Determinants of Corporate Borrowing: An Application of a Neural Network Approach

Jandhyala Sharma, Ravi Kamath and Sorin Tuluca

1. INTRODUCTION

The evidence presented in finance literature strongly suggests that the costs associated with the adoption of a non-optimal capital structure might be substantial. A non-optimal capital structure decision, for example, might result in higher direct and indirect economic costs to firms in form of lower stock price (Masulis (1983), DeAngelo and Masulis (1980)), higher cost of capital and lost growth opportunities (Myers and Majluf (1984)), increased probability of bankruptcy (Warner (1977,1983), (Castanias (1983)), higher agency costs (Jensen and Meckling (1976), (Barnea, Haugen and Sunbet (1985)), and possible wealth transfers from one group of investors to another. The underpinnings of the modern capital structure theories are robust and intellectually appealing, and consequently, this subject is able to maintain a sustained and an abiding interest of researchers even after nearly five decades of intense research in this area.

The capital structure theories forwarded in the literature are predicated upon the existence of imperfections in the capital market such as taxes, bankruptcy and agency costs, asymmetric information between insiders and outsiders and signaling costs.[1] Modigliani and Miller (1958) proved that in a perfect capital market, there is no relationship between the quantum of debt held by a firm and its value and moreover, the debt financing decisions have no bearing on the values of firms (and cost of capital). However, the existence of optional debt ratios for firms have been contended by numerous authors. Miller (1977) has argued in favor of the existence of optimal debt for the economy as a whole if not at a micro level. But interestingly, neither the investors nor managers can observe directly the optimal level of debt for a firm nor can they convincingly claim that the observed debt ratio is an optimal one. Empirically, it is infeasible to test either the existence of optimal capital structure for firms or the effect of debt on the value of the firm.

The main purpose of this paper is to revisit the capital structure decisions using the neural networks methodology. More specifically, the study compares the results derived from the classical linear statistical methods with neural networks results in explaining the observed cross-sectional debt ratios. Additionally, this paper attempts to determine if the neural networks in general, have promising applications to capital structure studies and in particular, whether it has better ability to predict debt ratios of firms.[2] The study is organized as follows: Section 2 briefly reviews the capital structure literature relevant to the present paper. The methodology and the data used in this paper are described in the third section. The empirical results are analyzed in Section 4. The summary of the paper and conclusions make up the final section.

2. CAPITAL STRUCTURE DECISIONS REVISITED

It is generally accepted that a firm value is determined by the value of the assets in place and the present

Dr. Jandhyala Sharma is Associate Professor of Finance, Cleveland State University, OH. His articles have appeared in *Financial Management, Applied Financial Economics,* among others.

Dr. Ravi Kamath is Professor of Finance, Cleveland State University, OH. His publications have appeared in such journals as *Financial Review* and *Journal of Economics and Finance.*

Dr. Sorin Tuluca is Assistant Professor of Finance, Fairleigh Dickinson University, NJ. He conducts research in the area of Derivative Securities.

value of its future growth opportunities. The separation theorem says that the firm's financing decisions can be separated from the investment decisions. Capital asset pricing models consider that the cost of capital is exogenously determined by firm's systematic risk, the return on the market portfolio, and the risk free rate. Further, the cost of capital of a firm is assumed to be influenced by the level of debt in its capital structure.

In spite of the fundamental theorem of value of a firm, the subject of the existence or the lack of an optimal capital structure for a firm remains a controversial topic. Many studies have concentrated their empirical research on the determinants of the level of debt (or observed debt ratios) of firms, to explain the cross sectional regularities in the level of debt. (See Friend and Hasbrouck (1988), Bradley Jarrell and Kim (1984), and Titman and Wessels (1988), among others). Even here there is no consensus, since the variables selected, the signs and statistical significance of the estimated regression coefficients vary from study to study and therefore, the resulting conclusions. For example, while Baskin (1989) found that "....debt leverage varies positively with past growth and inversely with past profits," Kim and Sorensen (1986) concluded that firms which have experienced higher growth tend to rely on less debt instead of more debt. The results of various studies support a number of different hypotheses, such as tax based theories (Modigliani and Miller (1963), Hamada and Scholes (1985), DeAngelo and Masulis (1980)), agency theory (Jensen and Meckling (1976), Myers (1977), and non-trivial bankruptcy costs (Warner (1977)), non debt tax shields (DeAngelo and Masulis (1980), (Brennan and Schwartz (1978)), and information content and signaling value of changes in the debt ratios,(Leyland and Pyle (1977), Noe (1988), Narayanan (1987), and Ross (1977)). On the whole, it appears, that the capital structure "puzzle" is still in a state of flux and evolving without any definitive theory or empirical guidance to the managers at the operational level (Myers (1984)).

