PREDICTING RECESSION USING THE YIELD CURVE: AN ARTIFICIAL INTELLIGENCE AND ECONOMETRIC COMPARISON

Mohamad Shaaf
University of Central Oklahoma

INTRODUCTION

Over the years, economists have provided evidence that both stock and bond market data contain information relevant for predicting future economic growth. Several authors investigated this subject empirically and found that the bond market predictions are more accurate. For example, Harvey [1989] showed that the bond market explained more than 30 percent of the variation in economic growth over 1953-89, while stock market variables explained only about 5 percent. This is due to variation in stock prices that reflect both changes in expected economic growth and changes in the expected risk of stock cash flows.

Other studies focused on the bond market or the yield curve as a predictor of recession. The yield curve most often used is the one representing the rates of return on Treasury bonds against their maturity dates. Normally, the yield curve has a positive slope and looks convex as is shown in Figure 1. However, when the curve gets flat or slopes downward, it is an indicator of economic recession in the future.

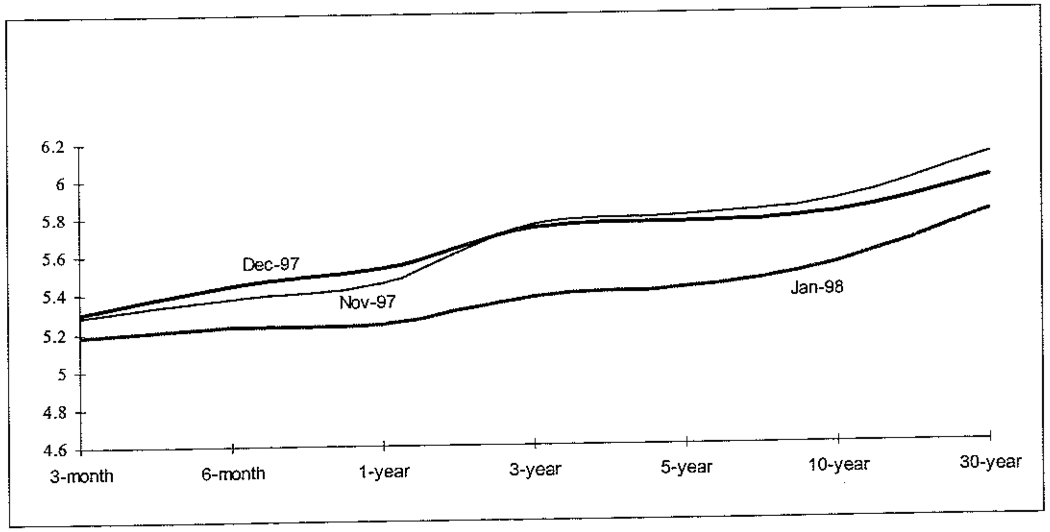
The theoretical basis of the yield curve goes back to Irving Fisher [1907]. There are several forward-looking hypotheses of the yield curve. One hypothesis suggests that tight monetary policy causes short-term interest rates to rise relative to the long-term rates. This action could reduce the spread between the long-term and the short-term interest rates, making the curve flat or downward sloping. Another hypothesis argues that tight monetary policy results in lower inflationary expectations and consequently reduces long-term interest rates, also reducing the spread and causing the curve to flatten.[1]

The empirical works of Harvey [1988; 1989; 1991; 1993], Mishkin [January 1990, August 1990; 1991], Estrella and Harduvelis, [1991], Hu [1993], Campbell [1995], Haubrich and Domrosky [1996], Dueker [1997], and others confirm the predictive power of the yield curve. These investigators used different types of econometric models to show the relationship between the yield curve spread, with or without other variables, and real-GDP growth.[2]

Mohamad Shaaf: Economics Department, University of Central Oklahoma, Edmond, OK 73034. E-mail: mshaaf@ucok.edu

Eastern Economic Journal, Vol. 26, No. 2, Spring 2000

FIGURE 1
Treasury Yield Curve



Source: Federal Reserve of St. Louis, *Monetary Trends*, March 1998

Estrella and Mishkin used a probit model to estimate the probability of recession given the shape of the yield curve. Using pseudo-R^2 as a measure of fit and '*t*' statistics in their model they concluded that the four-quarter lags of the Treasury spread:

> shows the best predictive performance across the range of horizons examined . . it is substantially outperformed by a number of other indicators, including stock price indices, the Commerce and Stock-Watson leading indicators, and some of the Commerce indicator's components [1995, 22].

In recent years the artificial intelligence technique and its forecast have been used successfully in various economic studies, including investment and financial forecast [Trippi and Turban, 1993; Hsieh, 1993; Swales and Yoon, 1992], effectiveness of fiscal and monetary policy [Shaaf, 1996], housing markets [Shaaf and Erfani, 1996], exports growth as the source of economic growth [Shaaf and Ahmadi, 1999], and others. The purpose of this study is twofold: The first, using an artificial intelligence model called "neural networks," is to investigate the power of the yield curve in predicting economic activity and recession. The second is to compare the prediction of the artificial intelligence model with those of the econometric approach. This study uses the same model, variables, and lag structures as those determined to be the "best-fit" by Estrella and Mishkin [1995].

The results of the artificial intelligence model confirm those of the econometric findings; a flat or negative yield curve is the sign of recession in the near future. The findings of out-of-sample simulations also suggest that the artificial intelligence approach to predicting recession is more accurate than that of the traditional econometric method. This paper proceeds as follows. The first section describes the architecture of artificial intelligence. The next section presents the empirical results of the artificial neural network and regression models, and those of the in-sample and out-of-sample simulations. The last section concludes the paper by summarizing the neural networks and the findings.

ARTIFICIAL INTELLIGENCE MODEL

One of the branches of the study of artificial intelligence is the neural network. Neural networks are multiple-layer configurations similar to the structure of the brain, consisting of simple processing elements or nodes that interact with each other through weighted connections in the system. This study utilizes the most widely used neural network model termed "back propagation," which is a non-parametric algorithm for adjusting connection weights in a multiple-layer network. Back propagation is a learning design by which the multi-layer network is set for pattern recognition utilizing actual cross section or time series data as the external teacher. A typical neural network consists of:

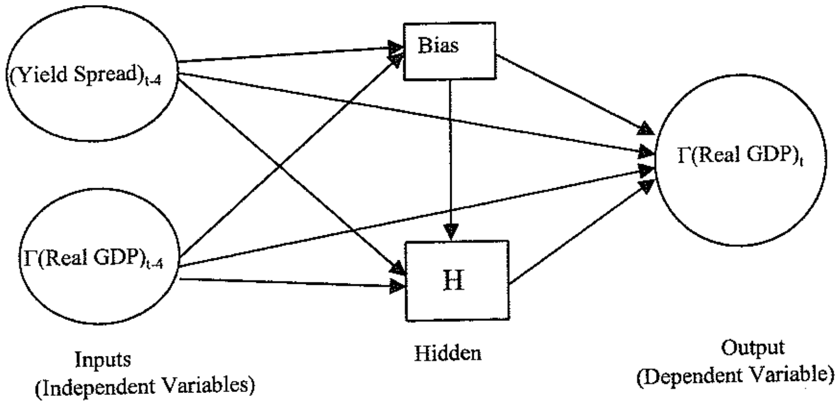
1. one input (independent-variable) layer
2. one or more intermediate or hidden layers
3. one output (dependent-variable) layer.

Similar to the simultaneous equations in econometric modeling, neural network architecture can have several independent variables and several dependent variables. A network similar to the one used in this study with two independent variables (inputs), a hidden layer, a bias, and one dependent variable (output) is shown in Figure 2. Each connection from the input layer to the hidden layer or from the hidden layer to the output layer has a weight. These weights represent the coefficients or the parameters of the model. The size of each weight represents the relative strength of the connection. Each node computes a weighted sum of the incoming values and passes this sum through a nonlinear or linear function as output. The output of one layer serves as the input to the next.

