Neural Comput & Applic (2006) 15: 117–123
DOI 10.1007/s00521-005-0014-x

ORIGINAL ARTICLE

Hsiao-Tien Pao · Yao-Yu Chih

Comparison of TSCS regression and neural network models for panel data forecasting: debt policy

Received: 18 April 2004 / Accepted: 1 September 2005 / Published online: 19 October 2005

Abstract Empirical studies of variations in debt ratios across firms have analyzed important determinants of capital structure using statistical models. Researchers, however, rarely employ nonlinear models to examine the determinants and make little effort to identify a superior prediction model among competing ones. This paper reviews the time-series cross-sectional (TSCS) regression and the predictive abilities of neural network (NN) utilizing panel data concerning debt ratio of high-tech industries in Taiwan. We built models with these two methods using the same set of measurements as determinants of debt ratio and compared the forecasting performance of five models, namely, three TSCS regression models and two NN models. Models built with neural network obtained the lowest mean square error and mean absolute error. These results reveal that the relationships between debt ratio and determinants are nonlinear and that NNs are more competent in modeling and forecasting the test panel data. We conclude that NN models can be used to solve panel data analysis and forecasting problems.

Keywords Neural networks · TSCS regression · Forecasting · Capital structure · Panel data

1 Introduction

One of the most perplexing issues facing financial managers is constructing a firm's capital structure, which comprises both debt and equity financing. Does the firm use too little or too much debt? Should different firms have different capital structures, and, if so, what should account for these differences? Although, how to decide on an optimal capital structure is complicated and not well understood, modern capital structure theories have provided insights into the effects of debt versus equity financing.

Modern capital structure theory suggests that each firm has an optimal capital structure, one that maximizes its value and minimizes its overall cost of capital. However, research on capital structure theory also points out that there are many contradictory issues regarding capital structure decisions, and the theory alone cannot be used to specify a precise optimal structure for a firm. In reality, business managers make actual capital structure decisions according to reasoned judgment supported by quantitative analysis together with an awareness of the theoretical issues. One of the quantitative analysis methods that financial economists most commonly use when making capital structure decisions is regression analysis.

A basic premise of this study is that the variability in the debt ratio is a function of the firm's dynamic behavior through time. Therefore, we employ a time-series cross-sectional approach (TSCS) to analyze the panel data and examine the debt policy of firms. TSCS contains the necessary mechanism to deal with both the intertemporal dynamics and the individuality of the firms being investigated. The methodological improvements gained from pooling TSCS data are well documented, e.g., Judge et al. [1]; Dielman [2]; and Hsiao [3]. Unfortunately, although many of these procedures have quantitative bases, there are few studies that evaluate the model's ability to predict. In addition, comparisons between linear and nonlinear models for debt policy are rare. This paper addresses two key issues. First, it analyzes the important determinants of capital structure using linear and nonlinear models. Second, it tries to identify a superior prediction model among competing ones. So far, there is no comparative analysis on debt ratio prediction using panel data. The purpose of this paper is to fill this void in the capital structure empirical literature.

H.-T. Pao (✉) · Y.-Y. Chih
Department of Management Science,
National Chiao Tung University,
Hsinchu, Taiwan, ROC
E-mail: htpao@cc.nctu.edu.tw
Tel.: 886-3-5131578

The paper is organized as follows: Sect. 2 presents the data source and variable definitions. Section 3 discusses the methodologies employed. Section 4 details the evaluation methods used for comparing the forecasting techniques. Section 5 analyzes the empirical results. The final section contains our summary and conclusions.

2 Data source and variable definition

Data used in this study came from 207 firms engaged in the high-tech industry of Taiwan between 1998 and 2001. Data from the first three years (621 data) served as training data, while that of the last year (207 data) as testing data. We used training data for model construction and testing data for performance evaluation of the model built. Each observation contains one dependent variable and eight independent variables. The Taiwan Economic Journal (TEJ) compiled all variables.