In addition to the theoretical studies and the empirical investigations conducted to validate the theoretical postulates, there are sample surveys of corporate financial managers to gather their views and practices (see for example, Norton (1989), Pinegar and Wilbricht (1989), Pruitt and Gitman (1991) and Kamath (1997)). In spite of only a small number of these surveys, their contribution can be viewed as attempting to create a linkage between the theoretic foundation, evidence emerging from empirical studies and the opinions and desires of financial decision makers.

For quite sometime, the methodology used in empirical research on capital structure theories has remained stagnant and not pursued new directions in addressing the crucial issues. This melody of not attempting significantly different methodology and not offering new insights into the theories continues to plague the capital structure, the cost of capital and related other issues. The limited success of the prior studies can be partly attributed to the methodology adopted in these studies. Classical statistical models require strong assumptions as to the distributional properties of the sample data (normal or log-normal distributions) and a well-defined functional relationship between a dependent variable and a set of uncorrelated independent variables (and many accounting variables used are highly correlated). Since much of the financial data might not possess these properties, the linear models are sensitive to specification errors.[3]

In the area of capital structure research like in many other areas in finance, the data is discrete, usually yearly intervals, cross sectional, noisy and highly correlated and non-normal. These properties often violate the basic assumptions of the underlying classical statistical models. This paper applies the neural network model to study the determinants of observed debt ratios. The paper compares the results of the neural networks model with those obtained by classical statistical inference. The basic rationale for the use of neural network is that no distributional assumptions are made about the sample data, such as normal or lognormal or sample path continuity. Moreover, this methodology does not require a priori specification of functional relationship between the variable of primary interest, namely, the dependent variable (debt ratio) and a set of independent variables. In spite of these relaxed assumptions about the sample data and the functional relationship, the neural network models are robust to specification errors. The neural network (NNW hereon) model observes the data and formulates an internal representation of the relationship between the input variables and the output variable. The NNW modeling essentially involves systematic pattern recognition of an outcome given a set of inputs (independent variables), which are independent of data generating process. Further, the NNW can process a large number of input variables even if the data is noisy and highly correlated.

Neural network process has been applied to solve problems in diverse fields ranging from speech and image recognition to the field of artificial neural nets to finance. The examples of NNW applications in finance include pricing of initial public offerings, IPOs (Jain and Nag (1995)), pricing and hedging derivative securities (Hutchinson, Lo and Poggio (1994)), predicting thrift failures (Salchenberger, Cinar and Lash (1992)), predicting bank failures (Tam and Kiang (1992)), bond ratings (Dudda and Shekhar (1988)), comparing Discriminant Analysis to artificial neural network performance (Yoon ,Swales and Margavio (1993)), and forecasting bankruptcy (Fletcher and Goss (1993)), among others. In a nutshell, the emerging popularity of this methodology can be attributed to its underlying technique of pattern recognition. In praising this methodology, Hammerstorm (1993) says that "The networks can recognize spatial, temporal, or other relationships and can perform such tasks as classification, prediction, and function estimation" (p.46). This assessment of the methodology and its application potential provides the motivation for the present investigation. In the next section, a brief introduction to the methodology is provided to pave the path to the empirical application itself.[4]

3. METHODOLOGY AND DATA

Methodology

An attempt is made here to describe the NNW approach itself and its merits. However, a rigorous explanation of the method and the underlying architecture is beyond the scope of this paper. The readers who are interested in the history as well as in the technical aspects of the method, Kilmasaukas (1988) and Hecht-Nielsen (1989) are highly recommended. This nonlinear approach parallel processes a number of different inputs and outputs. This attribute of NNW makes it an ideal method for capital structure studies. NNW modeling involves using a database consisting of a vector of input and output variable(s) to recognize a systematic relationship, if any, between input variables and output variable(s). The NNW learns by adjusting the weights to minimize the error of output(s). A feed forward back-propagation learning process constantly changes the weights by continuous iterations. A NNW consists of several layers of neurons; and an input layer, one or more hidden layers and an output layer. Each layer is connected to the previous layer by weights (synoptic nerves or perceptions). The input of one layer is fed to the next layer and the output of one layer feed the next layer or the output layer.

To understand the methodology, consider a simple NNW consisting of three layers - an input layer, a hidden layer and an output layer. The first layer, the input layer, distributes the inputs to the hidden layer and does not have any activation function. The learning process makes it more adoptive and responsive to structural changes in the data generating process.[5] For example:

1. The output of the hidden layer (treating the bias as another input)
   $h(j) = \Sigma\ (w(i,j) \times\ i(i)),\ i = 1, N$
   $s(j) = f\ (h(i))$
2. The output layer calculate:
   $h'\ (k) = \Sigma\ (w'\ (j,k) \times s(j))\ ,\ j = 1, M$
   $O(k)\ = f\ (h'(k))$

where: i(i) - are network input, (N= the number of inputs used in the network)
O(k) -are the network inputs (nodes) in the hidden layer
w(i,j) - are the weights connecting the neutrons in layer i to neutrons j in layer 2.
w' (j,k) -are the weights connecting the neutrons in layer 2 to neutron k in the next layer.
f (x) is the neuron transfer function[6], (the sigmoid function is the one used in the model)

