Sheffield Economic Research Paper Series

SERP Number: 2006007

Michael Dietrich*

**Neural networks and the evolution of firms and industries:
An application to UK SIC34 and SIC72.**

May 2006.

* Department of Economics
University of Sheffield
9 Mappin Street
Sheffield
S1 4DT
United Kingdom
www.shef.ac.uk/economics

Introduction

The basic objectives of this paper are twofold: (1) to show how neural networks might be applied to and used to analyse firms; and (2) to investigate the effect of different 'firm types' on evolutionary processes at the industry level. Applying neural networks to economic, managerial and financial issues is not, in itself, original. In business and finance, neural network analysis is not uncommon (see, for example, Altman et al, 1994; Wilson et al, 1995; Chiang et al, 1996; Wong et al, 1995; Jasic and Wood, 2005). Non-firm applications of neural network analysis within economics are becoming increasingly important, for example: Binner et al (2005) on inflation; Papadas and Hutchinson (2002) on input-output technology; Johnes (2000) on macroeconomic modelling; Franses and Homelen (1998) and Plasmans et al (1998) on exchange rate modelling. But previous published work that has applied a neural network framework to the firm (for example, Delgado et al, 2004) has not developed the analysis using actual data as is done in this paper.

The rest of the discussion is organised as follows. In the next section there is a brief introduction to neural networks and how they might be relevant for the analysis of the firm. Following this the data to be used in this study, i.e. firms from UK SIC34 and SIC72, is introduced and the empirical methodology discussed. This is followed by estimation and interpretation of neural networks for the two sectors. The final substantive section considers simulation results based on different firm types. Three firm types are considered here: (a) unchanged firms; (b) efficiency seeking firms, for which input substitution is possible; and (c) input use variability plus firm specific and variable expansion paths. It is shown that evolutionary characteristics differ for the three firm types and for the two sectors.

What are neural networks?

Artificial neural networks were originally developed by neural scientists to model brain functioning (see, for example, Bishop, 1995). A general principle involved is that we can understand the general functioning of the brain, as inputs of sensory data generating bio-chemical reactions. These reactions activate various interconnected neurons that in turn lead to (in principle) measurable 'output'. But the way in which this general model is understood and applied in specific circumstances to any one person's brain is too complex to be specifiable in terms of a simple functional relationship. For this reason neural scientists have developed neural network analysis to model the interactions inside the brain that are otherwise too complex to be modelled. A basic principle of this paper is that the firm is analogous: it is possible to specify general principles involved with firm activity but the actual functioning of any single firm is too complex to model. Arguably much economics oversimplifies both the general principles and the modelling. By implication we are therefore drawing an analogy between the functioning of the brain and the functioning of the firm. It is not, of course, original to claim that firm organisation is the brain of the firm; previous work based on this principle is Beer (1972).

With respect to the firm a traditional production function maps a vector of inputs into a vector of outputs:

$$X \rightarrow Q$$

A basic neural network adds an unobservable ‘hidden layer’ to this mapping:

$$X \rightarrow H \rightarrow Q$$

H is assumed (in our context) to define organisational functioning. For example Penrose (1959) argues that a simple mapping of factor inputs into outputs oversimplifies the nature of the firm. Instead she suggests that factor inputs are mapped into organisationally specific factor services and the services are mapped into outputs. In terms of neural network analysis these Penrosian factor services are a network’s hidden layer. Alternatively, Nelson and Winter (1982) suggest that a firm is guided by ‘routines’ that are equivalent to a firm’s genetic code and largely control firm activity. These routines would be equivalent to a network’s hidden layer.

Figure 1 here, see end.

With n inputs ($x_1, \ldots x_n$) and a single output (Q) we can specify a simple feed-forward network for a firm, as in figure 1: x_0 = an input bias (to be estimated); $h_1 \ldots h_m$ = m hidden units; h_0 = hidden unit bias. In figure 1 there are no network feed-back linkages, hence the term feed-forward network.

Systems of weights define the strength of the various network linkages. A generalised feed-forward neural network can be specified as follows (Delgado et al, 2004):

$$Q = F\left[\beta_0 + \sum_{j=1}^{m} G\left(\alpha_0 + \sum_{i=1}^{n} x_i \alpha_{ij}\right)\beta_j\right] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [1]$$

α_{ij} = weights from input unit i to hidden unit j

β_j = weights from hidden unit j to output.

We can assume that function F in [1] is linear in logs:

$$q = \beta_0 + \sum_{j=1}^{m} \beta_j g_j \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [1a]$$

$$q = \ln(Q); \qquad g_j = G\left(\alpha_0 + \sum_{i=1}^{n} x_i \alpha_{ij}\right)$$

In addition to estimation effectiveness, formulation [1a] has the advantage that the 'share' of hidden unit j can be readily calculated from the weights involved as $\beta_j / \sum \beta_j$. The relevance of this will be clear from later discussion.

The usual form of function G is a sigmoid defined by a logistic function (see Bishop, 1995):

$$q = \beta_0 + \sum_{j=1}^{m} \beta_j \frac{1}{1 + \exp(-z_j)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [1b]$$

$$z_j = \alpha_0 + \sum_{i=1}^{n} \alpha_{ij} \ln(x_i)$$

In a previous paper (Dietrich, 2005) it is shown that formulation [1b] models and predicts firm sales better than either Cobb-Douglas or trans-log formulations.

Estimating neural networks

Two firm inputs are used here: numbers employed (L) and the value of total assets (K). Firm output is measured by revenue (R). Estimation is undertaken for two sectors: SIC34 (vehicles manufacture) and SIC72 (computer software and services). Data is obtained from the FAME data base for 1998 and 2003, with monetary values for 2003 deflated to 1998 levels using the GDP deflator. The 1998 data is used to estimate neural networks for the two sectors. The 2003 data is used to simulate evolutionary processes. To generate these evolutionary processes two periods of comparable data is required (for reasons outlined below). Hence only firms that exist at both data points are used. This results in N=151 for SIC34 and N=1216 for SIC72. These two sectors are chosen because of their different characteristics, as described in table 1.

Table 1: basic sector characteristics

	NeqH 1998	NeqH 2003	mean K/L 1998	mean K/L 2003	mean revenue growth 1998-03	mean R/L 1998	mean R/L 2003
SIC34	10.17	10.03	89.89	122.00	1.54	111.01	135.95
SIC72	47.29	69.14	106.83	183.27	3.44	165.96	240.58

Note: NeqH is the numbers equivalent Herfindahl index calculated using firm revenue. Revenue growth is defined as R03/R98.