The hidden layer(s) is a processor of information and a bridge between the independent variable(s) and dependent variable(s). That layer(s) computes the information and does most of the work in the network. Since this layer is connected to independent variables (inputs) on one hand and dependent variables (output) on the other, and is hidden from the outside world, it is called the hidden layer. Besides its link to the hidden layer, the input layer can be directly connected to the output layer. Notice that the inputs of the network used in this study are also connected directly to the output layer (see Figure 2).

The network is "fully connected." That is, all the nodes are linked in adjacent layers. These links are the weights that can be strong or weak depending on their size. The weights are adjusted for the minimization of the mean square error as the

FIGURE 2
An Artificial Neural Network for the Yield Curve as a Predictor of Recession



objective function. Similar to the econometric method, the back propagation design is an optimization technique for finding optimum values of the weights as parameters. Furthermore, there is a bias-processing element that is similar to the intercept notion in equations of econometric models.

The three main phases in the operation of a network are learning, recall, and testing. In the learning phase, the neural network recognizes a pattern between independent variable(s) and dependent variable(s) and estimates the final value at the output layer. Subsequently, this estimated value is compared with the actual output, and errors are calculated. These errors become the factor used to adjust the weights and, subsequently, to reduce the errors and readjust the weights of the connections.

This process of weight adjustment continues further until the error declines to an acceptable level, when possible. Through this process, the neural network learns the rules and adapting patterns for processing the information. The resulting pattern of rules and adapting patterns linking the independent variables to the dependent variables in the learning process can be used for testing the accuracy of the networks. In addition, the pattern can be used for prediction of the dependent variable(s), given the independent variable(s).

To explain the mathematical equations of the neural networks, assume there are I number of independent variables, j hidden nodes in the hidden layer, and k output nodes in the output layer. Accordingly, back propagation computes the summation of multiplication of independent variables, X_I, by their corresponding weights (from the input layer to the hidden layer), and adds bias weights (intercept) to it as follows:

(1) $$u_j = w_{0j} + \sum_{i}^{I} w_{ij} * X_i$$

where w_{0j} is the bias weight j, and w_{ij} are the weights from the input i to hidden node j.

The weighted summations, u_j, in equation (1) are then transformed by a transfer function in the hidden layer. It is called the "transfer function" because it acts to "transfer" the internally generated sum to a potential output value. Three common nonlinear transfer functions used in artificial neural networks are the sigmoid, hyperbolic tangent, and sine functions. The sigmoid transfer function, which is considered appropriate and is used in this study, is a continuous monotonic mapping of the input into a limited-range value between zero and one. The sigmoid function of u variable is a smooth version of {0,1} step-shape function, and is defined as:

(2) $$y_j = f(u) = (1 + e^{-u})^{-1}$$

where e is the base of natural logarithms. Thus, the output of the hidden node J is a sigmoid-shaped hypersurface of I dimensions in a space of $(I+1)$ dimensions. This output, y_j, from the hidden layer enters the output node as input, and subsequently computes the output of the network with k output nodes, z_k, as:

(3) $$z_k = g(v_k), \quad k = 1, 2, ...K$$

and $g(v_k)$ is the sigmoid function of v_k as:

(4) $$v_k = w_{0k} + \sum_{j=1}^{J} w_{jk} y_j$$

where w_{jk} are weights on the links from hidden node j to the output k, and w_{ok} is the bias weight of output node k. Notice that at the output node the process works the same way as in the hidden nodes, and the output v in equation (4) enters the sigmoid function of equation (3) and transforms the information in the form of:

(5) $$z = f(u) = (1 + e^{-v})^{-1}$$

If the neural network is directly connected from the input nodes to the output nodes, the v_k function in equation (4) is modified as follows:

(6) $$v_k = w_{0k} + \sum_{j=1}^{J} w_{jk} y_j + \sum_{j=J+1}^{J+I} w_{jk} X_{j-J}$$

The third term in equation (6) represents the weighted summation of input nodes by the weights directly connecting them to the output nodes.

The next step in the neural network is the calculation of the error. Suppose the estimated final dependent variable at the output layer, which is produced by the network with N observations, is z_{nk}, and the actual dependent variable is A_{nk}, then the mean square error is computed as:

(7) $$E = \sum_{n=1}^{N} \sum_{k=1}^{K} (Z_{kn} - A_{kn})^2 / 2NK.$$

Here, E is the global error function which is differentiable of all the connection weights in the network.

The subsequent stage in the artificial neural network is to adjust the weights (parameters) of the model. The strength of the connections (weights) changes continuously for the minimization of the error as the objective function. The weights in the neural networks fall within two major categories The first category connects the hidden nodes to the output nodes. The second connects the input nodes to the hidden nodes. The equations for these two types of weight adjustment are explained next.

Hidden-to-Output-Node Weights

When the error E is determined, the network back propagates it for the increase or decrease of the weights. First, it selects some arbitrary point in weight space and computes the slope of the error surface at that point. Then, it computes the change or derivative of the error, E, with respect to the weight from the hidden node J to the output k, $dE/d\,w_{Jk}$, through a chain-rule multiplication of three terms as follows:

(8) $$\delta E/\delta w_{jk} = (\delta E/\delta z_k)(\delta z_k/\delta v_j)\,(\delta v_j/\delta w_{jk}).$$

The first term on the right side of the equation, $\delta E/\delta z_k$), is the derivative of the error with respect to the dependent variable k. The second term in the chain is the derivative of that node's output with respect to its weighted sum of inputs from the hidden layer to the output layer. The last term, $\delta v_j/\delta w_{jk}$), is the derivative of that sum with respect to the weight connected to the output node k. According to equation (4), the last term in equation (8) depends on the output in the hidden node, y_j, as:

(9) $$(\delta v_j/\delta w_{jk}) = y_j.$$

Input-to-Hidden-Node Weights

Similarly, the neural networks compute the relative change of the error, E, with respect to a change in the weights (from input nodes to the hidden nodes), $\delta E/\delta w_{jk}$, as a chain-rule multiplication of five terms as follows:

(10) $$\delta E/\delta w_{ij} = [\sum_{k=1}^{K} (\delta E/\delta z_k)(\delta z_k/\delta v_k)(\delta v_k/\delta y_j)](\delta y_j/\delta u_j)(\delta u_j/\delta w_{ij}).$$

The first two terms on the right-side of equation (10) are similar to those in equation (8). The third, $\delta v_k/\delta u_j$ term is the derivative of the input from the hidden layer to the output layer with respect to the output j to the hidden node. The fourth term, $\delta y_j/\delta u_j$, is the derivative of the output in the hidden node with respect to the weighted sum of inputs in that node. The fifth term is the derivative of that sum with respect to the weight connecting inputs to the hidden node j.

The next task of the neural network is to find the direction of the weight adjustment in which the error surface declines the most. This approach is called the "*gradient descent rule.*" Given a point in the weight space, including those connecting the input nodes to the hidden nodes and those connecting the hidden nodes to the output nodes, the network produces a certain mean square error. The network must determine the direction and the magnitude of the change in weight which results in the greatest reduction of error. The direction and magnitude of the weight change is computed according to the following equations:

(11) $$w_p = w_{p-1} - lc \sum_{n=1}^{N} (\delta E/\delta w_p)$$

or

(12) $$\Delta w_p = -lc \sum_{n=1}^{N} (\delta E/\delta w_p)$$

Where w_p, is the new weight at process number p, Dw is the weight change and lc is the learning coefficient as a parameter which can be set by the user. It actualizes the proportion and the speed of the weight change. Its best value is determined by trial and error.