A Sigmoid is the basic function given by $f(x) = 1 / (1+ e^{-ax})$

where "a" is the slope parameter. Of the Sigmoid function, by varying the slope, "a", one can obtain the sigmoid functions of different slopes. When "a" approaches zero (a ÿ 0,), the sigmoid function becomes a threshold function. The sigmoid function assumes continuous range of values from 0 to 1. An interesting property of sigmoid function is that it is differentiable, and differentiability is an important feature of neural network modeling (Haykin (1994), p.12),

Since the NNW constantly reestimates by adjusting the weights to minimize the error of output(s), the object of the choice of weights is to minimize the mean squared error $- \Sigma\ (\text{Actual (i)} - \text{Predicted (i)})^2$ summed over all the N values.

The Squared Mean Error $r = \Sigma\ (\Sigma(t\ (p,k) - O\ (p,k))$ ^2), $k = 1, N,\ p = 1, P$

Where O (p,k) is the Neural network output k for pattern p.
t (p,k) is the output training pattern p for output k.

Neural networks as information processing system is widely recognized in a host of different fields, including in as dissimilar fields as neurobiology and finance. A NNW system can be trained by feeding a set of data that is representative of the environment. If the environment is non-stationary, the statistical parameters of information generating process varies with time. In such cases, the traditional linear statistical methods used to generate parameters generally prove to be inadequate. On the other hand, the NNW modeling can easily overcome such shortcomings since an NNW can be trained to handle continuity, that is, to adopt to respond to every distinct input as a novel one. In a nutshell, the NNW relies on its " memory " to recall and exploit the past experience. This amazing attribute of the NNW methodology provides the underlying impetus for this study.

This study applies the novel and powerful approach of NNW to determine if there exists a systematic relationship between a set of generally accepted financial and accounting input variables and the observed cross-sectional measures of corporate use of debt in their capital structures. This investigation concurrently uses the linear statistical model relied upon by a host of previous studies on the subject. This procedure facilitates a comparison of the performance of the two models. In this study various hypotheses are tested using parametric as well as non-parametric tests. Within this framework, the estimated mean debt ratio for the sample by NNW is compared to the observed mean debt ratio of the sample and to the mean debt ratio generated by a linear statistical model. The sample size is large enough to permit additional testing with a holdout sample. Accordingly; the predictive powers of both models are tested with the help of the holdout sample.

Data

The sample data consists of all non-financial companies traded on the New York Stock Exchange, and for which financial information is available in an uninterrupted fashion on the COMPUSTAT tapes for the sample period. The 18-year sampling period spans from January 1973 to December 1990. For any firm to be included in the sample, the information had to be available to compute the 14 variables listed and described in Table 1. Thus, stock prices and bond prices were deemed necessary among a host of other characteristics. This procedure yielded a sample of 151 firms.[7]

Of the 14 variables listed in Table 1, first ten variables are utilized as input variables and the last four as output variables. Thus, the investigation attempts to ascertain if there is a systematic relationship between the first ten variable and the next four. For reasons of space, and to avoid rehashing, the description of the variables themselves as contained in Table 1 is considered adequate. The chosen variables are generally accepted in the literature[8] as good proxies of the financial health, performance, size, risk, liquidity, growth, and stockholder cashflows.

All values of the variables used in the analysis are averages of eighteen years. The rationale for averaging can be explained as follows. About one-third of NYSE listed firms admit that in raising new funds, they seek to maintain a target capital structure (Kamath (1997) and (Pinegar and Wilbricht (1989)), whether based on the book values or the market values. Yet, it is not always possible to be at the precise target at all points in time because of the constantly changing market values of common stock, preferred stock and bonds. Moreover, this task of achieving and maintaing a target capital structure is further complicated by the fluctuations in interest rates, costs of access to capital markets, and the economic costs of continuous rebalancing of the debt ratio. Thus, even though it is infeasible to maintain an optimal capital structure or a target debt ratio at all points in time, it is conceivable that firms actually strive to be around the desired long-term average debt ratio. Additional support for averaging procedure comes from the responses of the managers of the NYSE firms to the inquiry by Kamath (1997) regarding the dependence of their firm's debt ratios. The top three responses were found to be the "past profits", "average debt ratio in their respective industries", and "past growth". Clearly, these determinants according to the financial managers suggest long-term averages of the important determinants rather than their short-term values.