When defining a firm's capital structure, researchers usually employ several macroeconomic determinants to form complex models. The dependent variable is the firm's debt ratio. To be more comprehensive, we utilize both the total-debt ratio (TDR) and the long-term market-debt ratio (LDR) as dependent variables. We define TDR as total liabilities divided by total liabilities plus net worth, and LDR as total liabilities minus current liabilities divided by total liabilities minus current liabilities plus equity market value.

We choose variables to explain capital structure differences by considering the three principal theoretical models of capital structure: the static trade-off (STM), the pecking-order hypothesis (POH), and the agency theoretic framework (ATF) [4]. In each model, the choice between debt and equity depends on both firm-specific and institutional factors. In the STM, capital structure moves towards a target that reflects tax rates, asset type, business risk, profitability, and bankruptcy code. In the ATF, potential conflicts of interest between inside and outside investors determine an optimal capital structure that trades off agency costs against other financing costs. The nature of the firm's assets and growth opportunities are major factors determining the importance of these agency costs. In the POH, financial market imperfection is central. Transaction costs and asymmetric information link the firm's ability in undertaking new investments to its internally generated funds. If the firm must rely on external funds, as in the model of Myers and Majluf [5], then it prefers debt to equity due to the lesser impact of information asymmetries.

Empirically, distinguishing between these hypotheses has proven difficult. In cross-sectional tests, analysts might regard variables that describe the POH as STM or ATF variables and vice versa. Moreover, in time-series tests, Shyam-Sunder and Myers [6] showed that many of the current empirical tests lack sufficient statistical power to distinguish between the models. As a result, recent empirical research has focused on explaining capital structure choice with a variety of variables that can be justified using any or all of the three models. In this study, we consider the following eight independent variables.

Tax rate (TAXR): after relaxing irrelevant assumptions, firms with high-tax liabilities expect to utilize greater amounts of debt to take advantage of the deductibility of interest expense [7]. Zimmerman's [8] ratio of taxes paid to pretax income as the tax rate proxy can account for this deductible.

Firm size (SIZE): numerous studies have argued that the debt policy of firms may be affected by size, suggesting a positive relationship between the firm's size and debt ratio. Following Titman and Wessels [9], we employ the natural log of sales as proxy for size.

Profitability (PR): profitability reflects earnings to finance investment. Myers [10] suggested that managers have a pecking order in which retained earnings represent the first choice, followed by debt financing, and finally equity. If this were true, it would imply a negative relationship between profitability and debt ratio. A number of prior studies define profitability as the ratio of operating income to total assets, which is the proxy employed in this model.

Growth opportunities (GR): Myers [11] noted that high market-to-book ratios indicate the presence of growth opportunities [12], which can be thought of as real options. Given the agency costs attached to these options, it is, however, relatively more difficult to borrow against them than against tangible fixed assets.

Business risk (BR): Kim and Soresen [13] inferred that more the business risk a firm takes, the lesser its ability to issue secured debt. For our business risk proxy, we estimate the probability of financial distress as the variability of the return on assets over the available time period. We calculate the return on assets as the earnings before interest and tax divided by total assets. Increased variability in the return on assets implies an increase in the short-term operational component of business risk.

Collateral value of assets (CVA): Myers and Majluf have argued that the composition of collateral value of the firm's assets influences its financing sources. The greater the collateral value of the firm's assets, the greater its ability to issue secured debt, and, therefore, the lesser the need to reveal information about future profits will be. Following Titman and Wessels [9] and Mehran [14], we use the ratio of inventory plus gross plant and equipment to total assets as proxy for the firm's asset structure.

Uniqueness (UNI): Titman and Wessels argued that the capital structure decision should take into account the risk of bankruptcy. The more unique the firm, the higher the risk, and the higher the cost of bankruptcy would be. Hence, a negative relationship is expected between the firm's degree of uniqueness and its debt ratio. Titman and Wessels's measure for uniqueness, defined as the ratio of advertising plus research and development expense to annual sales, is also included in this model.

Long-term investment (INV): general argument assumes that more long-term investment can lower the volatility of earnings, and most, but not all, of the previous empirical studies have found the expected negative relationship between volatility and leverage. In other words, more long-term investment ratio, measured as long-term investment over total assets, indicates greater ability to support more leverage.