Significant market structure differences exist. In SIC34 approximately ten equivalently sized firms exist, whereas in SIC72 the number of equivalently sized firms is much greater. In addition, average market shares in SIC34 remained effectively constant over the 1998-2003 period but smaller firms have had a relative growth advantage in SIC72, as indicated by the increase in the numbers equivalent Herfindahl index (why this relative growth advantage might have occurred is examined below). Mean capital labour ratios are higher in SIC72. By a simple decomposition is it clear that a capital labour ratio is the product of factor productivities:

$$K/L = (K/R)(R/L)$$

The difference in R/L in the two sectors is indicated in the final columns. It follows that the greater capital intensity in SIC72, with apparently smaller firms on average, is the result of greater labour productivity as with constant K/R a higher R/L will produce higher K/L. Finally, mean revenue growth over the period 1998-03 is 2.23 times higher in SIC72 compared to SIC34. These characteristics are indicative of the different stages of the industry life-cycles of the two sectors.

Characteristically in economics production functions are estimated using frontier techniques, for instance, as adopted here, stochastic frontier methods. But it is not possible to estimate a non-linear stochastic frontier model. The following procedure is therefore adopted:

1. Use a 2nd order approximation of an unknown neural network. This approximation is, of course, a trans-log function. In this regard we can recognise that a generalised neural network is an arbitrarily complete specification of a data set (see White 1989). Hence any trans-log function is an approximation of this complete specification.
2. Estimate a trans-log stochastic frontier model, from which relative firm efficiencies can be derived.
3. Adjust firm revenues to imputed efficient levels and estimate neural networks with non-linear least squares as if all firms are efficient. Here we must recognise that any non-linear estimation is potentially sensitive to the algorithm starting point (Curry and Morgan, 1997) because of possible convergence to local solutions (Athanassopoulos and Curram, 1996).

The trans-log stochastic frontier models are well specified (see appendix one). Neural networks can be effectively estimated for both sectors with two hidden units. The general regression equation for these networks, based on formulation [1b], is

$$R_{EFF}98 = c_1 + c_2/\{1 + \exp[-1*(c_3 + c_4\ln(L_{98}) + c_5\ln(K_{98}))]\}$$

$$+ c_6/\{1 + \exp[-1*(c_7 + c_8\ln(L_{98}) + c_9\ln(K_{98}))]\}$$

where R_{EFF} is the imputed efficient level of firm revenue.

After elimination of insignificant and irrelevant coefficients the estimated neural networks are as reported in table 2

Table 2: Estimated neural networks

	SIC34	SIC72
c_2	6.686 (5.85)	10.905 (17.20)
c_3	2.932 (4.09)	
c_4	1.013 (5.30)	0.475 (8.35)
c_5	-0.623 (-4.74)	-0.256 (-8.87)
c_6	14.860 (3.35)	9.081 (15.11)
c_7	-3.643 (-5.20)	-3.655 (-13.57)
c_8	-0.028 (-0.31)	-0.453 (-7.71)
c_9	0.293 (3.03)	0.644 (12.68)
N	151	1216
R^2(adj)	0.970	0.886

Notes: t statistics in parentheses.

We can interpret these estimates by considering the structure of the two hidden units.

Hidden unit one: $c_2/\{1 + \exp[-1*(c_3 + c_4\ln(L_{98}) + c_5\ln(K_{98}))]\}$

$\partial R/\partial L > 0, \qquad \partial R/\partial K < 0.$

Hidden unit two: $c_6/\{1 + \exp[-1*(c_7 + c_8\ln(L_{98}) + c_9\ln(K_{98}))]\}$

$\partial R/\partial L \leq 0, \qquad \partial R/\partial K > 0.$

For hidden unit one any change in K is economically meaningless. An increase in K producing a reduction in R is commercially irrational. A reduction in K producing an increase in R is illogical given the frontier estimation method used here. It follows that hidden unit one defines the effect on R of changes in L with a given K. The significant coefficient on K (i.e. c_5) indicates the nature of the production function, i.e. how labour productivity varies with differing K/L. For hidden unit two an equivalent interpretation is possible. This indicates the effect on R of changing K with a given L, with the negative coefficient on L (i.e. estimated c_8) indicating the nature of the production function.

Figure 2 here, see end

With this interpretation of the two hidden units we can present the estimated neural networks in the manner presented in figure 2. For a firm with a capital labour ratio K_1/L_1 the two hidden units define the rays HU1 and HU2. The actual expansion path of the firm, i.e. the relative significance of these two orthogonal rays, is indicated by the relative shares of the hidden units as indicated by the network weights c_2 and c_6. For SIC34 $c_6/c_2 = 2.22$, for SIC72 the ratio is 0.83. It follows that firms in the vehicles manufacturing sector (SIC34) have a more capital intensive average expansion path than those in computer software and services (SIC72); a perhaps not surprising result.

Neural network simulations

The neural networks presented above can be used to simulate the evolutionary processes that characterise the two sectors. For all simulations period 0 involves actual 1998 data and period 1 uses actual 2003 data i.e. each simulated period involves an equivalent time period of five years. Three sets of simulations are presented here that embody different assumptions about the firm types involved.

Simulation one assumes unchanged firms. It is not being claimed that this is a necessarily realistic view of firm functioning, but instead it is a useful base with which to compare other firm types. Unchanged firms, in the current context involves K/L for each firm being fixed at 1998 levels and the expansion path for each firm being fixed as the estimated average for the sector involved. In terms of figure 2, expansion or contraction occurs along a given ray through a fixed point. The key feature of these first simulations is that non-viable firms are assumed to exit from the sector. The estimation of firm viability, and profitability, is set out in detail below. In each period firm profits are calculated. If firm profits are negative, inputs are adjusted to generate zero profit, using a constant K/L. If positive profit is not possible at any positive input levels the firm is deemed non-viable and exit occurs.

The general programming logic can be set out as follows. The 1998-03 (i.e. period 0 to 1) growth in firm turnover is calculated. The same input growth occurs over the period 1-2. The economic logic with this procedure is that 'unchanged firms', as defined here, are only consistent with actual constant returns to scale or that firms are ignorant of their production functions and so operate as if there are constant returns. In either case factor productivities can be assumed constant and input growth or decline is a simple scaling up or down in line with revenue growth. With this input growth, period 1-2 turnover growth can be calculated. In turn this allows period 2-3 input growth to be calculated, etc.

