Paper 245-27

Neural Networks and Genetic Algorithms as Tools for Forecasting Demand in Consumer Durables (Automobiles)

Paul D. McNelis, Georgetown University, Washington D.C.
Jerry Nickelsburg, Economic and Financial Consultant, Malibu, CA

I. INTRODUCTION

This paper attempts to make the tools of neural networks and genetic algorithms accessible, "user-friendly", and operational, for the broader population of economists, analysts, and financial professionals, who see to become more efficient in forecasting market conditions. The particular market chosen for the analysis is that of automobiles in the United States. Selecting a consumer durable allows for a demonstration of the power of temporal linkages in neural network forecasting.

Our empirical focus in this paper is the automobile market. This market is well developed and there is a wealth of research on the theoretical foundations and the empirical behavior of the market. Since the work of Chow(1960) demonstrating that it is one of the more stable consumer durable markets, empirical analysis has focused on improving the aggregate and disaggregated market forecasting with traditional time series and pooled time series cross sectional methodologies, McCarthy(1996). Moreover, the general stability in the underlying market structure and the recursive nature of producer versus consumer decision making, has made this market amenable to less complex estimation methods. As research suggests this is precisely the kind of market where linear time series forecasting will perform well, we take it as an appropriate test of the alternative of neural network modeling and forecasting.

The paper proceeds with a brief description of the two methodologies, a description of the model and the data, an analysis of the econometric and neural network out of sample forecast results, and a concluding section with comments on the use of SAS® and alternatives in the generation of neural network forecasting models.

II. FORECASTING MODELS

In forecasting one usually starts with the linear regression model, given by the following equation.

(1) $$Y_t = \Sigma \beta_\kappa X_{\kappa,t} + \varepsilon_t$$

(2) $$\varepsilon_t \sim N(0, \sigma^2)$$

Where the variable ε_t is a random disturbance term, normally distributed with mean zero and variance σ^2, and $\{\beta_\kappa\}$ represent the parameters to be estimated. The set of estimated parameters is denoted $\{\beta e\kappa\}$ while the set of forecasts of Y generated by the model with the coefficients $\{\beta e\kappa\}$ is denoted by {Yet}. The goal is to select $\{\beta e\kappa\}$ in order to minimize the sum of squared differences between the actual observations Y and the observations predicted by the linear model Ye. While there are a number of different computational methods for selecting $\{\beta e\kappa\}$ depending on the structure of the model, the basic concept is to use the implied stochastic process of the error term to generate a maximum likelihood solution to the problem. Commonly this results in the autoregressive linear forecasting model:

(3) $$Yt = \sum_{t=1}^{K^*} \beta_i Y_{t-i} + \sum_{j=1}^{K} \gamma_j X_{j,t} + \varepsilon_t$$

in which there are K independent X variables, with coefficient γ_j for each x_j, and K* lags for the dependent variable y with, of course, K+K* parameters $\{\beta\}$ and $\{\gamma\}$ to estimate. Thus the longer the lag structure, the larger the number of parameters to estimate, and the smaller are the degrees of freedom of the overall regression estimates.

The number of output variables may be more than one. In the case of orthogonal error terms ε, each may be estimated as a single equation as above. In the case of correlation between the error terms, a vector auto-regression approach is appropriate. The linear model has the useful property of having a closed form solution for solving the regression problem of minimizing the sum of squared differences between {Y} and $\{Y_e\}$. Thus the linear method is quick and easily understood.

For short run forecasting, the linear model is a reasonable starting point or benchmark. This is because in many markets one observes only small symmetric changes in the variable to be predicted about the long term trend. However it may not be especially accurate for volatile markets nor for markets which exist during times of structural or rapid change. The linear approximation to the market forces breaks down when the nonlinear behavior of the market as evidenced in the data cannot be modeled as an error process with a Gaussian probability structure. Asset prices, including consumer durable prices, are often observed to creep up to historically high levels and suddenly plummet away from the historical trend. This phenomenon, often studied as the economics of speculative bubbles, is a result of a mismatch between current expectations, and the expectations (held moments before) that support the higher price. Its applicability is wider than Tulip and dot.com bubbles and affects a wide variety of asset prices at one time or another. As this is the most important of the asset price movements to capture, one is led to forecasting with non-linear models.