The last term on the right side of equations (11) and (12) is the total derivative of error with respect to each weight in the weight surface. The negative signs in those equations change the direction of the error so that it declines. With the new weight values, the process continues until the magnitude of the error approaches the minimum point in the error surface.[3]

The "explain" command in the neural network reveals the information about the extent of contributions of each independent variable on the dependent variable. In the literature of the artificial neural network, this measure is called "dithering," which is similar to the measure of elasticity in the econometric approach.[4]

EMPIRICAL FINDINGS

Quarterly data from the first quarter of 1959 to the first quarter of 1997 for real-GDP growth and the spread between one-year and three-month Treasury bill rates were used for two input mixes.[5] The contemporaneous quarter of the real-GDP growth as the dependent variable (output) and four-quarter lag of each variable in the input mix were utilized (see Figure 1). One input mix has only one independent variable (four-quarter lag of yield spread), and the other has two independent variables (four-quarter lag of the yield spread and four-quarter lag of the real-GDP growth).

Table 1 presents the learning results of the neural networks for each of the two data mixtures. The input value column represents results of the last output value of the node where the connection originates. Notice that the value for the bias is one, similar to the value of the intercept in the econometric method. The final estimated weights (parameters) of the connection between the input node and the output are also shown in the table.

The resulting signs of the weight (parameter) of the lag of the Treasury spreads, as expected, are positive for both mixes of data. This suggests that the magnitude of

TABLE 1
Results of the Artificial Intelligence Approach to the Yield Curve as a Predictor of Recession [a]

Input (Independent Variables)	Input	Weight Value	Elasticity AVE	(5% Dithering) SD
Without GDP_{t-4}				
Bias	+1.0000	-0.0570		
$(\text{Yield Spread})_{t-4}$	+0.4608	+0.3751	9.33	0.053
Hidden	+0.5021	-0.0425		
With GDP_{t-4}				
Bias	+1.0000	-0.1148		
$(\text{Yield Spread})_{t-4}$	+0.6501	+0.4024	10.01	0.06
$(GDP)_{t-4}$	+0.6968	+0.0719		
Hidden	+0.4764	-0.0485		

a. AVE = Average; SD = Standard Deviation; Number of runs = 20,000.

the yield spread between long- and short-term interest rates moves in the same direction as the growth of real-GDP. That is, a lower yield spread implies a slowdown in economic activity in the next four quarters. Similar conclusions were derived from the results of the traditional regression approach for each of the input mixes.[6]

The neural networks also calculated the relative sensitivity of real-GDP growth (elasticity) for each observation and each data mix. It represents the percentage changes of real-GDP growth as a result of a 5 percent change in the Treasury spread. Table 1 presents the averages of this elasticity measure for all the observations. Accordingly, a 5 percent increase (decrease) in the yield spread resulted in 9.33 percent increase (decrease) in the growth of real GDP, almost double. When the lag of the real-GDP growth is added to the input mix, the elasticity slightly increased.

In addition, Table 1 presents the average standard deviation (SD) of the elasticity of real-GDP growth to the changes in those input variables for each observation. The low SD of each of the input mixes implies that the elasticity measure did not vary across observations.

In conclusion, the results of the artificial neural network approach imply that a low or negative yield spread indicates a probable recession within a year.

SIMULATION AND COMPARISON

To test the accuracy and validity of the neural-network prediction, two types of simulations for each of the input mixes were performed: in-sample simulation, and out-of-sample simulation.

In-Sample Simulation

For the in-sample simulation, the whole data from the first quarter of 1959 to the first quarter of 1997 were used to test the predictive power of the yield curve. The

TABLE 2
Actual versus in-Sample Forecast of Real-GDP Growth with Lag of Yield Spread and Real-GDP Growth as Input: Artificial Intelligence and Regression Approach [a]

		Forecast						Forecast			
		Art Intel		Regression				Art Intel		Regression	
Year	Actual	WOT	WIT	WOT	WIT	Year	Actual	WOT	WIT	WOT	WIT
1960.1	3.26	2.48	2.52	2.59	0.67	1978.2	0.18	3.30	3.58	3.67	-3.49
1960.2	5.9	2.92	2.76	2.82	3.08	1978.3	1.76	2.86	2.80	2.87	-1.11
1960.3	4.95	3.05	2.92	2.97	1.98	1978.4	1.30	2.23	2.16	2.24	-0.94
1960.4	7.8	3.13	2.91	2.96	4.84	1979.1	1.64	1.88	1.67	1.73	-0.09
1961.1	5.74	3.03	2.99	3.06	2.68	1979.2	-9.49	1.88	1.66	1.72	-11.21
1961.2	3.95	3.07	3.11	3.18	0.77	1979.3	0.03	1.67	1.49	1.55	-1.52
1961.3	3.03	3.19	3.22	3.28	-0.25	1979.4	7.48	1.12	0.88	0.95	6.53
1961.4	1.35	3.08	3.18	3.25	-1.90	1980.1	4.84	1.14	0.91	0.98	3.86
1962.1	4.02	2.94	2.96	3.03	0.99	1980.2	-2.19	2.64	2.18	2.22	-4.41
1962.2	5.9	2.84	2.80	2.87	3.03	1980.3	2.66	3.27	3.15	3.21	-0.55
1962.3	6.76	2.84	2.77	2.84	3.92	1980.4	-5.03	1.25	1.21	1.29	-6.32
1962.4	2.74	2.80	2.68	2.74	0.00	1981.1	-5.52	1.09	0.96	1.03	-6.55
1963.1	9.81	2.72	2.68	2.75	7.06	1981.2	2.18	1.27	0.94	1.00	1.18
1963.2	3.9	2.75	2.77	2.84	1.06	1981.3	-0.91	1.92	1.78	1.85	-2.76
1963.3	4.34	2.56	2.59	2.66	1.68	1981.4	1.27	3.64	3.39	3.44	-2.17
1963.4	1.56	2.48	2.38	2.45	-0.89	1982.1	2.12	3.06	2.75	2.80	-0.68
1964.1	7.61	2.50	2.62	2.70	4.91	1982.2	8.76	3.08	3.01	3.07	5.69
1964.2	5.78	2.55	2.49	2.56	3.22	1982.3	6.60	4.61	4.56	4.62	1.98
1964.3	6.58	2.52	2.48	2.55	4.03	1982.4	6.84	3.92	3.88	3.94	2.90
1964.4	9.26	2.39	2.25	2.31	6.95	1983.1	7.68	3.71	3.69	3.75	3.93
1965.1	8.27	2.27	2.30	2.38	5.89	1983.2	4.74	3.50	3.66	3.74	1.00
1965.2	1.59	2.29	2.27	2.34	-0.75	1983.3	2.50	3.74	3.85	3.92	-1.42
1965.3	3.46	2.32	2.32	2.40	1.06	1983.4	2.86	4.01	4.14	4.22	-1.36
1965.4	2.6	2.26	2.35	2.43	0.17	1984.1	3.39	3.92	4.08	4.16	-0.77
1966.1	3.33	2.17	2.21	2.29	1.04	1984.2	2.54	4.35	4.44	4.52	-1.98
1966.2	1.75	2.19	2.03	2.10	-0.35	1984.3	6.04	3.78	3.77	3.83	2.21
1966.3	3.2	2.13	2.02	2.09	1.11	1984.4	2.11	4.04	4.06	4.13	-2.02
1966.4	3.13	1.92	1.77	1.84	1.29	1985.1	4.88	4.35	4.40	4.48	0.40
1967.1	5.76	2.11	2.00	2.07	3.69	1985.2	0.33	4.31	4.34	4.41	-4.08
1967.2	6.39	2.84	2.74	2.80	3.59	1985.3	2.23	4.23	4.36	4.44	-2.21
1967.3	2.67	2.70	2.63	2.70	-0.03	1985.4	2.32	3.80	3.79	3.85	-1.53
1967.4	2.19	2.66	2.59	2.65	-0.46	1986.1	2.08	3.18	3.20	3.27	-1.19
1968.1	4.99	2.44	2.43	2.50	2.49	1986.2	3.86	3.04	2.92	2.97	0.89
1968.2	0.68	2.21	2.20	2.28	-1.60	1986.3	3.27	3.27	3.21	3.27	0.00
1968.3	2.54	2.23	2.12	2.18	0.36	1986.4	5.77	3.35	3.30	3.36	2.41
1968.4	-1.2	2.18	2.04	2.11	-3.34	1987.1	2.35	3.17	3.11	3.17	-0.82
1969.1	-0.2	2.12	2.06	2.14	-2.34	1987.2	3.79	3.86	3.90	3.97	-0.18
1969.2	-0.5	2.16	1.98	2.04	-2.54	1987.3	2.40	3.97	4.00	4.07	-1.67
1969.3	4.48	1.95	1.81	1.88	2.60	1987.4	4.82	4.25	4.37	4.45	0.37
1969.4	-2.6	2.02	1.77	1.83	-4.42	1988.1	3.60	3.88	3.87	3.94	-0.34