Simulation two can be characterised as being based on efficiency seeking firms. This involves each firm in each period varying K/L to increase profits. But with this new K/L the expansion path is based on the neural network weights c_2 and c_6. In economic terms we can view these efficiency seeking firms as Marshallian (1881). In his writing the role of the entrepreneur involved discovery of the firm's production function by trial and error. But the firm was representative of the competitive industry, which in the current context implies an average and exogenous expansion path. Apart from this variation in K/L the general programming logic is the same as for unchanged firms, including the assessment of viability.

Simulation three is based on firms with, what can be called, strategic flexibility. This involves efficiency seeking, as in simulation two, plus firm specific expansion paths that can be changed in each period. In economic terms we might view this ability to shift expansion paths as a limited form of Schumpeterian (1976) change. In neural network terminology this involves network learning. Apart from this network learning, firm exit responds to non-viability as with the simulations one and two.

The estimation of firm specific expansion paths involves calculation of firm specific hidden unit weights. For the k'th firm, and for a two hidden unit network, we define the share of hidden unit one as s^1_k, and hence the share of the second hidden unit is $s^2_k = 1 - s^1_k$. This allows us to define the slope of the k'th firm's expansion path, in a diagram such as figure 2, as s^2_k/s^1_k. The value of these shares can be estimated assuming that the sum of the hidden unit weights is the same for each firm, interpreted as each firm having the same organisational effectiveness. For firm k we can specify the network as follows:

$$\ln(R_{EFFk}) = s_k^1 \frac{c_2 + c_6}{c_2} HU_{1k} + s_k^2 \frac{c_2 + c_6}{c_6} HU_{2k} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [2]$$

$HU_1 = c_2/\{1 + \exp[-1*(c_3 + c_4\ln(L_k) + c_5\ln(K_k))]\}$

$HU_2 = c_6/\{1 + \exp[-1*(c_7 + c_8\ln(L_k) + c_9\ln(K_k))]\}$

If firm k is in SIC34 and has $s^2_k/s^1_k = 2.22$, formulation [2] collapses back to the earlier average neural network for SIC34. If a firm has an expansion path with a fixed capital stock, i.e. $s^1_k = 1$, the network weight for hidden unit one is increased by $(c_2+c_6)/c_2$. For a firm that is expanding with a fixed labour force, i.e. $s^2_k =1$, the network weight for hidden unit two is increased by $(c_2 + c_6)/c_6$. In this sense all firms have the same organisational effectiveness no matter what their characteristics in terms of hidden unit shares.

To simplify notation we can re-write [2] as

$$\ln(R_{EFFk}) = s_k^1 x_k^1 + s_k^2 x_k^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [2a]$$

$$x_k^1 = \frac{c_2 + c_6}{c_2} HU_{1k}; \qquad x_k^2 = \frac{c_2 + c_6}{c_6} HU_{2k}$$

As $s_k^1 = 1 - s_k^2$, [2a] can be re-written as

$$\ln(R_{EFFk}) - x_k^1 = s_k^2(x_k^2 - x_k^1) \quad \text{..................................[2b]}$$

and in turn

$$s_k^2 = \frac{\ln\left(R_{EFFk}\right) - x_k^1}{x_k^2 - x_k^1} \quad \text{..[2c]}$$

In [2c] the right hand side is calculable, based on direct observation or use of parameter estimates. It follows that 1998 firm specific hidden unit shares can be calculated.

The change in s^2_k, over the period 1998-2003 is modelled using a partial adjustment mechanism plus a firm specific shock that is intended to track the influence of market pressures. The partial adjustment mechanism for a shift in s^2 for firm k is modelled as

$$s_{k,t+1}^2 = s_{k,t}^2 + d\left(s_k^{*2} - s_{k,t}^2\right) = ds_k^{*2} + (1-d)s_{k,t}^2$$

where s_k^{*2} = the desired level of s_k^2, and d = the partial adjustment coefficient. The average desired level of s^2 for all firms (s^{*2}) can be readily estimated using least squares regression, as set out below. The calculation of the firm specific equivalent is considered shortly.

The firm specific shock is calculated using a time specific multiplier ($m_{k,t}$) that is based on the ratio of firm sales growth, over the period t-1 to t, ($g_{k,t}$) compared to average sales growth for all firms over the same period (gav_t). This multiplier is used in two ways: to allow for deviations from planned adjustment in s^2 because of external shocks and to allow for changes in desired s^2 in the light of external shocks. The multiplier is specified as:

$$m_{k,t} = \left(\frac{g_{k,t}}{gav_t} \right)^e$$

The adjustment process and shock are combined as follows:

$$s^2_{k,t+1} = m_{k,t}\left[ds^{*2} + (1-d)s^2_{k,t}\right] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..[3]$$

Formulation [3] can be substituted into [2b] specified for period t+1:

$$\ln(R_{EFFk,t+1}) - x^1_{k,t+1} = (x^2_{k,t+1} - x^1_{k,t+1})m_{k,t}[ds^{*2} + (1-d)s^2_{k,t}] \ldots\ldots\ldots[3a]$$

Equation [3a] can be readily estimated using non-linear least squares. Results are reported in appendix two and summarised in table 3. Given the parameter estimates, firm specific s^{*2} can be calculated as

$$s^{*2}_k = \frac{\ln(R_{EFFk,t+1}) - x^1_{k,t+1}}{\left(x^2_{k,t+1} - x^1_{k,t+1}\right)dm_{k,t}} - \frac{(1-d)}{d}s^2_{k,t}$$

<u>Table 3: Adjustment process model parameter estimates</u>

	d	e	s^{*2}	MSD learning network	MSD basic network
SIC34	0.694	-0.025	0.683	4046.49	206895.11
SIC72	0.920	0.060	0.434	34746.65	62171.45

Note: MSD is the mean squared deviation of actual from predicted 2003 firm sales.

As can be seen from table 3 the predictive accuracy of a network with the adjustment process is significantly improved in both sectors compared to a basic network. In terms of model parameters, the partial adjustment coefficient (d) suggests that for SIC72 firms almost complete adjustment to desired expansion paths is achieved over the 5 year interval 1998-2003, but for SIC34 firms the ability to shift expansion paths is more constrained. The shock parameter (e) indicates that in the vehicles sector lower firm growth than the sector average leads to a shift towards more labour intensive production, whereas for software companies the same effect leads to greater capital intensity. This difference may reflect input market constraints and technology differences for the two sectors. Finally the value of s^{*2} implies a desired expansion path, as defined in terms of figure 2 above, of K/L = 2.15 for SIC34 firms and K/L = 0.77. This indicates marginal long-run shifts towards greater labour intensity in both sectors.