TABLE 1
Input Variables Used to Train Neural Network and to Estimate the Regression coefficients

Variable Description	Name
(1) Log of total assets	SIZE
(2) Non-debt tax shields (Depreciation/total assets)	NDT
(3) Tangible Assets Ratio (Inventories + Gross Plant)/ total assets	TAR
(4) Growth rates (Year to year Percentage change in sales)	GROW
(5) R &D/Total Assets (Research and Development)	TECH
(6) Risk (Standard deviation of net income)	RISK
(7) Net Operating Income /Total assets	PROF
(8) Cash flows/net worth	CSFLOW
(9) Earnings / Price (Inverse of P/E ratios)	EARYLD
(10) Dividend payout ratio (Dividends) / Earnings)	DIVPAY
(11) Debt Ratio (Debt/ Debt + Equity), using market values	DEBTMV
(12) Debt ratio (Debt)/ (debt + equity), using book values	DEBTBV
(13) Long-term debt , using book values	LTDBV
(14) Long-term debt, using market values (Market values)	LTDMV

1. All values are the averages of 18 year period (1973-1990). The RISK variable is a Standard Deviation of 18 annual net income figures.

4. FINDINGS

The descriptive statistics of all 14 variables utilized in this study for the 151 sample firms are exhibited in Table 2. With respect to the dividend payout variable, the computational procedure measures the ratio of dividends paid by a firm in a test year divided by its reported earnings for that year. Thus a negative value for this variable is possible. As a result, this variable does display a very high level of dispersion. The coefficient of variation of DIVPAY can be computed to be 2.5584, while the same for the RISK variable is 1.6785, the two largest values of the variables utilized in this study. In the parts of this investigation where predictive powers of the OLS and NNW are compared, the sample data is randomly divided into two sets: one training set of 130 firms and a test data set of 21 firms to serve as a holdout sample.

First using the entire sample of 151 firms, the OLS methodology is applied to determine the best model as suggested by the adjusted R^2 criterion. Four sets of regression coefficients are estimated using all ten independent variables with four different definitions of debt utilization as dependent variables. The findings are displayed in Table 3. The results indicate that the

TABLE 2
Descriptive Statistics of the Variables (1973-1990)

VARIABLE	MEAN	VARIANCE	KURTOSIS	SKEWNESS	MINIMUM	MAXIMUM
1. SIZE	2.547	.251	-.028	.139	1.21	3.87
2. NTD	.036	.001	.931	.835	.01	.08
3. TAR	.719	.025	-.228	.245	.37	1.21
4. GROW	.094	.002	3.757	1.586	.01	.28
5. TECH	.023	.001	1.595	1.268	.00	.09
6. RISK	.930	2.436	7.557	2.693	.00	8.79
7. PROF	.105	.001	.857	.780	.02	.24
8. CSFLOW	.093	.001	-.638	.007	.02	.15
9. EARYLD	.096	.001	.054	.513	.02	.22
10. DIVPAY	.588	2.315	109.025	9.963	-.57	17.68
11. DEBTMV	.204	.012	-.293	.315	.02	.55
12. DEBTBV	.246	.010	.874	.590	.06	.59
13. LTDBV	.208	.010	1.205	.688	.00	.61
14. LTDMV	.175	.011	-.359	.388	.00	.49

TABLE 3
The Results of OLS Models on the Entire Sample with All variables (1973-1990) [1,2]

VARIABLE	DEBTMV	DEBTBV	LTDMV	LTDBV
SIZE	.020123	.0608	.0147	.0475
	[1.224]	[3.29]***	[.914]	[2.56]*
NDT	-1.6321	-.7644	-1.517	-.6316
	[-2.47]**	[-1.03]	[-2.34]**	[-.849]
TAR	.158651	.1567	.1779	.1961
	[2.845]***	[2.50]***	[3.26]***	[3.12]***
GROW	-.18092	.0089	-.1275	.0803
	[-1.26]	[.055]	[-.911]	[.498]
TECH	-1.4846	-1.597	-1.554	-1.858
	[-4.157]***	[-3.98]***	[-4.44]***	[-4.62]***
RISK	-.00034	-.0078	.00083	-.0055
	[-.067]	[-1.37]	[.168]	[-.970]
PROF	-1.5644	-1.016	-1.426	-1.0305
	[-9.561]***	[-5.53]***	[-8.90]***	[-5.59]***
CSFLOW	-.1869	-.9130	-.1441	-.7946
	[-.609]	[-2.64]***	[-.480]	[-2.30]**
EARYLD	.1285	.0485	.1318	.0717
	[.721]	[.242]	[.755]	[.357]
DIVPAY	-.0054	-.0016	-.0057	-.0023
	[-1.37]	[-.367]	[-1.50]	[-.527]
CONSTANT	.3244	.2433	.2675	.1910
	[5.93]***	[3.95]***	[4.99]***	[3.10]***
ADJ R^2	.5715	.3465	.5632	.3844
SSE	.6998	.8834	.6699	.8867

1.Tabulated values are the estimated coefficients, and the [t-values]
2.*, **, *** , represent 10%, 5% and 1% significance levels, respectively

model which relies on the market value based measure of debt ratio (DEBTMV) as the dependent variable yields the largest adjusted R^2 (0.5715).