TABLE 2 (Cont.)
Actual versus in-Sample Forecast of Real-GDP Growth with Lag of Yield Spread and Real-GDP Growth as Input: Artificial Intelligence and Regression Approach [a]

		Forecast						Forecast			
		Art Intel		Regression				Art Intel		Regression	
Year	Actual	WOT	WIT	WOT	WIT	Year	Actual	WOT	WIT	WOT	WIT
1970.1	9.36	2.17	1.96	2.02	7.34	1988.2	2.60	3.88	3.91	3.98	-1.38
1970.2	1.12	2.75	2.58	2.64	-1.52	1988.3	2.15	3.47	3.43	3.50	-1.35
1970.3	2.05	2.82	2.80	2.87	-0.82	1988.4	0.47	2.89	2.89	2.95	-2.48
1970.4	1.22	3.07	2.86	2.91	-1.69	1989.1	3.78	2.51	2.44	2.51	1.27
1971.1	7.71	3.53	3.71	3.79	3.92	1989.2	1.25	2.30	2.19	2.25	-1.00
1971.2	8.47	3.41	3.33	3.39	5.08	1989.3	-1.75	2.24	2.11	2.17	-3.92
1971.3	3.82	3.05	2.98	3.04	0.78	1989.4	-4.12	2.23	2.05	2.11	-6.23
1971.4	7.3	3.18	3.09	3.15	4.15	1990.1	-2.43	2.50	2.44	2.51	-4.94
1972.1	8.97	3.80	3.95	4.03	4.94	1990.2	1.89	2.69	2.56	2.62	-0.73
1972.2	1.98	3.66	3.82	3.90	-1.92	1990.3	0.97	2.88	2.68	2.73	-1.76
1972.3	0.16	3.46	3.46	3.53	-3.37	1990.4	0.88	3.01	2.75	2.79	-1.91
1972.4	2.21	3.08	3.16	3.24	-1.03	1991.1	4.61	3.41	3.22	3.28	1.34
1973.1	-2.6	2.67	2.77	2.85	-5.44	1991.2	2.44	3.79	3.76	3.83	-1.39
1973.2	1.71	2.20	2.06	2.12	-0.41	1991.3	3.07	3.78	3.73	3.79	-0.72
1973.3	-3.6	1.31	1.05	1.11	-4.70	1991.4	4.37	3.95	3.91	3.97	0.40
1973.4	-1	1.55	1.37	1.44	-2.47	1992.1	-0.04	4.35	4.45	4.52	-4.56
1974.1	-5.8	1.67	1.35	1.41	-7.24	1992.2	1.90	4.54	4.59	4.66	-2.76
1974.2	4.09	1.64	1.46	1.52	2.57	1992.3	2.57	4.44	4.49	4.57	-2.00
1974.3	6.28	1.90	1.57	1.63	4.65	1992.4	4.97	4.52	4.63	4.70	0.27
1974.4	4.98	2.27	2.04	2.10	2.88	1993.1	2.62	4.29	4.24	4.31	-1.69
1975.1	7.89	3.27	2.97	3.01	4.88	1993.2	4.96	4.09	4.09	4.15	0.81
1975.2	2.14	3.85	3.90	3.96	-1.82	1993.3	3.81	3.82	3.82	3.89	-0.08
1975.3	1.64	3.39	3.46	3.54	-1.90	1993.4	3.38	3.78	3.84	3.91	-0.53
1975.4	4.03	3.70	3.76	3.83	0.20	1994.1	0.66	3.97	3.97	4.04	-3.38
1976.1	5	3.97	4.13	4.21	0.79	1994.2	0.93	4.14	4.23	4.30	-3.37
1976.2	6.24	3.82	3.80	3.86	2.38	1994.3	3.95	3.98	4.02	4.09	-0.14
1976.3	5.82	3.80	3.76	3.83	1.99	1994.4	1.01	3.78	3.80	3.87	-2.86
1976.4	0.02	3.76	3.79	3.86	-3.84	1995.1	2.45	3.23	3.13	3.19	-0.74
1977.1	2.93	3.89	3.97	4.04	-1.11	1995.2	5.17	2.75	2.62	2.68	2.49
1977.2	13.2	3.76	3.87	3.94	9.21	1995.3	2.76	2.70	2.66	2.72	0.04
1977.3	3.2	3.31	3.37	3.44	-0.24	1995.4	4.28	2.48	2.33	2.40	1.88
1977.4	4.48	3.07	2.93	2.99	1.49	1996.1	6.59	2.72	2.63	2.70	3.89
1978.1	0.64	4.40	4.51	2.37	6.30						

	Art Intel		Regression	
	WOT	WIT	WOT	WIT
MSE	10.02092	10.00789	10.00717	9.901107
SD	16.91955	16.61081	16.59822	5.442661

a. Art Intel = Artificial Intelligence; WOT = Without Lag of Real-GDP Growth; WIT = With Lag of Real-GDP Growth; MSE = Mean Square Error; SD = Standard Deviation Error.

yield spread from four quarters prior to the dependent variable was used as the independent variable. Two methods of in-sample simulation were performed. First, using the estimated weights (parameters) of the neural network the real-GDP growth was estimated and compared with those of the actual. This comparison is shown in Table 2 and Figures 3 and 4 for each of those two mixes. Second, the traditional regression forecast of the real-GDP growth was estimated and compared with those of the actual for the same input mixes.[7] This comparison is shown in Table 2, and Figures 5 and 6.

The mean square error (MSE) and the standard deviation (SD) of the error for both the artificial neural networks and regression methodology are calculated and shown at the end of Table 2. Accordingly, the MSE and the SD of the error for the artificial intelligence method and that of the regression for the input mix, without the lag of the real-GDP growth, are fairly similar in value. However, the MSE and SD of the error for the regression with the input mix of the lag of the real-GDP growth are slightly lower.

Out-of-Sample Simulation

As a better test of the forecast accuracy of the artificial neural networks and the regression models, the out-of-sample simulation was performed. The data was divided into two equal parts—1959.1 to 1977.3 and 1977.4 to 1997.1. The first data set was used to estimate the parameters of each model (neural network, and regression). Then, the input data of the second part were injected into each of the estimated models to forecast the output (real-GDP growth).

The estimated forecast of the real-GDP growth of the neural networks is compared with that of actual real-GDP growth in Table 3 and Figures 7 and 8 for two input mixes. Similarly, the estimated forecast of the real-GDP growth of the regression approach is compared with that of actual growth in Table 3 and Figures 9 and 10 for two input mixes.[8]

For comparison, the MSE and the SD of the error for both artificial neural networks and regression method are calculated and shown at the end of Table 3. Accordingly, the MSE for the artificial intelligence method for each of the input mixes (12.8 and 13.68) is far lower than those of the regression approach (28.33 and 37.84). Thus, the findings of the out-of-sample simulation strongly suggest that the real-GDP growth prediction of the artificial intelligence approach is far more accurate than that of the traditional regression method.