Firm viability and profitability are assessed in the same way in all three sets of simulations. Viability is defined in terms of an ability to avoid losses at some input use and output level. To operationalise this definition a zero profit condition for each firm must be used. Such a condition is derived assuming an unknown distribution of firm profitability across the actual (not simulated) population of firms, but that a subset of the actual firms is operating at zero profit. These marginal, zero profit, firms are assumed to define a lower bound for the population. Hence a zero profit condition can be derived by estimating this lower bound. To derive this bound we can recognise that for any one firm using labour and capital inputs

$$\pi = R - p_L L - p_k K$$

i.e. $$R/L = \pi/L + p_L + p_K(K/L),$$

where: π = firm profit; p_L and p_K = input prices.

It follows that for a zero profit, marginal firm

$$R/L = p_L + p_K(K/L)$$

and for a positive profit, non-marginal firm

$$R/L > p_L + p_K(K/L).$$

It is assumed that p_L and p_K are the same for all marginal firms and that labour and capital prices are no lower for non-marginal than marginal firms.

To derive the lower bound the firms are ranked in ascending order of their degree of technical efficiency, as derived from the frontier estimates reported in appendix 1. The marginal firms are assumed to be the least efficient. The following procedure is used to estimate the lower bound viability condition:

1. Rank the series R/L and K/L in terms of the degree of technical inefficiency.
2. Starting with the first 10 observations of the ranked series identified in (1), formulate the regression equation:

   $$(R/L) = \beta_0 + \beta_1(K/L).$$

3. Add successive observations from the ranked series and re-estimate the same equation. The equation that minimises the estimate of β_1 is assumed to identify the lower bound.

This procedure results in the lower bound being identified by 12 firms in SIC34 and 70 firms in SIC72. The estimates of β_0 and β_1 are also used to simulate firm profitability. Regression results from the use of the procedure are reported in appendix 3.

Neural network simulation results

For all simulations reported below:

Gav = average growth rate for the sector;

RoSav = average return on sales for the sector;

RoSmin = minimum return on sales for the sector;

NeqH = numbers equivalent Herfindahl index, measured on the right hand axis.

<u>Simulation one</u>: fixed K/L, single sector expansion path.

These base simulations are shown in figures 3a and 3b. The key characteristics involved can be summarised as follows. For SIC34 average growth is initially (approximately) constant and reaches an approximate equilibrium after 20 periods; this is equivalent to100 years. Average profitability has a small rise as non-viable firms exit. The exit of non-viable firms results in profits of marginal firms eventually reaching (approximately) zero. Market structure, as indicated by the numbers equivalent Herfindahl index, cycles for 15 periods, following which there is an increasing trend that converges on 22 equivalently sized firms. This long-run trend indicates that the largest firms have long-run relative growth disadvantages but in the short and medium terms an equivalent conclusion cannot be drawn.

For SIC72 average growth exhibits an initial decline and following this reaches an approximate steady state after 20 periods. Average profitability indicates a marked increase with the exit of marginal firms. This firm shake-out requires 17 periods to be completed. With regard to market structure, the initial cycle is less enduring than for SIC34. The steady state number of equivalently sized firms shows an increase compared to the actual 1998 and 2003 figures reported above in table 1. This indicates that the largest firms have a continuing growth disadvantages, if we assume unchanged firms.

Figure 3a here, see end.

Figure 3b here, see end

<u>Simulation two</u>: variable K/L, single sector expansion path.

The results of these second simulations are reported in figures 4a and 4b and depict what might be called efficiency seeking firm behaviour. For SIC34 short-run average growth is higher than with fixed capital intensity i.e. factor substitution is a clear contributor to industry growth. Short-run profitability exhibits a similar trend as fixed K/L but converges on a higher level reflecting the efficiency advantages. The characteristics of marginal firms shows a cycling between viability and non-viability with eventual convergence on clear positive profitability. This steady state characteristic indicates the potential importance of long-run firm entry in this sector, if this responds to marginal firm profitability. With regard to market structure effects in figure 4a, while a similar general trend is evident as compared to figure 3a, there is a marginal increase in the number of equivalently sized firms in the short, medium and long terms. The approximate steady state is 27 firms, compared to 22 with fixed capital intensity. This difference implies that factor substitution offers a growth advantage to smaller firms.

For SIC72 factor substitution generates a minor growth advantage in the short term when figures 3b and 4b are compared. The average and minimum profitability advantages also indicate only minor long-run increases. The clearest difference, when variable and fixed capital intensity are compared, is with market structure. Factor substitution generates a clear increase in the number of equivalently sized firms.

Figure 4a here, see end

Figure 4b here, see end

<u>Simulation three</u>: variable K/L, firm specific and adapting expansion paths.

These final simulations, reported in figures 5a and 5b, are based on factor substitutability and evolving firm specific expansion paths, denoted here as network learning. This possibility might be called factor and strategic flexibility. Note that in economic theoretical terms this flexibility is not a scale economy effect. For SIC34 firms, factor and strategic flexibility produces a significant short-run average growth spike compared to simulation two. In addition, steady state growth is not achieved, even in an approximate sense, after the 20 periods shown here. The growth spike is mirrored by an average profitability spike. But long-run average profitability shows no real change compared to figure 4a. This similar long-run profitability in figures 4a and 5a suggests that strategic flexibility has short-medium term profit advantages, but does not influence long-run profitability. With regard to minimum return on sales, the combination of factor and strategic flexibility does not appear to improve the ability to survive, when figure 5a and 4a are compared. This may reflect the market shock that is part of the expansion learning process in simulation three. For SIC34 market structure, there is apparently a greater degree of short and medium term variability. In addition, the general number of equivalently sized firms is lower than in figures 4a and 3a. This indicates that the strategic flexibility of simulation three benefits relatively larger companies.

For SIC72 firms, in figure 5b, similar differences appear to exist when these simulation results are compared to figures 3b and 4b, as are apparent with SIC34. There is a significant short-run growth spike that is even greater than in SIC34. But the related profitability change hardly appears to occur. In addition, following the rapid decline in average growth there is a significant average profit decline. Marginal firms appear not to be affected by strategic flexibility. Finally, market structure shows a greater degree of volatility and, as with SIC34, has a long-run trend with fewer equivalently sized firms. This indicates that in both sectors strategic flexibility operates to the relative advantage of larger firms.

