbb{R}^O$, where O is the number of outputs (2 in our case), and $\mathbf{w}^o \in \mathbb{R}^{M\times O}$ is the weight vector. The learning process takes the form of successive adjustments of the weights, with the objective of minimizing a (squared) error term.[3] Inputs are passed through the neural network to determine an output; this happens through transfer (or squashing) functions, like the sigmoid, to allow for nonlinear transformations. Then supervised learning takes place in the sense that at each iteration the network output is compared with a known correct answer, and weights are adjusted in the direction that reduces the error (the so called 'gradient descent method'). The learning process is stopped once a threshold level for the error has been attained, or a fixed number of iterations has elapsed. Thus, the working of a network (with one hidden layer) may be summarized as follows. The feed forward phase is given by

$$\underset{1\times O}{\hat{\mathbf{y}}} = g\left[g\left(\underset{1\times N}{\mathbf{x}} \cdot \underset{N\times M}{\mathbf{w}^h}\right)\underset{M\times O}{\mathbf{w}^o}\right]$$

where $g(\cdot)$ is the sigmoid function that is applied both to the input to the hidden layer and to the output. And the error vector is:

$$\boldsymbol{\varepsilon} = \left\{(y_j - \hat{y}_j)^2\right\}, j = 1, ..., O$$

[3] The network may be seen as a (nonlinear) regression model. The inputs are the independent variables, the outputs are the dependent variables, and the weights are equivalent to the regression coefficients.

Total error is then calculated:

$$\xi = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = \sum_{j=1}^{o} \varepsilon_j$$

where $\mathbf{y}$ is the true value of the function, corresponding to the input vector $\mathbf{x}$.

This information is then propagated backwards as the weights are adjusted according to the learning algorithm, that aims at minimizing the total error, ξ. The gradient of ξ with respect to the output-layer weights is

$$\frac{\partial \xi}{\partial \mathbf{w}^o} = -2\left(\mathbf{y} - \hat{\mathbf{y}}\right)\left[\hat{\mathbf{y}}(\mathbf{1} - \hat{\mathbf{y}})\right] g\left(\mathbf{x} \cdot \mathbf{w}^h\right),$$

since for the sigmoid function, $\partial \hat{\mathbf{y}} / \partial \mathbf{w}^o = \hat{\mathbf{y}}(\mathbf{1} - \hat{\mathbf{y}})$.

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

$$\frac{\partial \xi}{\partial \mathbf{w}^h} = -2\left[\left(\mathbf{y} - \hat{\mathbf{y}}\right)\left[\hat{\mathbf{y}}(\mathbf{1} - \hat{\mathbf{y}})\right] g\left(\mathbf{x} \cdot \mathbf{w}^h\right)\right] \mathbf{w}^o g'(\mathbf{x} \cdot \mathbf{w}^h)\mathbf{x}.$$

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. Define $\boldsymbol{\delta}^o = .5\left(\mathbf{y} - \hat{\mathbf{y}}\right)\left[\hat{\mathbf{y}}(\mathbf{1} - \hat{\mathbf{y}})\right]$. We then have the weight adjustment for the output layer as

$$\mathbf{w}^o(t+1) = \mathbf{w}^o(t) + \eta \delta^o g\left(\mathbf{x} \cdot \mathbf{w}^h\right)$$

Similarly, for the hidden layer,

$$\mathbf{w}^h(t+1) = \mathbf{w}^h(t) + \eta \boldsymbol{\delta}^h \mathbf{x},$$

where $\boldsymbol{\delta}^h = \mathbf{g}'(\mathbf{x} \cdot \mathbf{w}^h)\boldsymbol{\delta}^{\mathbf{o}}\mathbf{w}^o$. When the updating of weights is finished, the firm views the next input pattern and repeats the weight-update process.

B Derivation of the Profit Function

This appendix briefly describes the process that leads to equations 1 and 2, that in the present paper is left in the shadow. Details on the model can be found in BS (forthcoming).

B.1 Cournot Competition in an Uncertain Environment

We have two Cournot duopolists facing the demand function

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

where α_t changes and is *ex ante* unknown to firms. Assume that production costs are zero. Then, the best response function, were α_t known, would be given by

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

with a Nash Equilibrium of

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

When deciding output, firms do not know α, but have to estimate it. They only know that it depends on a set of observable environmental variables $\mathbf{x} \in \mathbb{R}^N$:

$$\alpha_t = \alpha(\mathbf{x}_t)$$

where $\alpha(\cdot)$ is unknown. Each period, the firm views an environmental vector $\mathbf{x}$ and uses this information to estimate the value of $\alpha\left(\mathbf{x}\right)$.