From the Weierstrass Theorem, a polynomial expansion around a set of inputs x with a progressively larger power P is capable of approximating to a given degree of precision any unknown but continuous function y = g(x). Consider for example a second degree polynomial approximation of three variables, $[x_{1t}, x_{2t}, x_{3t}]$ and a continuous but unknown function g. The approximation formula for g would be :

$$(4) \quad y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x^2_{1t} + \beta_5 x^2_{2t} + \beta_6 x^2_{3t} + \beta_7 x_{1t} x_{2t} + \beta_8 x_{1t} x_{3t} + \beta_9 x_{2t} x_{3t}$$

Note that the second degree polynomial expansion with three arguments or dimensions has three cross terms $\{\beta_7, \beta_8, \beta_9\}$ which represent the cross effects of the exogenous variables on y. This simple expansion requires the estimation of ten parameters and the degrees of freedom in the estimate will decline exponentially as the degree of the polynomial in the expansion increases. This "curse of dimensionality" is an exacting price to pay for accuracy in the fitting of the non-linear model to the data in the econometric estimation.

As an alternative to these traditional methods, the neural network approach is a more parsimonious estimation technique. The reason why one uses the neural network is simple and straightforward. The goal is to find an approach or method which forecasts well data which are generated by often unknown and highly non0linear processes with as few parameters as possible.

Like the linear and polynomial approximation methods, a neural network relates a set of input variables $\{x_i\}$, i=1,...,k to a set of one or more output variables$\{y_j\}$, j=1,...,k*. The difference between a neural network and the other approximations methods is that the neural network makes use of one or more "hidden layers", in which the input variables are transformed by a logistic or log-sigmoid transformation. The appeal of the log-sigmoid transform function comes from its "threshold behavior" which characterizes many types of economic responses to changes in fundamental variables. For example, if interest rates are already very low or very high, small changes in this rate will have very little effect on the decision to purchase an automobile. However within critical ranges between these two extremes, small changes may signal significant upward or downward movements and therefore create a pronounced impact on automobile demand.

Furthermore, the shape of the logsigmoid function reflects a kind of learning behavior. Often used to characterize learning by doing, the function becomes increasingly steep until some inflection point. Thereafter the function becomes increasingly flat up and its slope moves exponentially to zero. Following the same example, as interest rates begin to increase from low levels, consumers will judge the probability of a sharp uptick or downtick in the interest rate based on the currently advertised financing packages. The more experience they have, up to some level, the more apt they are to interpret this signal as the time to take advantage of the current interest rate, or the time to postpone a purchase. The results are markedly different than those experienced at other points on the temporal history of interest rates.

The following system of equations are most commonly used in the "feed forward" neural network:

$$(5) \qquad n_{k,t} = \omega_{k,0} + \sum_{i=1}^{i*} \omega_{k,i} x_{i,t}$$

$$(6) \qquad N_{k,t} = 1 / (1 + e^{-ni,t})$$

$$(7) \qquad y_t = \gamma_0 + \sum_{k=1}^{k*} \gamma_k N_{k,t}$$

In this system there are i* input variables {x} and k* neurons. Equations (5) through (7) are interpreted as: at any time t the a convex combination of the input variables x (equation (5)) are transformed by the logsigmoidal transform (equation (6)) and are input into the output equation (7) to generate a forecast of the variables of interest y. It is easy to see that this is simply a nonlinear expansion of the function g for purposes of estimation. Researchers have used other transform functions such as the hyperbolic tangent function as well, however the logsigmoidal remains the most successful transform function to date.

III. THE AUTOMOBILE MARKET

The structure of the automobile (new vehicle) market is recursive. Manufacturers evaluate and forecast the demanded stock of automobiles, the number of retirements, and their market share. Adding a dose of strategic planning they decide on how much to produce. These decisions occur well before production and distribution take place. Manufacturers are providing a flow of capital goods which augment an existing stock. Consumers decide at the time of purchase and based on income, price and utility requirements, what stock is optimal. To the extent that consumer decisions to expand the stock of the asset coincide with or exceed the amount of production by manufacturers, price will adjust to revised the optimal stock and clear the market. To the extent they fall short, the number of retirements of automobiles will increase and the price of new vehicles will fall to clear the market. Chow(1960), Hess(1977) and McCarthy(1996) show how forecasting the demand in the markets is a sufficient proxy to modeling the optimal stock decision.