Figure 5a here, see end

Figure 5b here, see end

Table 4: Summary measures of non-steady state simulation results.
(Cumulative difference compared to simulation one)

	Simulation two				Simulation three		
SIC34	Gav	RoSav	NeqH		Gav	RoSav	NeqH
Periods 0-5	1.008	0.442	4.501 (2.98)		3.274	4.426	-16.572 (-10.97)
Periods 0-10	1.055	0.912	17.796 (11.79)		3.849	4.998	-28.002 (-18.54)
SIC72							
Periods 0-5	0.549	0.528	131.662 (10.83)		11.998	-2.637	-84.177 (-6.92)
Periods 0-10	0.526	0.868	264.043 (21.71)		11.935	-1.790	-209.590 (-17.24)

Note: Figures in parentheses are the changes in the numbers equivalent index as a percentage of the number of actual firms i.e. 151 for SIC34 and 1216 for SIC72.

The observations based on the graphical results are summarised in table 4. It is clear that growth advantages from factor substitution (i.e. simulation two) are more significant in SIC34 compared to SIC72. This is perhaps a logical result that reflects the differing technologies involved. The table confirms that simulation three results in significant growth advantages, particularly in SIC72. These growth advantages are much greater than when factor substitution alone is allowed i.e. with simulation two. For SIC34 simulation three also indicates significant profitability gains that are small with simulation two. But for SIC72 simulation three generates profit disadvantages, even thought there are significant growth gains. A possible explanation here is the way in which the competitive environment is modelled and its impact on the evolution of firm expansion paths. These results are consistent with SIC72 having more competitive markets than SIC34.

Finally, table 4 indicates the way in which the two simulations generate different market structure implications, as summarised by the numbers equivalent Herfindhal index. Factor substitution alone (i.e. simulation two) increases the number of equivalently sized firms. In absolute and relative terms these increases are greater in SIC72 compared to SIC34 indicating that efficiency gains arising from factor substitution offer greater advantages to relatively smaller firms in this sector. But with simulation three, that incorporates both factor and strategic flexibility, there is a clear reduction in the number of equivalently sized firms in both sectors i.e. larger firms are able to exploit advantages of general flexibility to a greater extent than smaller firms. We can link these findings back to the general sector characteristics described above in table 1. It was reported above that in SIC72 there was a significant increase in the numbers equivalent Herfindahl index between 1998 and 2003. Based on simulation results, we might account for this finding by suggesting that strategic change is relatively unimportant compared to efficiency seeking behaviour. But in SIC34 we found earlier that the numbers equivalent index was effectively constant over the period 1998 to 2003. This might be explained either in terms of no change in the population of firms (i.e. the equivalent of simulation one) or that the effects of efficiency seeking and strategic change have cancelled each other in terms of the firm size effects.

Conclusion

This paper has considered how neural networks might be used to analyse firm activity and the evolution of industries. It has been shown that neural networks can be effectively estimated using firm data and that the resulting estimates have an intuitive interpretation that is based on an average estimated expansion path for the firms concerned. Using this interpretation it is possible to impute different firm 'types' and analyse the implications involved. The three firm 'types' considered here are (a) unchanged firms; (b) firms for which input substitution is possible (denoted as efficiency seeking behaviour); and (c) firms that are able to exploit input use variability plus firm specific and variable expansion paths (denoted as factor and strategic flexibility). Expansion path variability is modelled effectively in terms of a partial adjustment mechanism plus a firm specific shock that impacts on the adjustment process and the long-run desired expansion characteristics. It is shown that this type (c) model has significant forecasting advantages compared to a basic neural network.

The key findings from the simulation results can be summarised as follows. While efficiency seeking behaviour has growth advantages, compared to unchanged firms, these are small compared to the growth advantages that are displayed with type (c) firms. In addition the two sectors analysed here (UK SIC34 and SIC72) show different profit implications of these growth advantages. In SIC34 an increase in firm growth caused by strategic flexibility coincides with an increase in profitability, whereas in SIC72 the increase in firm growth coincides with a profitability reduction. This difference is explained in terms of the differing market structures in the two sectors along with the differing effects of market shocks. Finally the market structure effects of differing firm types have been analysed. It is shown that factor flexibility generates relative growth advantages that benefit smaller firms. But strategic flexibility generates relative growth effects that benefit larger firms.

The neural network framework used here, while displaying key comparative advantages in terms of the analysis of the firm, can obviously be developed beyond the current simple structure. Two possible improvements would seem to be relevant. First, networks can be estimated with more than two inputs. An obvious possibility here is to separate labour and capital inputs that are operational compared to managerial. In principle this might provide more explicit insights into the organisational characteristics of firms. Secondly, more complex neural networks might be relevant. Two possibilities would seem to be appropriate here: feed-forward networks with more than one hidden layer, and networks that contain feed-back effects. Among other things this might allow, for example, the endogenisation of relative firm efficiencies given output and competitive characteristics. Such developments might yield important insights into firm behaviour but are topics for future research.

Appendix One

Trans-log stochastic frontier estimates

Trans-log stochastic frontier estimates were undertaken using 1998 data. Using a standard formulation, the total error term is divided into two elements: a random part (v) and the degree of technical inefficiency (u). The random element has the usual characteristics of being independently N(0, σ_v^2) distributed over the observations. The inefficiency element is assumed independently half-normal $N^+(0, \sigma_u^2)$ distributed. Figures in parentheses are z statistics.

SIC34: dependent variable ln(R)

cons	6.421 (6.17)
lnL	0.379 (0.99)
lnK	-0.195 (-0.50)
(lnL)^2	-0.246 (-5.32)
(lnK)^2	-0.044 (-1.17)
lnL*lnK	0.287 (3.77)
Wald chi^2	2954.25
Log likelihood	-66.822
$\ln\sigma^2_v$	-2.232 (-7.00)
$\ln\sigma^2_u$	-2.343 (-2.42)

SIC72: dependent variable ln(R)

cons	3.683 (9.52)
lnL	0.394 (3.29)
lnK	0.355 (2.76)
(lnL)^2	0.027 (1.67)
(lnK)^2	0.031 (2.46)
lnL*lnK	-0.043 (-1.68)
Wald chi^2	6085.15
Log likelihood	-1114.7622
$\ln\sigma^2_v$	-1.675 (-18.15)
$\ln\sigma^2_u$	-.654 (-5.98)

Appendix Two

Adjustment process model regression estimates

The general adjustment model, as set out in the text, is

$$\ln(R_{EFFk,t+1}) - x^1_{k,t+1} = (x^2_{k,t+1} - x^1_{k,t+1})m_{k,t}[ds^{*2} + (1-d)s^2_{k,t}]$$

$$x^1_k = \frac{c_2 + c_6}{c_2}HU_{1k}; \qquad x^2_k = \frac{c_2 + c_6}{c_6}HU_{2k}; \; m_{k,t} = \left(\frac{g_{k,t}}{gav_t}\right)^e$$

This implies the following general non-linear regression model

$$y_{t+1} = (x^2_{k,t+1} - x^1_{k,t+1})\left(\frac{g_{k,t}}{gav_t}\right)^{\alpha 3}\left[\alpha 1 + (1-\alpha 2)s^2_{k,t}\right]$$

where α1, α2, and α3 are coefficients to be estimated.