The first column of Table 3 shows that the estimated coefficients of only four independent variables, namely, NDT, TAR, TECH, and PROF are statistically significant. Moreover, the estimated coefficients of TAR, TECH, and PROF are found to be significant in all four variations of the linear model. This fact prompted the present study to pursue the investigations of more parsimonious specifications by including only these independent variables which had statistically significant coefficients. The results emerging from such reduced specifications (four independent variables for market value based dependent variables and five independent variables for the book value based dependent variables) are summarized in Table 4. Again, the model which utilizes the market value based debt ratio DEBTMV, as the dependent variables exhibits the largest adjusted R^2 value (0.5704). This value of adjusted R^2, 0.5704, in the reduced model is very close to the 0.5715 value obtained for the 10-variable model. The results suggest that as far as linear OLS models which utilize DEBTMV as the dependent variable and either the full set of 10 or a reduced set of 4 independent variable display the best explanatory powers. These regression results of the present study are comparable with the results of numerous other empirical studies which have attempted to explain the determinants of debt ratios, although the relative sig-

TABLE 4
The Results of the Reduced [1] OLS Models on the Entire Sample (1973-1990)[2,3]

151 firms and "reduced" models				
VARIABLE	DEBTMV	DEBTBV	LTDMV	LTDBV
SIZE		.0438 [3.00]***		.0341 [2.34]**
NDT	-1.692 [-2.82]***		-1.510 [-2.59]***	
TAR	.1960 [3.83]***	.1216 [2.28]**	.2073 [4.16]***	.1627 [3.06]***
TECH	-1.380 [-4.19]***	-1.736 [-4.71]***	-1.479 [1.23]	-1.965 [-5.35]***
PROF	-1.624 [-10.5]***	-1.038 [-5.96]***	-1.470 [-9.82]***	-1.022 [-5.88]***
CSFLOW		-1.023 [-3.39]***		-.8698 [-2.89]***
CONSTANT	.3276 [8.79]***	.2958 [6.35]***	.2699 [7.43]***	.2417 [5.21]***
ADJ R^2	.5704	.3511	.5651	.3943
SSE	.7316	.9085	.6954	.9036

1.The Models in this table utilize the only input variables which were found significant in Table 3
2.Tabulated values are the estimated coefficients, and the [t-values]
3. *, **, *** represent 10%, 5%, 1% significance levels, respectively

nificance of variables differs.

Next, the neural network models for both, the 4-variable specification, as well as the 10-variable specification are calibrated for the entire sample. The sum of the squared errors (SSEs) for different number of cycles of training of NNW model is shown in Table 5. For the 4-independent variable specification (with DEBTMV as the dependent variable), the SSE value is minimized with 15,000 training cycles. For the 10-independent variable model, the training was terminated after 45,000 cycles since no further improvement in terms of reducing the SSE could be obtained. It is definitely worth noting the improvement accomplished by the NNW models as compared with the comparable linear models. Specifically, for a 4-variable reduced model, the SSE under the NNW framework is 0.6660 (Table 5) as compared to the SSE figure of 0.7316 (Table 4). Similarly, the NNW resulting SSE for a 10-variable model can be found to be 0.4961 (Table 5) while the linear regression produced an SSE value of 0.6998 (Table 3). Thus, NNW modeling is definitely successful in reducing the SSE figures by 9 to 29 percent over the comparable linear regression framework.

Encouraged by the relative success of the NNW models in reducing the SSEs, this study set out to compare the predictive abilities of the two models, namely, the NNW and the OLS models. To attain this objective, the models were first estimated using a randomly selected sample of 130 firms out of the 151 firm total sample. The estimated models, and/or the resulting SSEs from them are shown in Table 6. With the OLS methodology (using the DEBTMV as the dependent variable), the adjusted R^2 values of 0.5724 and 0.5809 are obtained under the 10-variable and 4-variable specifications, respectively. More importantly, the SSE values obtained are 0.5665 and 0.5833, respectively. Once again, the SSE values obtained with NNW models shown in the lower panel of Table 6, for the two specifications (the full and the reduced) are lower than the OLS method. In case of the 10-variable model, opting to choose the OLS method would have resulted in virtually doubling the SSE (0.5565 instead of 0.2966). The relative gain from the reliance on an NNW was a modest 8 percent or so, for a 4-variable model. (SSE going from 0.5419 to 0.5833). Inspite of this modest gain, we have selected the reduced model for

TABLE 5
Neural Network Results as a Function of Training Cycles[1]

Entire Sample (151 firms) and Four Independent Variables				
No. of Cycles	5000	10000	15000	25000
SSE	.6958	.672	.666	.6740
Entire sample (151 firms) Ten Independent variables				
No. of Cycles	15000	25000	45000	
SSE	.5326	.51584	.4961	

1. The dependant variable is DEBTMV

the subsequent parts of this study. Basically, this investigation was attempting ascertain the predictive power advantage of an NNW, if any, over the OLS model, even when the models used for prediction (Table 6) had not suggested a whopping advantage to the NNW model.