Of interest to manufacturers and distributors in this market is the prediction of future price and quantity variables, especially when there is rapid change in the market. When focusing on a particular consumer durable market, it is often reasonable to assume that the impact of the rest of the economy on the market is captured in expectations of future important variables such as interest rates and income. Therefore a model based solely on the equilibrium in the durable goods market suffices for forecasting purposes and is commonly applied in this market Dyckman (1965). The linear time series model which has been explicated above and forms linear regressions of the relevant variables on present exogenous variables and lagged exogenous and endogenous variables has been used as a benchmark for the analysis here. The same inputs were used to characterize the model as a nonlinear process and neural network techniques were employed.

The purchase of a durable good on the part of a consumer is a decision to purchase a stream of benefits over time. Therefore the decision depends on the consumers rate of time preference as well as price and permanent income. Forward looking expectations on the part of consumers dictate their demand decisions. Following the body of rational expectations theory of economics, such forward looking expectations depend on the body of information known to the consumer at the time of the decision. This can be modeled as a time series lag structure on known and unknown variables. The unknown variables form the error term of a partial equilibrium demand model.

We model consumer demand for automobiles as a function of present and expected interest rates, and permanent income following. The model is:

(8) $$\mathbf{Q_{Dt} = F(P_t, Y_{pt}, i_t, W_t)}$$

where $\mathbf{Y_{pt}}$ is permanent Income at time t, $\mathbf{I_t}$ is the interest rate at time t, $\mathbf{P_t}$ is the price, $\mathbf{Q_{dt}}$ is the quantity demanded and $\mathbf{W_t}$ is wealth.

We model producer supply of automobiles as a function of the expected price in the market when the automobiles arrive at the dealers and the expected level of aggregate production in the economy:

(9) $$\mathbf{Q_{St} = G(E_{t-1}\{P_t, Y_{A,t}\})}$$

Where E_T represents the mathematical expectation of the vector conditional on all of the information available at time T.

Combining equations 1 and 2 yield the reduced form of the model as:

(10) $$\mathbf{Q = R(E_{t-1}\{P_t, Y_{At}\}, YP_t, i_t, W_t)}$$

(11) $$\mathbf{P = S(E_{t-1}\{P_t, Y_{At}\}, YP_t, i_t, W_t)}$$

Substituting a linear distributed lag structure for the expectations and for the calculation of permanent income yields the typical vector autoregressive model used in time series forecasting. Identification requires restrictions on some parameters as developed by the theory for the particular good. In the case of automobiles, the production decision is made considerably in advance of the purchase decision and hence a recursive structure identifies the model.

Dyckman (1966) pointed out that heterogeneity in the product would produce biased estimates using the linear time series model. His solution was to first difference the data arguing that changes in the heterogeneous structure would occur slowly and would be washed out by differencing. Trandel (1991) analyzed this bias in a cross sectional study and found Dyckman's conclusion to be valid. Bordley and McDonald (1993) also looked at heterogeneity in a cross sectional study and found quality to have a significant impact on income elasticity estimates.

The implications of these studies, as well as Dyckman's finding of significant non-linear effects of credit conditions on demand suggest that the stability of the linear relationship may mask important movements in the variables due to non-linear decision making. In particular, innovations in new vehicles such as small pick-up trucks and SUV's create a demand for new vehicles without dramatically altering the optimal stock. Thus, the company that takes a risk in defining a new product and changing the heterogeneity of the market offering, is apt to find demand for its product exceeding supply, a common occurrence in the automobile industry. While we don't suggest that models can forecast strategic moves in product development, situations such as these violate the basic assumptions of the linear models. Substituting a neural network structure permits a more complex relationship of the variables and has the potential for more accurate forecasting and for accounting for rapid changes in demand due to innovations or other structural changes. The estimation in the next section will test that hypothesis.