All model parameters can be identified: d = α2, e = α3, s^{*2} = α1/α2.

With period t being 1998 and t+1 being 2003 results are as follows

<u>SIC34</u>

Coefficient	Estimate	t statistic
α1	0.474	159.21
α2	0.694	161.33
α3	-0.025	-6.30
N	151	
adj R^2	0.939	

SIC72

Coefficient	Estimate	t statistic
$\alpha 1$	0.399	91.14
$\alpha 2$	0.920	114.08
$\alpha 3$	0.060	12.98
N	1215	
adj R^2	0.906	

It follows that the model parameters, as reported in the text, are

SIC34: d = 0.694; e = -0.025; s^{*2} = 0.683.

SIC72: d = 0.920; e = 0.060; s^{*2} = 0.434.

Appendix 3

Viability condition regression estimates

In the text the rationale for using the viability condition $R/L = p_L + p_K(K/L)$ was set out. This appendix reports regression estimates for this condition using 1998 data in SIC34 and SIC72. The two series R/L and K/L are ranked by technical inefficiency. Increasing the number of observations from the ranked series allows successive regression equations of the form

$$(R/L) = \beta_0 + \beta_1(K/L)$$

to be computed. The regression with the lowest estimated value for β_1 is assumed to define the lower bound i.e. the firm viability condition. The resulting lower bound regressions for the two sectors are as follows.

SIC34

Coefficient	Estimate	t statistic
β_0	30.859	6.51
β_1	0.186	5.48
N	12	
adj R^2	0.725	

SIC72

Coefficient	Estimate	t statistic
β_0	55.974	4.64
β_1	0.226	2.70
N	70	
adj R^2	0.932	

References

Altman EI, Marco G, Varetto FC (1994), "Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks (the Italian experience)", Journal of Banking and Finance, 18, 505-529.

Athanassopoulos AD, Curram SP (1996), "A comparison of data envelopment analysis and artificial neural networks as tolls for assessing the efficiency of decision making units", Journal of the Operational Research Society, 47, 1000-16.

Beer S (1972), Brain of the firm: the managerial cybernetics of organization, London: Allen Lane.

Binner JM, Bissoondeeal RK, Elger T, Gazely AM, Mullineux AW (2005), "A comparison of linear forecasting models and neural networks: an application to Euro inflation and Euro Divisia", Applied Economics, 37(6), 665-680.

Bishop CM (1995), Neural networks for pattern recognition, Oxford, Clarendon Press.

Chiang WC, Urban TL, Baldridge GW (1996), "A neural network approach to mutual fund asset value forecasting", OMEGA, 24, 205-16.

Curry B, Morgan P (1997), "Neural networks: a need for caution", OMEGA, 25, 123-33.

Delgado FJ, Valiño A, Santín D (2004), "The measurement of technical efficiency: a neural network approach", Applied Economics, 36(6), 627-635.

Dietrich M (2005), "Using simple neural networks to analyse firm activity", Sheffield Economic Research Papers, No 2005014.

Franses PH, Homelen Pv (1998), "On forecasting exchange rates using neural networks", Applied Financial Economics, 8 (6), 589-596.

Jasic T, Wood D (2004), "The profitability of daily stock market indices trades based on neural network predictions: case study for the S&P 500, the DAX, the TOPIX and the FTSE in the period 1965-1999", Applied Financial Economics, 14 (4), 285-297.

Johnes G (2000), Up Around the Bend: linear and nonlinear models of the UK economy compared”, International Review of Applied Economics, 14 (4), 485-493.

Marshall A (1881), The Economics of Industry, London, Macmillan, 2nd edition.

Nelson RR and Winter SG (1982), An evolutionary theory of economic change, London & Cambridge Mass, Belknap Press.

Papadas CT, Hutchinson WG (2002), “Neural network forecasts of input–output technology”, Applied Economics, 34 (13), 1607-1615.

Penrose ET (1959), The theory of the growth of the firm, Oxford, Blackwell

Plasmans J, Verkooijen W, Daniels H (1998), “Estimating structural exchange rate models by artificial neural networks”, Applied Financial Economics, 8 (5), 541-551.

Schumpeter JA (1976), Capitalism, socialism and democracy, London, Allen & Unwin.

White H (1989), “Learning in artificial neural networks: a statistical perspective”, Neural Computation, 1, 425-64.

Wilson N, Chong KS, Peel MJ (1995), “Neural network simulation and the prediction of corporate outcomes”, International Journal of the Economics of Business, 2, 31-50.

Wilson N, Chong KS, Peel MJ (1995), “Neural network simulation and the prediction of corporate outcomes”, International Journal of the Economics of Business, 2, 31-50.

Wong BK, Bodnovich TAE, Selvi Y (1995), “A bibliography of neural network business applications research: 1988-1994”, Expert Systems, 12, 253-261.

Acknowledgements: I would like to thank Mohamed El-Farra for research assistance in the collection of the data used in this paper. In addition I would like to thank colleagues at the University of Sheffield, and in particular Sarah Brown, for commenting on an early version of the material presented here.