One of the crucial objectives of this study, namely, that of critically comparing the predictive powers of the two models can now be pursued. The procedure adopted for meeting the task at hand is as following. Using the estimated coefficients (or the neural weights) from the 130 firm sample of Table 6 are now used to predict the mean "market value based debt ratio" (DEBTMV) of the 21 firm holdout sample. The estimated mean debt ratio of the holdout sample is compared to the actual mean debt ratio of the holdout sample. This provides a barometer for evaluating the predictive power of each method.

Table 7 shows the SSE values obtained for the holdout sample of 21 firms with two methodologies using 4-input variables. Once again, the SSE with the NNW model can claim its superiority over the OLS generated SSE. Also, exhibited in this table are the Theil's U inequality coefficients. Theil's U provide an elegant way of systematically measuring the prediction accuracy of econometric models. The inequality coefficient, U, ranges between 0 and 1. The better the forecasting performance, the smaller will be the inequality coefficient with a value of 0.0 representing a perfect prediction. The values of U

TABLE 6
OLS and Neural Network Results for the Selected 130 firms [1, 2, 3]

VARIABLE	OLS Model 10 Variable Model	4- Variable Model
SIZE	.00701	
	[.411]	
NDT	-1.083	-1.1712
	[-1.40]	[-1.68]*
TAR	.1753	.1999
	[2.83]***	[3.60]***
	[.005]	[.000]
GROW	-.1212	
	[.153]	
TECH	-1.478	-1.4514
	[-3.49]***	[-3.84]***
RISK	.00052	
	[.097]	
PROF	-1.473	-1.5045
	[-8.83]***	[-9.67]***
CSFLOW	-.1472	
	[-.453]	
EARYLD	.1602	
	[.861]	
DIVPAY	-.0047	
	[-1.22]	
CONSTANT		.2996 .2931
	[5.14]***	[7.50]***
ADJ R^2	.5724	.5809
SSE	.5665	.5833
THE NEURAL NETWORK MODEL		
SSE	.2966	.5419

1. The dependant variable is DEBTMV
2. The tabulated values for the OLS methodology are the estimated coefficients and the [t-values]
3. *, and *** represent 10% and 1% levels of significance , respectively

TABLE 7
A Comparison of Predictive Capabilities of NNW and OLS Models
The Holdout Sample of 21 firms (1973-1990)[1, 2]

	Neural Network Model	*OLS Regression Model*
SSE	0.1461	0.1586
Theil's "U"	0.182	0.443

1. SSE are for the predicted debt ratios
2. Theil's "U" Coefficient measures the inequality between the predicted and the actual debt ratios

presented in Table 7 are 0.443 and 0.182 with the OLS model and the NNW model, respectively. These values therefore illustrate the commanding superiority of the NNW approach, in terms of its predictive ability.

Next, a simulation is performed, both, with the NNW model as well as the OLS regression model. The simulation consists of applying an increase and a decrease of 10 percent to each of the four independent variables (NDT, TAR, TECH, and PROF). With the 10 percent changes applied to the independent variables, the dependent variable, DEBTMV, is predicted for the holdout sample of 21 firms. The objective of the simulation exercise is to ascertain the sensitivity of the predicted debt ratios with respect to the changes in the independent variables. Moreover, the simulation would indicate if the relationships between the dependent and independent variables are non-linear. The findings of the simulation are summarized in Table 8. The tabulated results represent the p-values associated with the pairwise tests on the forecasts obtained with the two models (NNW and OLS). In all 3 tests, for both, the upward changes, as well as, the downward changes, only one independent variable (NDT) is found to be responsible for statistically significant differences. In other words, the variable which measures the non-debt tax shields gives rise to significant differences in the forecasted debt ratios with the two methods. One possible interpretation of this important finding is that the dependence of the corporate debt ratio on the non-debt tax shields may not be linear in nature. Therefore, the non-linear approach, namely, the NNW approach, is able to evaluate the influence of this crucial variable on the debt ratio better than the linear regression model. The only other variable which appears to bring about significant differences in predicted debt ratios of two models is the profitability measure (PROF). Yet, this variable has significant impact only when it is increased by 10 percent and not when decreased. This finding also suggests that the relationship between the debt ratio and the past profitability measure may not be linear (changing slopes).

SUMMARY

Ascertaining what determines the corporate debt ratio has long remained an enigmatic subject. For the most part, previous studies have relied on linear statistical methods to investigate the impact of a set of variables on the corporate use of debt. In this empirical investigation, a non- linear methodology, namely, the neural network method is applied to study the determinants of corporate borrowing. The documented applications of this novel approach range over a wide spectrum of subjects, including finance.

In this paper, the investigation centers on a sample of 151 non-financial forms over the 1973-1990 period. The findings of this study definitely show promise of the NNW method in the area of capital structure research. The results indicate that the SSEs (sum of squared errors) from the NNW model are consistently smaller than those obtained with the OLS regression models using the same set of variables. The reduction of SSEs found in this study range from a low of 9 percent to a high of 29 percent.