Furthermore, the resulting MSE was higher for each of the input mixes with the inclusion of the lag of growth for both the neural networks and the regression methods. This outcome suggests that adding a variable(s) in the input mix (over-fitting) does not necessarily increase the accuracy of the forecast of the out-of-sample simulation. In fact, over-fitting can result in less accuracy, as was the case here with a higher MSE.[9]

In addition, the SD of the forecast error of the artificial intelligence model and that of the regression method were calculated and shown in the Table 3. The SD for the model without real-GDP lag is about equal for the artificial intelligence method compared to that of the regression method, while the SD for the model with real-GDP

FIGURE 3
Actual and In-Sample Forecast of Real-GDP Growth with Yield Spread as Input: Artificial Intelligence

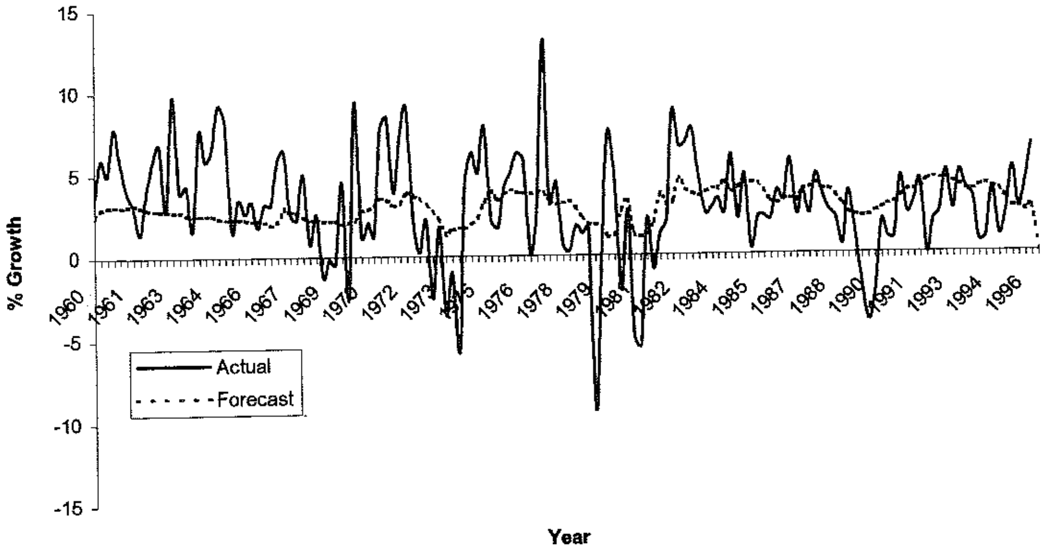


FIGURE 4
Actual and In-Sample Forecast of Real-GDP Growth with Yield Spread and GDP as Inputs: Artificial Intelligence

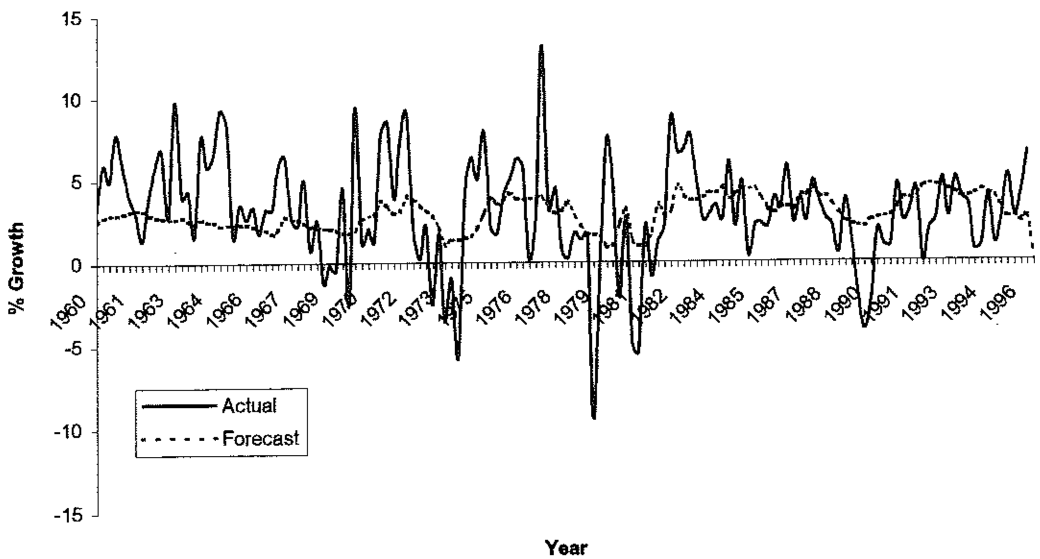


FIGURE 5
Actual and In-Sample Forecast of Real-GDP Growth with Yield Spread as Input: Regression Method

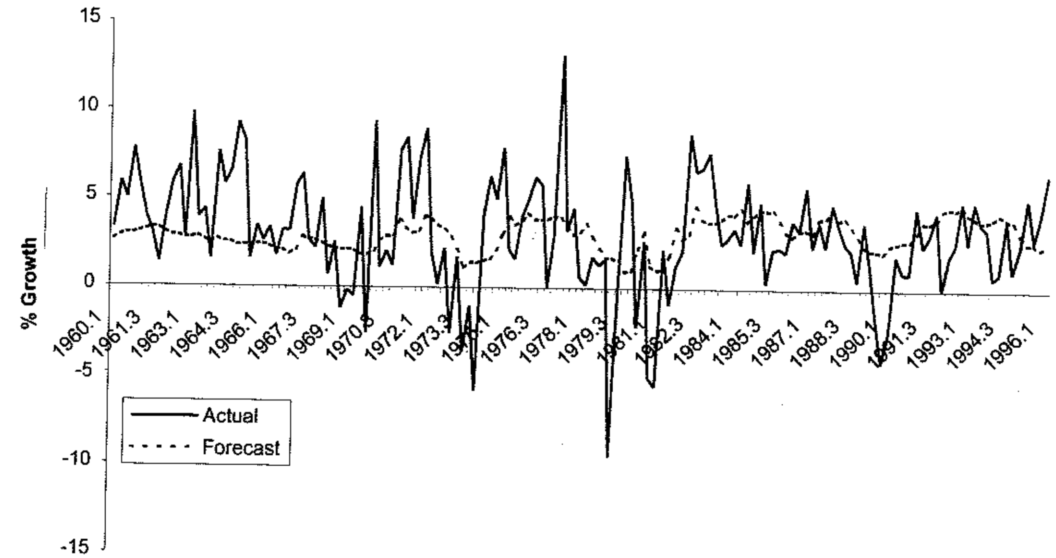


FIGURE 6
Actual and In-Sample Forecast of Real-GDP Growth with Yield Spread GDP as Inputs: Regression Method

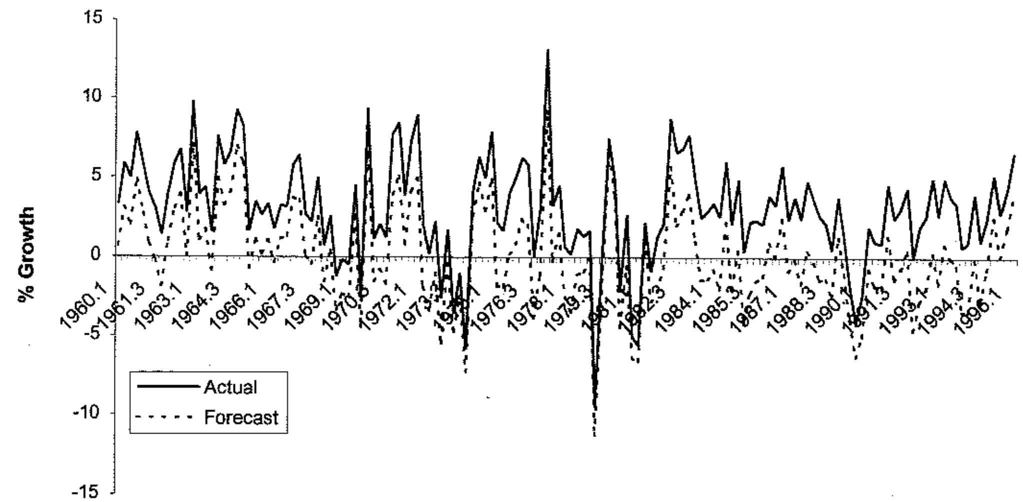


TABLE 3
Actual versus Out-of Sample Forecast of Real-GDP Growth with Lag of Yield Spread and Real-GDP Growth as Inputs: Artificial Intelligence and Regression Approach, 1978.1 to 1996.1 [a]