Figures and diagrams

Figure 1: A simple feed-forward neural network

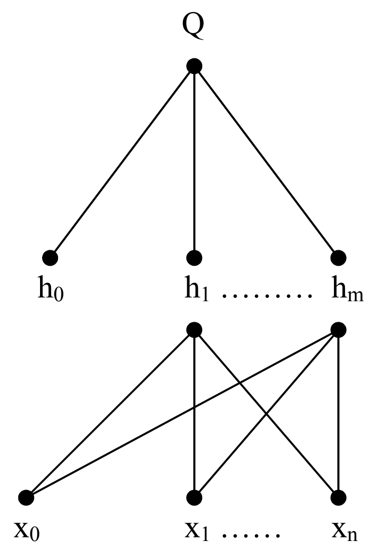

Figure 2: Interpreting the estimated neural networks

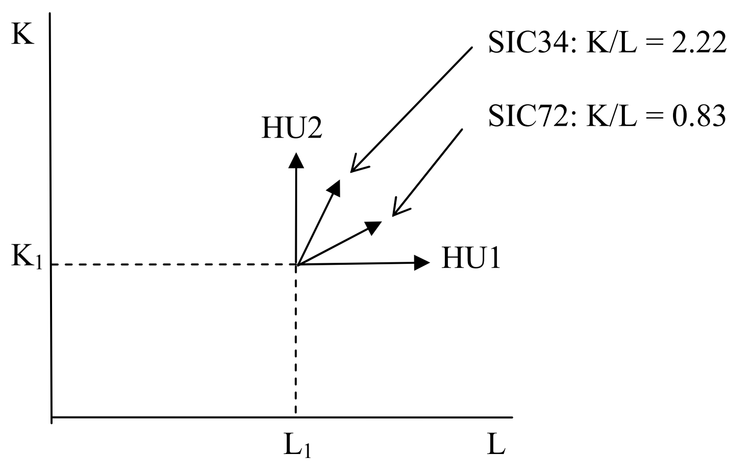

Figure 3a

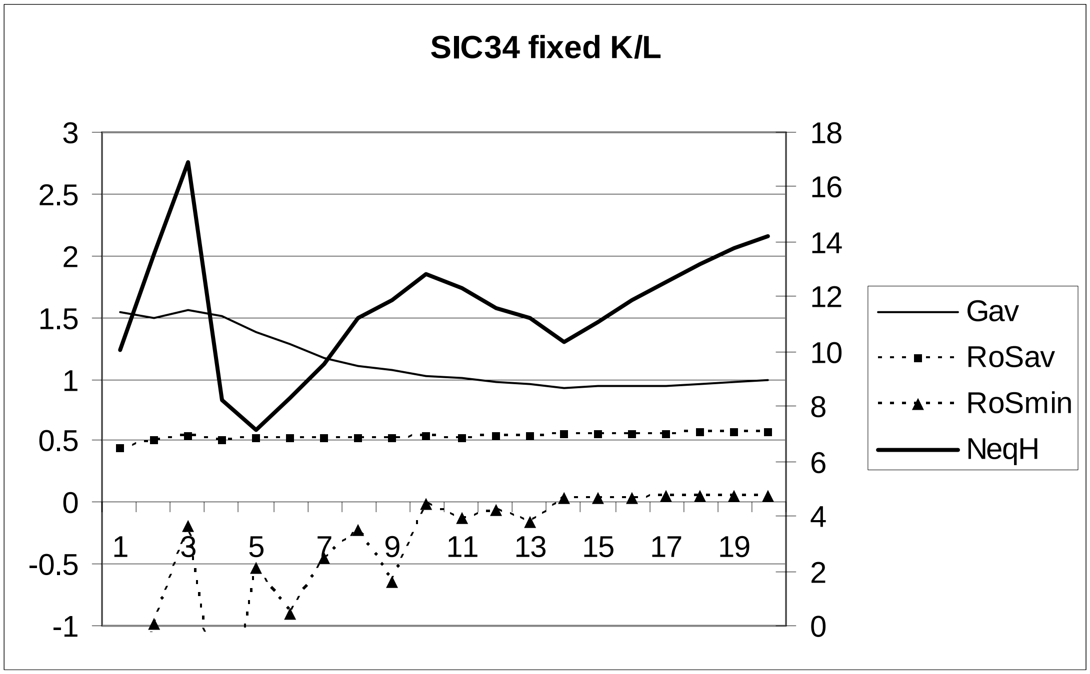