This paper also compares the predictive capabilities of an NNW model with the linear regression model. To accomplish this task, the paper uses the data of 130 randomly selected firms to predict the debt ratios of a holdout sample of 21 firms. The results show that NNW model indeed possesses superior predictive ability. The Theil's U inequality measures used in this study strongly suggest that the NNW method can be successfully adopted to offer better prediction accuracy.

Some of the glaring disadvantages of the standard OLS regression technique used in capital structure studies are the assumptions regarding the linear relationships, normality or log-normality, and the uncorrelated independent variables. The NNW approach does not rest on such assumptions. In this paper, a simulation is performed to evaluate the impact of four independent variables on the differences between the debt ratios predicted by the two methodologies. The results indicate that the non-debt tax

TABLE 8
Simulation Tests [1,2]

	t-test		Wilcoxon 2-tail significance level		Sign test	
Increase/ Decrease In Variable	+10%	-10%	+10%	-10%	+10%	-10%
NDT	.028**	.074*	.0325**	.0792*	.0784*	.1892
TAR	.681	.614	.0537	.7151	.6636	1.00
TECH	.125	.206	.1592	.357	.3833	.6636
PROF	.089*	.296	.0853*	.5663	.1892	1.00

1. The tests determine if the differences in the predicted values of debt ratio by the two models are significant when on e of the four of the independent variables is increased or decreased by 10%
2. * and ** represent the significance levels of 10% and 5%, respectively

shields variable produces consistently significant differences in the predicted debt ratios by the NNW and the OLS models. For the positive changes, the profitability variable also shows a similar impact. Thus, the findings suggest that the relationship between these variables and the debt ratios might be non-linear. It is believed that this paper paves the path for further research in the area of capital structure utilizing the neural network modeling technique.

NOTES

1 For exhaustive review of capital structure theories , see Harris and Raviv (1991)

2 To the best of authors' knowledge, none of the prior empirical studies on capital structure theory have tested predictive power of their models (usually regressions) or the stability of the independent variables. Almost all of the studies tested for the statistical significance of the estimated regression coefficients and explanatory power of the model. So, it is difficult to conclude whether the estimated coefficients have any predictive power.

3 For, example, the extent of insider ownership in a firm can lead to two conflicting hypotheses as to the level of debt. If the owner-managers are perceived as risk averse, then the insider ownership may be negatively related to level of debt. But on the contrary if owner managers are rational and follow modern portfolio theory they will diversify their holdings by resorting to high level of debt at firm level by reducing their equity stake in the firm and at the same time retain effective control of the firm. Under signaling hypothesis firms with high insider ownership can support relatively higher levels of debt, for investors might perceive it as a positive signal. A high debt may also reduce agency costs (Jensen and Meckling). Another example might be, the size (defined by total assets) might be positively or negatively to leverage depending upon the assumptions. Large firms which are well diversified and have lower bankruptcy cost allow them to borrow more funds and thus leverage might be positively related to size. On the contrary if one assumes that large firms are usually mature firms with limited growth opportunities might have large free cash flows and might be able to finance most of their needs from internally generated cash flows and hence, the leverage can be negatively related to size. Similarly, a high dividend payout ratio can be interpreted as either positively or negatively related to levels of debt depending upon the behavioral assumptions about managers. Another example is that a high level of profitability can support either high or low level of debt, depending upon the level of "free cash flows". Firms with high cash flows will use all of its internally generated resources before resorting to external debt. On the contrary, high cash flows can be interpreted as a growth firm and finance high growth with additional debt.

4 Hsieh (1993) in fact suggests that artificial neural systems could be used successfully in a wide range of financial management issues. One of the potential areas mentioned in this paper is the determination of capital structure.

5 The basic description of the problem is adopted from Yaron Danon, the creator of WinNN canned back-propagation neural network software.

6 The three most commonly used transfer functions are : (1) Linear, f(x) = xT, (2) Sigmoid f(x) = 1/(1+e-ax), and hyperbolic Tan: f(x)= tanh (xT). The choice of function, to certain extent, is arbitrary.

7 Authors thank Arjun Chatrath for generously allowing us to use the data in the present study.

8 See Harris and Raviv (1991) for an excellent and an exhaustive review of capital structure studies and the variables commonly used as surrogates.

REFERENCES:

Barnea, A., Haugen, R., and Sunbet, L.,1985, *Agency Problems and Financial Contracting*, Prentice Hall, Englewood Cliffs., NJ.

Baskin J., 1989, An empirical investigation of the pecking order hypothesis, *Financial Management* 18, 26-35.

Bradley, M., Jarrell, G. and Kim, H., 1984, On the existence of an optimal capital structure: theory and evidence., *Journal of Finance*., 39, 857-878.

Brennan, M., and Schwartz, E., 1978, Corporate income taxes, valuation and problems of optimal capital structure, *Journal of Business* , 39, 103-114.

Brennan, M., and Kraus, A., 1987 Efficient financing under asymmetric information, *Journal of Finance*, 42, Pages 1225-1243.

Castanias, R., 1983, Bankruptcy risk and optimal capital structure, *Journal of Finance*, Volume 38, pages 1617-1635.