		Forecast						Forecast			
		Art Intel		Regression				Art Intel		Regression	
Year	Actual	WOT	WIT	WOT	WIT	Year	Actual	WOT	WIT	WOT	WIT
78.1	0.6415	4.4006	4.5086	2.3736	6.3033	87.2	3.79	5.5878	5.8642	-0.51	7.4416
78.2	0.18	4.649	4.5204	-2.647	19.82	87.3	2.4	5.7645	6.0847	-0.319	6.702
78.3	1.76	3.9011	3.9186	-3.261	6.6279	87.4	4.82	6.2251	6.5523	-2.41	9.9356
78.4	1.3	2.8364	2.6396	-0.429	8.3322	88.1	3.6	5.6099	5.9298	0.1247	5.452
79.1	1.64	2.2378	2.047	0.6345	3.249	88.2	2.6	5.6099	5.8917	-0.651	7.3833
79.2	-9.49	2.2378	2.0598	1.8452	2.961	88.3	2.15	4.9301	5.1394	-1.9	5.5598
79.3	0.03	1.8921	1.6122	-12.62	4.7749	88.4	0.47	3.958	3.9405	-1.507	8.8055
79.4	7.48	0.9474	0.5236	0.3196	3.9531	89.1	3.78	3.3184	3.2272	-2.132	7.0983
80.1	4.84	0.9819	0.5543	11.404	4.4781	89.2	1.25	2.9628	2.8394	3.0913	5.8499
80.2	-2.19	3.5359	3.8415	7.9372	-10.03	89.3	-1.75	2.8594	2.731	0.3102	5.3414
80.3	2.66	4.6039	4.8241	-5.457	2.4136	89.4	-4.12	2.8479	2.7642	-3.444	3.1898
80.4	-5.03	1.1776	0.6205	-0.436	12.478	90.1	-2.43	3.307	3.2088	-6.564	7.5146
81.1	-5.52	0.9014	0.3722	-4.962	9.0039	90.2	1.89	3.6159	3.6393	-4.947	4.0477
81.2	2.18	1.2121	0.9287	-5.276	-0.506	90.3	0.97	3.9466	4.1074	0.1026	0.1103
81.3	-0.91	2.3185	2.0851	3.9981	5.9911	90.4	0.88	4.1626	4.4236	-1.699	-3.019
81.4	1.27	5.2098	5.6628	-1.515	-4.232	91.1	4.61	4.8291	5.1521	-2.227	-0.893
82.1	2.12	4.242	4.554	-3.329	-4.904	91.2	2.44	5.4658	5.775	1.6698	4.8776
82.2	8.76	4.276	4.3838	-0.786	5.0791	91.3	3.07	5.4547	5.7866	-1.956	3.6439
82.3	6.6	6.8091	7.3966	8.0547	1.0584	91.4	4.37	5.7315	6.1094	-1.153	3.4876
82.4	6.84	5.6762	6.0352	1.3394	3.9321	92.1	-0.04	6.3883	6.7711	0.1155	8.5427
83.1	7.68	5.3324	5.6142	3.4183	5.0308	92.2	1.9	6.7017	7.188	-6.469	5.6198
83.2	4.74	4.9862	5.0332	5.0554	13.888	92.3	2.57	6.5291	6.9732	-4.491	6.4328
83.3	2.5	5.3769	5.5463	1.9929	11.099	92.4	4.97	6.6694	7.1015	-3.343	8.0816
83.4	2.86	5.8196	6.0544	-1.581	11.406	93.1	2.62	6.2905	6.7786	-0.394	2.3223
84.1	3.39	5.6873	5.8784	-1.773	12.5	93.2	4.96	5.9626	6.3498	-3.025	4.8184
84.2	2.54	6.3774	6.7551	-0.857	8.6403	93.3	3.81	5.5213	5.8213	0.5693	5.7365
84.3	6.04	5.4436	5.7331	-3.099	5.5802	93.4	3.38	5.4436	5.6673	-0.237	8.9201
84.4	2.11	5.8747	6.223	2.772	6.1687	94.1	0.66	5.7535	6.0891	-0.596	5.8933
85.1	4.88	6.3774	6.7901	-2.966	6.7893	94.2	0.93	6.0393	6.3585	-4.674	8.9771
85.2	0.33	6.3231	6.7494	-0.106	5.797	94.3	3.95	5.7756	6.0832	-4.674	7.371
85.3	2.23	6.1924	6.5074	-5.96	10.367	94.4	1.01	5.4547	5.7225	-0.396	6.8871
85.4	2.32	5.488	5.7949	-3.179	5.1715	95.1	2.45	4.5362	4.7282	-3.744	3.2521
86.1	2.08	4.4458	4.5082	-2.138	8.8383	95.2	5.17	3.7187	3.7681	-0.584	3.5308
86.2	3.86	4.2193	4.3682	-0.779	2.7755	95.3	2.76	3.6387	3.5917	4.1936	7.5901
86.3	3.27	4.5927	4.7513	1.7138	5.3028	95.4	4.28	3.2726	3.2452	1.2939	3.6618
86.4	5.77	4.7279	4.9063	0.4505	5.35	96.1	6.59	3.673	3.673	3.7169	5.4981
87.1	2.35	4.4345	4.5712	3.4949	5.1313						

	Art Intel		Regression	
	WOT	WIT	WOT	WIT
MSE	12.8	13.678	28.333	37.842
SD	3.9751	4.5571	3.7156	21.824

a. Art Intel = Artificial Intelligence; WOT = Without Lag of Real-GDP Growth; WIT = With Lag of Real-GDP Growth; MSE = Mean Square Error; SD = Standard Deviation Error.

FIGURE 7
Actual and Out-of-Sample Forecast of Real-GDP Growth with Yield Spread as Input: Artificial Intelligence

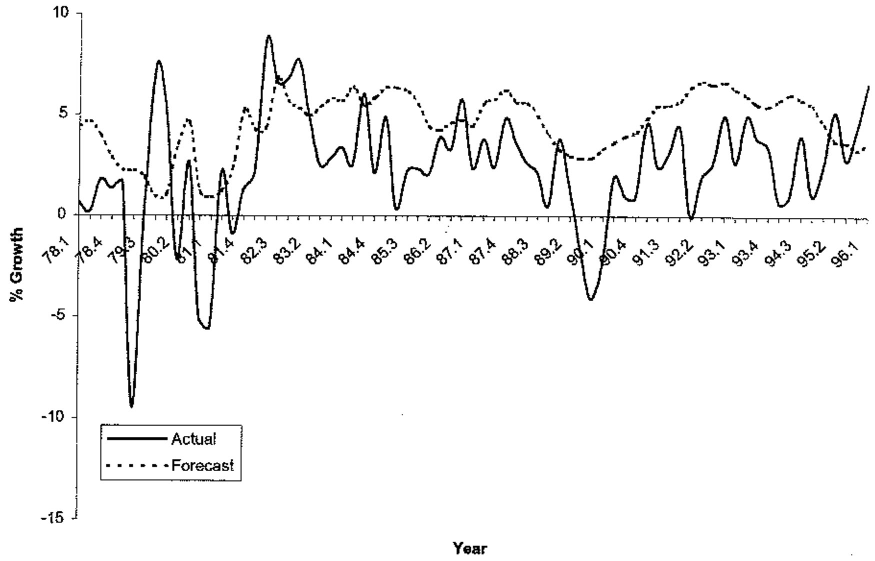


FIGURE 8
Actual and Out-of-Sample Forecast of Real-GDP Growth with Yield Spread and GDP as Inputs: Artificial Intelligence