Figure 3b

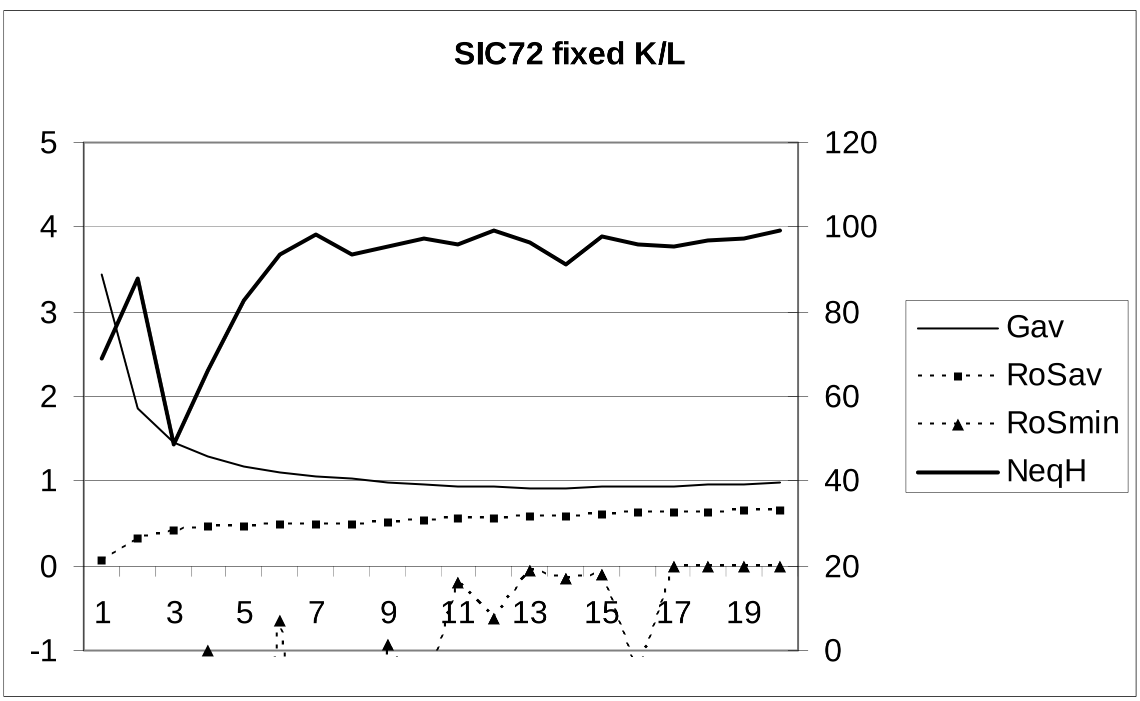

Figure 4a

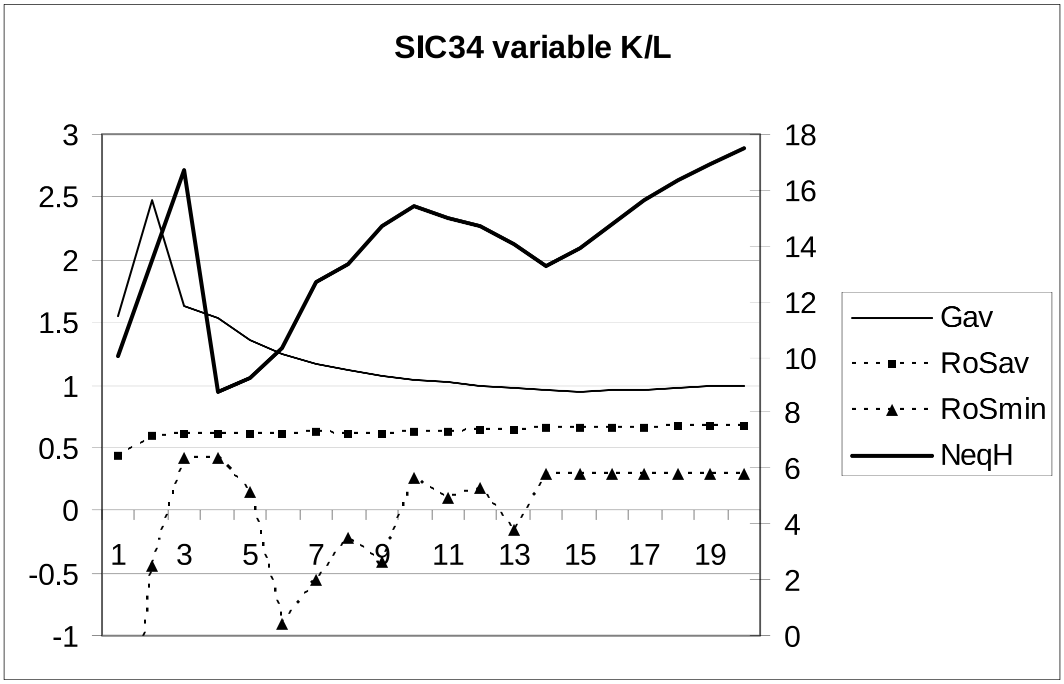

Figure 4b

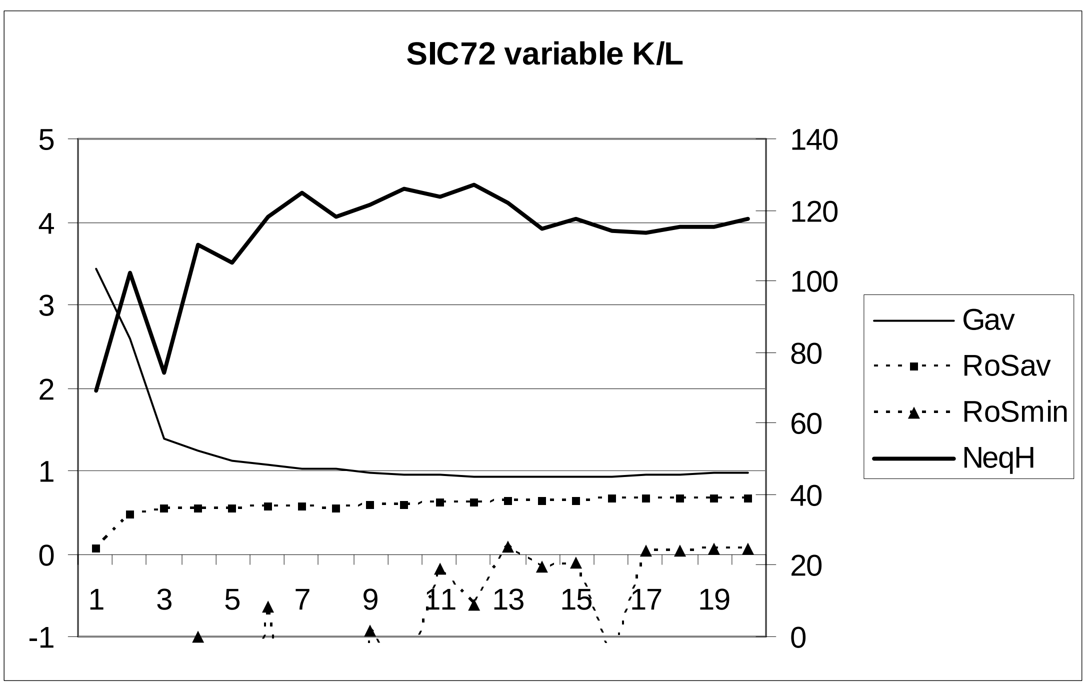

Figure 5a

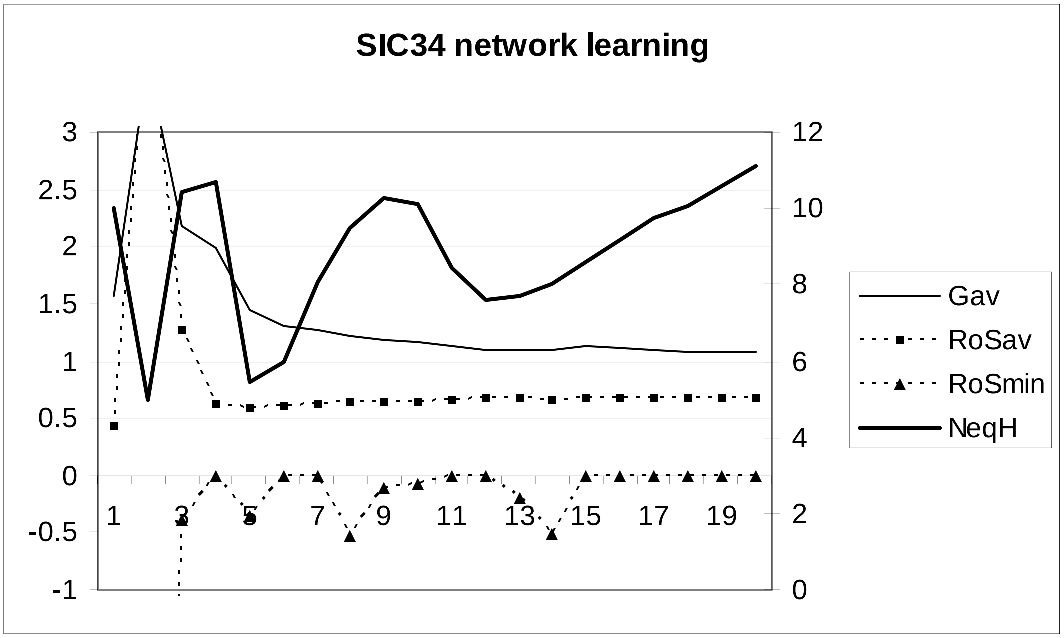

Figure 5b