DeAngelo, H., and Masulis, R., 1980, Optimal capital structure under corporate and personal taxation, *Journal of Financial Economics*, 8, 3-29.

Dudda, S., and Shekhar, S., 1988, Bond rating: A non-conservative application of neural networks, Proceedings *of IEEE international Conference on neural Networks*, 2, 443-450.

Fletcher, D., and Goss, E., 1993, Forecasting with neural networks: An application using bankruptcy data, *Information and Management*, 24, 159-167.

Friend, I., and Hashbrouck, J., 1988, Determinants of capital structure, in *Research in Finance,* Andy Chen, editor , 7, 1-19.

Harris, M., and Raviv, A., 1990, Capital structure and the informational role of debt, *Journal of Finance*, 45, 321-349.

Hamada, R.., and Scholes, M., 1985 Taxes and corporate financial management, in, *Recent Advances In Corporate Finance*, Ed Altman, editor, Homewood, IL .

Hammerstrom, Dan., 1993, Working with neural networks, *IEEE Spectrum*, 46-53.

Hecht, Nielsen., 1989, *Neuro-Computing*, Redding, MA: Addison-Wesley.

Hsieh, C., 1993, Some potential applications of artificial neural systems in financial management, *Journal of Systems Management*, 12-15

Hutchison, J., Lo, A., and Poggio, T., 1994, Non-parametric approach to pricing and hedging derivative securities via learning networks, *Journal of Finance*, 49, 851-889.

Jain, V. and Nag, B., 1995, Artificial neural network models for pricing initial public offering, *Decision Sciences*, 26, 283-302.

Jensen, M., and Meckling W., 1976, Theory of the Firm: Managerial behavior, agency costs and ownership structure, *Journal of Financial Economics*, 3, 305-360.

Kamath, R. 1997, Long-term financing decisions: Views and practices of financial managers of NYSE firms, *Financial Review*, 32, 331-356.

Kim, W., and Sorensen, E., Evidence on the impact of the agency costs of debt on Corporate debt policy, *Journal of Financial and Quantitative Analysis*, 21, 131-144.

Klimasauskas, C., 1988, *An Introduction to Neural Computing*, Sewickley, PA: NeuralWare, Inc.

Leyland, H., and Pyle, D., 1977, Information asymmetric, financial structure, and financial intermediation, *Journal of Finance*, 32, 371-388.

Masulis, R., 1980, The effects of capital structure change on security prices: A study of exchange offers, *Journal of Financial Economics*, 8, 139-178.

Miller, M., 1977, Debt and Taxes, *Journal of Finance*, 32, 261-275.

Modigliani, F., and Miller, M., 1958, The cost of capital, corporation finance, and the theory of investment, *American Economic Review*, 48, 261-297.

Modigliani, F, and Miller, M., 1963, Corporate income taxes and cost of capital: a correction, *American Economic Review*, 53, 433-443.

Myers, S., 1977, Determinants of corporate borrowings, *Journal of Financial Economics*, 5, 147-176.

Myers S., 1984, The Capital Structure Puzzle, *Journal of Finance*, 39, 575-592.

Myers, S. and Majluf, N., 1984, Corporate financing and investment decisions, when firms have information that investors do not have *Journal of Financial Economics*, 13, 187-221.

Narayanan, M., 1987, Debt versus equity under asymmetric information, *Journal of Financial and Quantitative Analysis*, 23, 39-51.

Noe, T., 1988, Capital structure and signaling game equilibria, *Review of Financial Studies*, 1, 331-356.

Norton, E., 1989, Determinants of capital structure: A survey, in *Advances in financial planning and forecasting*, editor C. F. Lee, Greenwich, CT, Jai Press, 3.

Pinegar, J., and Wilbricht, L., 1989, What managers think of capital structure theory: A survey, *Financial Management*, 18, 82-91.

Pruitt, S., and Gitman, L., The intersection between the investment financing and dividend decisions of major U.S. firms, *Financial Review*, 26, 409-430

Rajan, R., and Zingales, L., 1995, What do you know about capital structure? Some evidence from international data, *Journal of Finance*, 50, 1421-1460.

Ross, S., 1977, The determination of financial structure: The incentive signaling approach,*Bell Journal of Economics*, 8, 23-40.

Salchenberger, L., Cinar, E., and Lash, N., 1992, Neural networks: A new tool for predicting thrift failures, *Decision Sciences*, 23, 899-916.

Tam, K., and Kiang, M., 1992, Managerial application of neural networks: The case of bank failure prediction, *Management Science*, 38, 926-947.

Tttman, S., and Wessels, R., 1988, The determinants of capital structure choice, *Journal of Finance*, 43, 1-19.

Warner, J., 1977, Bankruptcy, absolute priority, and pricing of risky debt claims, *Journal of Financial Economics*, 4, 239-276.

Warner, J., 1983, Bankruptcy costs and: Some evidence, *Journal of Finance*, 32, 337-347.