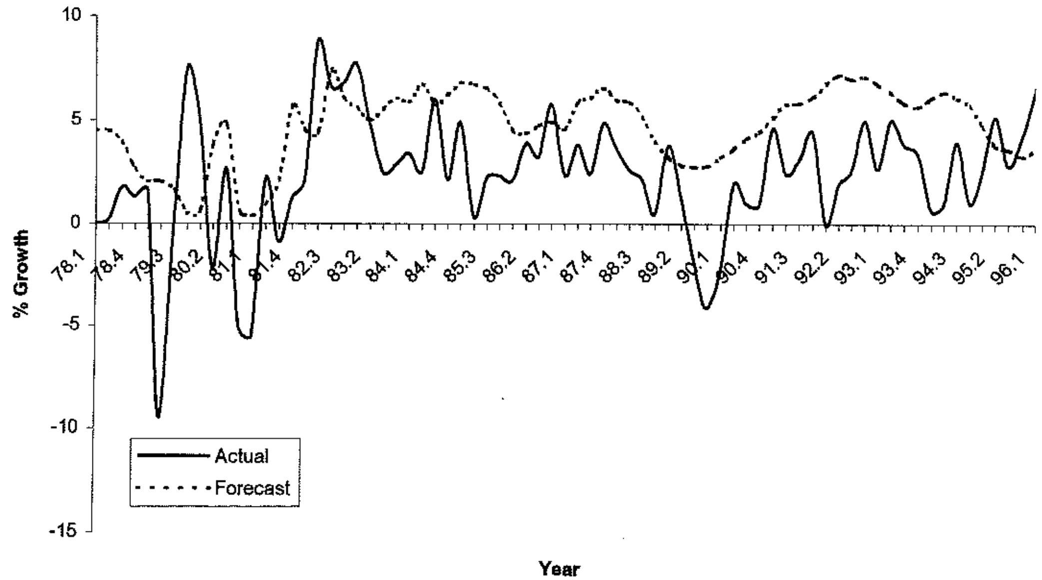


FIGURE 9
Actual and Out-of-Sample Forecast of Real-GDP Growth with Yield Spread as Input: Regression Method

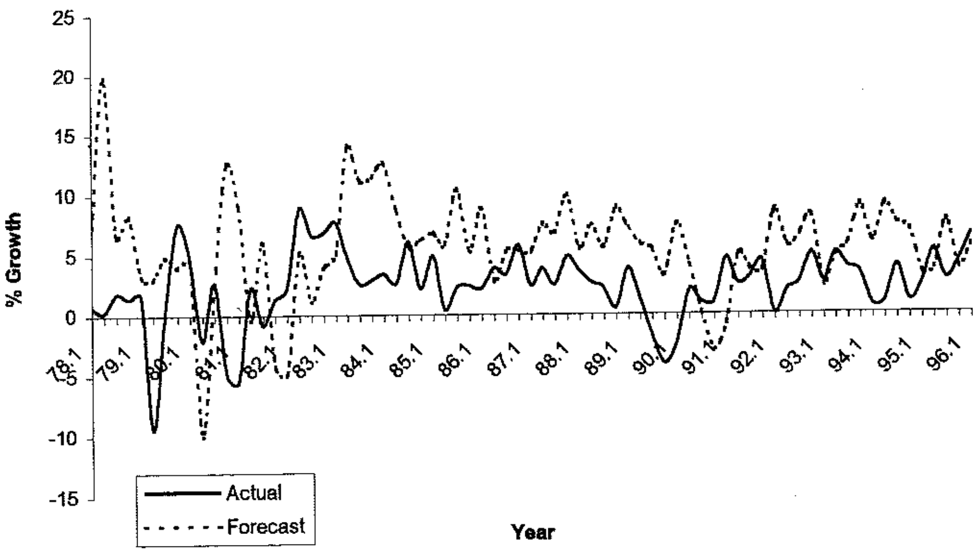
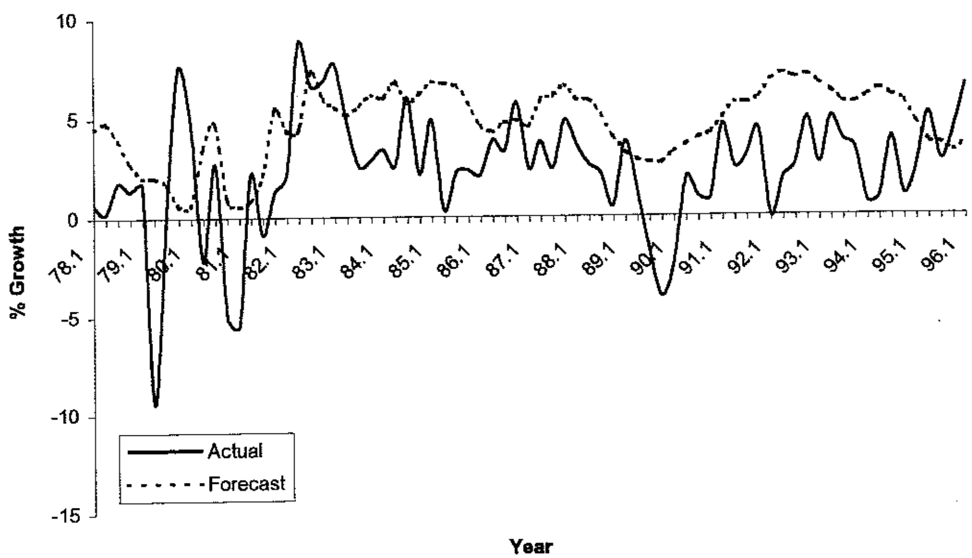


FIGURE 10
Actual and Out-of-Sample Forecast of Real-GDP Growth with Yield Spread and GDP as Inputs: Regression Method



lag is lower for the artificial intelligence approach compared to that of the regression method. Thus, the findings of the out-of-sample forecasts suggest that the artificial intelligence method is more accurate with less error and less variation than those of the regression approach.

SUMMARY AND CONCLUSIONS

Economists typically utilize the shape of the yield curve to predict recessions. Using econometric models they concluded that a flat or a downward-sloping yield curve signals a slowdown in economic activity in the near future. This study used one branch of artificial intelligence model called neural networks as another method to confirm that conclusion and to compare the results with those of the econometric models.

The computation of the artificial network starts with loading a set of values of independent variables into the input layer of the network for each observation. Then, each hidden node calculates the weighted summation of the inputs and incorporates the threshold, or bias, weight. Next, each hidden node computes the sigmoid function of the summation, which represses the value of the summation to a range between one and zero. After that, each hidden node sends the results to the output nodes. In the output layer, computations similar to those in the hidden layer take place, and finally, in that layer, the value of the dependent variable(s) is estimated. The networks then compute the error as the difference between the estimated and the actual dependent variable(s). The output node propagates the amount of its error on each observation back through its links to the hidden nodes. This process continues and, based on this error minimization, the weights are determined.

Using quarterly data from the first quarter of 1959 to the first quarter of 1997, the findings confirm those of the earlier studies employing econometric methodology that a flat or negative-sloping yield curve is a reliable predictor of recession within a year. Furthermore, the findings of the out-of-sample simulation suggest that the forecast of the artificial neural networks is more accurate with less error and lower variation than those of the regression.