IV. ESTIMATION AND TESTING OF FORECASTING MODELS OF AUTOMOBILE MARKETS

In this section the model from Section III is estimated in two different ways and the two methods are compared and tested for their forecasting prowess. Data were collected on the aggregate production of new vehicles excluding heavy trucks and machinery (Bureau of Economic

Analysis), automobile prices as represented by a price index (Bureau of Labor Statistics), Home Mortgage Rates (Federal Reserve Board of Governors) and Personal Disposable Income (Bureau of Economic Analysis) and are publicly available from the stated sources or the authors. Home Mortgage Rates were chosen as the relevant interest rate following the results of Hess(1977) which show that the consumers consider housing and automobile decisions jointly. Personal Disposable Income was generated from Consumption and Savings Rate Data. The Consumption Data was the average over the quarter to better reflect the permanent income concept. All data were converted to log first differences to negate the effect of strong secular trends.

For the linear expansion models the estimation method was a time series seemingly unrelated regressions method utilizing the standard routines of SAS/STAT®, SAS/ETS®. For the neural network models alternative software solutions were employed. The choice of software for the estimation was a function of the availability at the time of estimation.

The models were parameterized in two ways. First, the optimal choice of lag structure was determined by that which had the best out of sample performance in the neural network framework. Second, the optimal lag structure was specified by that which had the best out of sample performance using the linear autoregression model. In both cases the parameterized models were tested against each other to determine the out of sample forecasting accuracy using standard test statistics.

Table 1 gives the results for the models using the lag structure found to be optimal by the neural network parameterization. The first measure of fit of the models is the in sample analysis. The R-Square and the H-Q Statistic give us two measures of this. The H-Q Statistic penalizes a model for loss of degrees of freedom in estimation and tends to correct for the impact of over parameterization on the goodness of fit criteria. The price equation shows a .8 R-Square compared to a .18 for the linear model and the quantity equation yields a .72 R-Square compared to a .42 for the linear model. This relationship is maintained in the H-Q Statistic and not surprisingly, the Neural Model shows a better in sample fit.

The out of sample criteria used here are the Root Mean Square Error and the Success Ratio as descriptive statistics and the Diebold-Mariano Statistic for tests of significant differences between the models. For both the price and quantity equations the RMSE is lower and the Success Ratio for the sign of the forecast is higher for the Neural Network Model. The Diebold-Mariano Statistic yields .23 and .05 respectively on the hypothesis that the two models yield statistically equivalent forecasts. In both cases Linear Model does not yield average absolute forecast errors significantly below those of the Neural Network Model.

As an alternative approach, the models were estimated with the lag structure that would obtain with the optimal Linear Time Series Model. These results are presented in

Table 1
New Automobile Price and Quantity Estimate and Forecasting Results
Lag Structure Parameterized by Best Neural Network
Out-of-sample: last 10% of available data
Monthly Data 1991-2000

	Price Equation		Quantity Equation	
	Linear	**Neural Net**	**Linear**	**Neural Net**
Lag structures for p, q, i, Y		3 - 3 - 5 -4		2 - 4 - 4 - 2
Neurons		5		3
R-Square	0.180342	0.816423	0.415916	0.750713
Hannan-Quinn Statistic	-569.359	-628.335	-39.0256	-88.636
Q-stat: marginal significance	0.668029	0.572111	0.569535	0.346452
Q-stat (squared error): marginal sig.	0.099631	0.438805	0.021948	0.007917
Engle-Ng symmetry: marginal sig.	0.920335	0.60789	0.013047	0.361376
J-B normality: marginal sig.	0.145069	0	0.036102	3.52E-11
Root Mean Squared Error	0.017179	0.010158	0.072469	0.065297
Diebold-Mariano: marginal sig.	0.232233		0.052245	
Success Ratio on signs of forecasts	0.5	0.666667	0.666667	0.916667

Table 2
New Automobile Price and Quantity Estimate and Forecasting Results
Lag Structure Parameterized by Best Linear Model
Out-of-sample: last 10% of available data
Monthly Data 1991-2000

	Price Equation		Quantity Equation	
	Linear	Neural Net	Linear	Neural Net
Lag structure for p, q, i, Y	5 - 6- 3- 2		3 - 1 - 5 - 3	
neurons		2		5
R-Square	0.147862	0.655957	0.318	0.80305
Hannan-Quinn Statistic	-558.333	-628.016	-21.3721	-70.4972
Q-stat: marginal significance	0.578823	0.731858	0.022394	0.932838
Q-stat (squared error): marginal sig.	0.027216	0.154352	0.053655	0.99696
Engle-Ng symmetry: marginal sig.	0.835742	0.153479	0.039422	0.848198
J-B normality: marginal sig.	0.119761	3.4E-06	0.039238	0
Root Mean Squared Error	0.016684	0.2217	0.064669	0.114494
Diebold-Mariano: marginal sig.	0.875269		0.988257	0
Success Ratio on signs of forecasts	0.583333	0.5	0.916667	0.5