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

1. Each firm observes an environmental state vector $\mathbf{x}$.

2. Based on that each firm estimates a value of the intercept parameter, $\hat{\alpha}_j$. 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. It then observes the true value of and α, and q_{-j}, and uses this information to determine its errors using the following rules:

$$\varepsilon_{1j} = \left(\hat{\alpha}_j - \alpha\right)^2 \tag{6}$$

$$\varepsilon_{2j} = \left(\hat{q}^j_{-j} - q_{-j}\right)^2 \tag{7}$$

4. Based on these errors, the firm updates the weight values in its network.

To summarize, 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

This process repeats for a number $T = 250$ of iterations. At the end, we can compute the average profit for the two firms as

$$\pi_i = \frac{1}{T} \sum_{t=1}^{T} q_{it}(\alpha_t - (q_{1t} + q_{2t})). \tag{8}$$

C Regression Results for Profit

Equation 2 was derived by using the model described in the preceding appendix. We built a data set by making random draws of $n \in [5, 50]$, $m_i \in [2, 20]$. We ran the Cournot competition process for $T = 250$ iterations (random initial conditions were appropriately taken care of by averaging over multiple runs). We recorded average profit for the two firms computed as in eq. 8, and the values of m_1, m_2, and n. This was repeated 10000 times, in order to obtain a large data set. We then ran a regression to obtain a precise polynomial form for profit as a function of sizes and environmental complexity. Table 1 gives the complete results of the regression, that is reflected in equation 2.

Dependent Variable: $10000 * \pi_1$			
Variable	**Coefficient**	**Std. Error**	**t-Statistic**
$Const$	287.505	1.947	147.64
m_1	7.530	0.674	11.17
m_2	-1.166	0.105	-11.15
n	-2.954	0.122	-24.23
n^2	-0.015	0.004	-3.61
n^3	0.00029	0.00005	6.02
m_1^2	-0.812	0.104	-7.83
m_1^3	0.031	0.007	4.65
m_1^4	-0.0005	0.00015	-3.08
$m_1 * m_2$	-0.340	0.021	-16.24
$(m_1 * m_2)^2$	0.001	0.00009	10.75
$(m_1 * m_2)^3$	-0.000002	0.00000015	-10.35
$m_1 * n$	0.079	0.003	25.81
$m_2 * n$	0.088	0.005	18.34
$(m_2 * n)^2$	-0.000035	0.000003	-12.00
$m_1 * m_2 * n$	0.001	0.0003	3.36
R-squared	0.8799	Mean dependent var	225.44
Adjusted R-squared	0.8797	S.D. dependent var	23.79
Nobs.: 10,000		S.E. of regression	8.25

Table 1: Profit function for firm 1. White Heteroskedasticity-Consistent Standard Errors.

References

[1] Barr, J. and Saraceno, F. (forthcoming). "Cournot Competition, Organization and Learning." *Journal of Economic Dynamics and Control.*

[2] Barr, J. and Saraceno. F. (2004). "Organization, Learning and Cooperation." Rutgers University, Newark Working Paper #2004-001.

[3] Barr, J. and F. Saraceno (2002), "A Computational Theory of the Firm," *Journal of Economic Behavior and Organization*, 49, 345-361.

[4] Chandler, Jr., A. D. (1962). *Strategy and Structure: Chapters in the History of the American Industrial Enterprise.* The MIT Press: Cambridge.

[5] Chang, M-H, and Harrington, J.E. (forthcoming). "Agent-Based Models of Organizations." *Handbook of Computational Economics,* Vol 2. Eds K.L. Judd and L. Tesfatsion.

[6] Cyert. R. M. and March J.G. (1963). *A Behavioral Theory of the Firm.* Prentice-Hall, New Jersey.

[7] Currie, M. and Metcalfe, S. (2001). "Firm Routines, Customer Switching and Market Selection Under Duopoly." *Journal of Evolutionary Economics*, 11, 433-456.

[8] DeCanio, S. J. and W. E. Watkins (1998), "Information Processing and Organizational Structure," *Journal of Economic Behavior and Organization*, 36, 275-294.

[9] Fudenberg, D. and Tirole, J. (1996). *Game Theory.* The MIT Press, Cambridge.

[10] Kopel, M. (1996), "Simple and Complex Adjustment Dynamics in Cournot Duopoly" *Chaos, Soliton & Fractals*, 7, 2031-2048.

[11] Nelson, R. R. and Winter, S. G. (1982). *An Evolutionary Theory of Economic Change.* Belknap Press of Harvard University Press: Cambridge.

[12] Puu, T. (2003). *Attractors, Bifurcations & Chaos: Nonlinear Phenomena in Economics.* 2nd Ed. Springer, New York.

[13] Radner, R. (1993), "The Organization of Decentralized Information Processing," *Econometrica*, 61, 1109-1146.

[14] Simon, H. A. (1982). "Rational Choice and the Structure of the Environment." *Psychology Review.* 63(2): 129-138. Reprinted in: *Behavioral Economics and Business Organizations.* The MIT Press: Cambridge.

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