NOTES

The author gratefully acknowledges two anonymous referees whose comments helped to improve the content and style of this work. The usual disclaimer applies. This research was made possible by the Office of Sponsored Research and Grants of the University of Central Oklahoma.

1. For more discussion of the theoretical basis of the yield curve as a predictor of recession see Kozicki[1997] and Dueker [1997].
2. The yield curve is also used both to gauge whether or not monetary policy is expansionary or not and to predict inflation. See for example, Frankel and Lown [1994], and Mishkin [January 1990; August 1990; 1991].

There are three basic questions regarding the topology of the artificial neural networks. The first is the number of hidden layers. When using a functional link, the first step is to use no hidden layer, and to connect the processing elements to both the inputs and the functional link. When accuracy is not achieved, a functional link layer should not be used, and one hidden layer can be added. Normally, most problems such as those with a small number of inputs and one output can be solved with a single hidden layer, using conventional statistical methods. A theoretical upper boundary for the number of hidden units can be determined by the number of training observations. The rule is that there should be at least five observations for each weight.

3. The second question is the number of nodes in each hidden layer. Generally, the fewer the number of processing elements in the hidden layers, the better the network will "generalize." Generalization is the power to learn and interpolate between previously seen examples generalized to the future prediction.

Finally, the third question relates to the learning rate and the learning rule. The learning rate for the last hidden layer must be twice that of the output layer. The delta rule measures error and changes the weights for each observation. Two extensions of the delta rule are the cumulative delta rule, which accumulates weight changes over several observations, and the normalized-cumulative delta rule. In general, the normalized cumulative delta rule works well. Furthermore, the *epoch*, which is the number of presentations over which weight changes are accumulated, is set to 16, unless the data are very noisy. In this study, the normalized-cumulative delta rule, with and *epoch* of 16 was considered appropriate.

4. The elasticity concept, called "*dithering*" in the artificial neural network literature, is not necessarily linear.

5. Theoretically, the real business cycle moves with real interest rates. Since real rates are not available, it is assumed that the spread between nominal rates is approximately equal to the spread between real rates. Furthermore, interest rates and spreads are measured on a quarterly average basis. While the precise starting date does not seem to be crucial, the date chosen maximizes the availability of comparable data for all series.

6. The regression models used are as follows:

(1N) $$\Gamma(R - GDP)_t = \alpha + \beta(\textit{Treasury Spread})_{t-4}$$

(2N) $$\Gamma(R - GDP)_t = \alpha + \beta(\textit{Treasury Spread})_{t-4} + \gamma\Gamma(R - GDP)_{t-4}$$

where α, β, and γ are the parameters of the equations, and Γ is the rate of growth. A similar structure is used for the artificial intelligence model.

7. The estimated results of equation (1N) for the whole period of the first quarter of 1959 to the first quarter of 1997 are as follows:

Variable	Estimated Coefficient	Standard Error	t-statistic
Constant	1.98934	0.398767	4.98873
(Treasury Spread) $_{t-4}$	0.733436	0.214375	3.42128

and the estimated results of equation (2N) for the same period are as follows:

Variable	Estimated Coefficient	Standard Error	t-statistic
Constant	1.90886	0.446987	4.27051
(Treasury Spread) $_{t-4}$	0.722515	0.216701	3.33415
$\Gamma(R-GDP)_{t-4}$	0.032472	0.080549	0.403139

Accordingly the '*t*' ratios for *(Treasury Spread)* $_{t-4}$ are statistically significant for both equations. However, the '*t*' ratio for $\Gamma(R-GDP)_{t-4}$ is not significant.

8. The estimated results of equation (1N) for the first quarter of 1959 to the third quarter of 1977 is as follows:

Variable	Estimated Coefficient	Standard Error	t-statistic
Constant	2.37356	0.556824	4.26268
(Treasury Spread) $_{t-4}$	1.29820	0.404322	3.21080

The estimated results of equation (2N) for the same period is as follows:

Variable	Estimated Coefficient	Standard Error	t-statistic
Constant	2.45043	0.643621	3.80726
(Treasury Spread) $_{t-4}$	1.32126	0.417947	3.16132
$\Gamma(R-GDP)_{t-4}$	-.028728	0.118199	-.243050

Again, the results of the '*t*' ratios for the *(Treasury Spread)* $_{t-4}$ are statistically significant for both equations. But, the '*t*' ratio for the $\Gamma(R-GDP)_{t-4}$ is not significant.

9. Estrella and Mishkin [1995] came to the same conclusion from their regression equations.

REFERENCES

Campbell, J.Y. Some Lessons from the Yield Curve. *Journal of Economic Perspectives*. Summer 1995, 129-52.

Dueker, M. J. Strengthening the Case for the Yield Curve as a Predictor of U.S. Recessions. *Economic Review*. Federal Reserve Bank of St. Louis. March/April 1997, 41-51.

Estrella, A. and Harduvelis, G. A. The Term Structure as a Predictor of Real Economic Activity. *Journal of Finance*, June 1991, 555-76.

Estrella, A. and Mishkin, F. S. Predicting U.S. Recessions: Financial Variables as Leading Indicators. National Bureau of Economic Research, Working Paper No. 5379, 1995, 1-32.

Fisher, I. *The Rate of Interest*. New York: Macmillan. 1907.

Frankel, J. and Lown, C. An Indicator of Future Inflation Extracted from the Steepness of the Interest Rate Yield Curve Along its Entire Length. *The Quarterly Journal of Economics*. May 1994, 517-30.

Harvey, C. R. The Term Structure and Consumption Growth. *Journal of Financial Economics*. December 1988, 305-33.

________. Forecasts of Economic Growth from the Bond and Stock Markets. *Financial Analysts Journal*. September/October 1989, 38-45.

________. Term Structure and World Economic Growth. *Journal of Fixed Income*. June 1991, 7-19.

________. Term Structure Forecasts Economic Growth. *Financial Analysts Journal*. May-June 1993, 6-8.

Haubrich, J. and Dombrosky, S. Predicting Real Growth Using the Yield Curve. *Economic Review*. Federal Reserve Bank of Cleveland, Quarter 1, 1996, 26-35.

Hsieh, C. Some Potential Applications of Artificial Neural Systems in Financial Management. *Journal of Systems Management*. April 1993, 12-15.

Hu, Z.. The Yield Curve and Real Activity. *International Monetary Fund Staff Papers*. December 1993, 781-806.

Kozicki, S. Predicting Real Growth and inflation with the Yield Spread. *Economic Review*, Fourth Quarter 1997, 39-57.

Mishkin, F. What Does the Term Structure Tell Us About Future Inflation? *Journal of Monetary Economics*, January 1990, 77-95.

________. The Information in the Longer-Maturity Term Structure About Future Inflation. *Quarterly Journal of Economics*. August 1990, 815-28.

________. A Multi-Country Study of the Information in the Term Structure About Future Inflation. *Journal of International Money and Finance*. 19 March 1991, 2-22

NeuralWare, Inc. Neural Computing, *Neural Works Professional II/PLUS and Neural Works Explorer*. NeuralWare, Inc., Pittsburgh, PA, 1991.

Shaaf, M. Fiscal and Monetary Actions: A Neural Network and Econometric Comparison of Their Relative Importance. *Studies in Economics and Finance*. Fall 1996, 69-87.

Shaaf, M. and Ahmadi, A. An Artificial Intelligence Approach to the Role of Exports in Economic Development of Malaysia. *Atlantic Economic Journal*. December 1999, 363-75.

Shaaf, M. and Erfani, G. Air-Pollution and the Housing Market: A Neural Network Approach. *International Advances in Economic Research*, November 1996, 484-95.

Swales, G. S. and Yoon, Y. Applying Artificial Neural Networks to Investment Analysis. *Financial Analyst Journal*. September-October 1992, 70-80.

Trippi, R. and Turban, E., eds. *Neural Networks in Finance and Investing*. Probus Publishing Company, Chicago, 1993.