Table 2. Surprisingly, the in sample goodness of fit criteria are clearly superior for the Neural Network Model. We expected that all of the criteria would favor the Linear Model since this was used in the optimization criteria. Clearly the step of finding the optimal number of neurons (logsigmoidal transforms) dramatically improves the model in sample fit. The out of sample forecast measures yield different results than before. The RMSE and Success Ratio on signs of forecasts both favor the linear model. The Diebold-Mariano Statistic is still not significantly different from zero.

In comparing the two sets of results the following conclusions emerge. The Neural Network Models proved to be the better fitting models for the data. The out of sample forecasting of the models estimated with the “best linear configuration” tended to be better for the Linear Models, but the performance of these was inferior to the best Neural Network Models. The statistical tests of the out of sample forecasts were inconclusive, perhaps due to the stability of the underlying processes during the out of sample time frame.

V. SUMMARY AND COMMENTS

This paper described the use of Neural Network Models in application to industrial forecasting. Much of the early development of neural network analysis has been within the disciplines of Psychology, Neuro-sciences, and Engineering, often related to problems of "pattern recognition". Genetic algorithms have followed a pattern of development within applied mathematics: optimization of dynamic non-linear and/or discrete systems. In applied fields there is an "interface" problem for professionals whose training is in classical Statistical Methods. The analysis here has shown that even in the least favorable case, Neural Network Model forecasting can be a useful tool for business analysis.

The appeal of Neural Network Modeling is in its efficacy in picking up structural or nonlinear behavior in the underlying data. As a result, one can expect an increasing use of these techniques in the future. The software employed in our analysis was a mixture of data mining tools available with SAS/STAT®, SAS/ETS® and other software packages. In particular, the complex calculations required to parameterize a vector Neural Network Model were performed with other packages. The primary driving force behind the choice of software was the authors' familiarity with the packages. Most of the applied developments in this area are using other packages. For the SAS user and the Neural Network modeler the challenge is to overcome the hurdles to using one umbrella software package. As Neural Network Modeling becomes more prevalent, this challenge becomes more important.

VI. REFERENCES

Bordley, Robert F. and James B. McDonald, "Estimating Aggregate Automobile Income Elasticities From The Population Income-Share Elasticity" _Journal of Business and Economic Statistics_, 11 (1993), 209-214.

Chow, Gregory, "Statistical Demand Functions For Automobiles and Their Use For Forecasting" in Arnold C. Harberger (ed.) _The Demand For Durable Goods_, University of Chicago Press, (Chicago 1960), 149-178.

Dyckman, Thomas R. "An Aggregate Demand Model For Automobiles" _Journal of Business_, 38 (1965), 252-265.

Hess, Alan C. "A Comparison of Automobile Demand Functions" _Econometrica_, 45 (1977), 683-701.

Lubet, Lloyd. "Combinatorics with Dynamic Regression" www.sas.com/service/library/periodicals/obs/obswww20/index.html.

McCarthy, Patrick S. "Market Price and Income Elasticities of New Vehicles" _The Review of Economics and Statistics_, 78 (1996), 543-548.

McNelis, Paul _Neural Networks and Genetic Algorithms: Frontiers of Forecasting, Risk Analysis and Decision Making In Volatile Financial Markets_, Georgetown University (2001).

SAS® ETS, Version 8. Cary, NC: SAS Instiute.

Trandel, Gregory A. "The Bias Due To Omitting Quality When Estimating Automobile Demand" _The Review of Economics and Statistics_, 73 (1991), 522-525.

VII. CONTACT INFORMATION

Paul D. McNelis
Gasson Professor of Macroeconomics and International Finance
Boston College
Department of Economics, Georgetown University
Washington D.C. 20057-1036
McNelisP@Georgetown.edu

Jerry Nickelsburg
Economic and Financial Consultant
5746 Busch Drive
Malibu, CA 90265
JQNickels@aol.com