In general, the following model formulas express the relationship: TDR = f (TAXR, SIZE, PR, GR, BR, CVA, UNI, INV) and LDR = f (TAXR, SIZE, PR, GR, BR, CVA, UNI, INV).

3 Methodology

The time-series cross-sectional regression and the neural network BP model are the necessary foundation for analyzing and predicting debt ratios. To summarize:

3.1 TSCS regression model building

The TSCS regression (TSCSREG) model analyzes a class of linear econometric models that commonly arise when analyzes time-series and cross-sectional data. The TSCSREG procedure deals with panel data sets that consist of time-series observations on each of several cross-sectional units.

We can view such models as two-way designs with covariates expressed as follows:

$$Y_{it} = \sum_{k=1}^{8} X_{itk}\beta_k + u_{it} \quad i = 1, \ldots, 207; \quad t = 1, 2, 3,$$

where Y is the TDR or LDR, X_1 is the tax rate (TAXR), X_2 is the firm size (SIZE), X_3 is the profitability (PR), X_4 represents the growth opportunities (GR), X_5 is the business risk (RISK), X_6 is the collateral value of assets (CVA), X_7 is the uniqueness (UNI), and X_8 is the long-term investment (INV). The total number of firms is 207 and length of the time-series for each firm is three.

The performance of any estimation procedure for the model regression parameters depends on the statistical characteristics of the error components in the model. The TSCSREG procedure estimates the regression parameters in the preceding model under three common error structures. The error structures are as follows.

(a) Variance components model (VC)

$$u_{\mathrm{it}} = v_{\mathrm{i}} + e_{\mathrm{t}} + \varepsilon_{\mathrm{it}}$$

To determine if there are heteroscedasticity problems, we used the Fuller-Battese [15] method to estimate this model. The error structure of this model is similar to the common two-way random effects model with covariates. It estimates the variance components by the fitting-of-constants method and estimates the regression parameters with generalized least squares (GLS).

(b) First-order autoregressive model (AR)

$$u_{it} = \rho_i u_{i,t-1} + \varepsilon_{it}$$

To determine if there are autocorrelation problems, we used the Parks [16] method to estimate this model. This model assumes a first-order autoregressive error structure with contemporaneous correlation between cross-sections. A two-stage procedure leading to the estimation of model regression parameters by GLS estimates the covariance matrix.

(c) Variance-component moving average model (VCMA)

$$u_{it} = \alpha_{\mathrm{i}} + b_t + e_{it}$$
$$e_{it} = \alpha_0\varepsilon_t + \alpha_1\varepsilon_{\mathrm{t}-1} + \cdots + \alpha_m\varepsilon_{t-m}$$

To determine if there are heteroscedasticity and autocorrelation problems, we used the Da Silva [17] method to estimate the mixed variance-component moving average model. This method estimates the regression parameters using a two-step GLS-type estimator.

3.2 Neural network model building

The BP neural network consists of an input layer, an output layer, and one or more intervening layers also referred to as hidden layers. The hidden layers can capture the nonlinear relationship between variables. Each layer consists of multiple neurons that are connected to neurons in adjacent layers. Since these networks contain many interacting nonlinear neurons in multiple layers, the networks can capture relatively complex phenomena [18, 19].

The historical data of a panel data set can train a neural network, in order to capture the characteristics of this data set. A process of minimizing the forecast errors will iteratively adjust the model parameters (connection weights and node biased). For each training procedure, we randomly selected an input vector from the training set and submitted it to the input layer of the network being trained [20]. We then propagated the output of each processing unit forward through each layer of the network.

As shown in Fig. 1, the NN model consists of an input layer with eight input nodes, one hidden layer comprising h nodes, and an output layer with a single output node. The input to the NN includes eight variables used in the TSCS model. During training, we presented a set of n pairs of input vectors and corresponding output, $((X(1),y(1)), (X(2),y(2)), \ldots, (X(n),y(n)))$ to the network, one pair at a time (Fig. 1).

We calculated a weighted sum of the inputs, $\mathrm{NET}_t = \sum_{i=1}^{N} w_{ti}x_i + b_t$, at tth hidden node. Each hidden

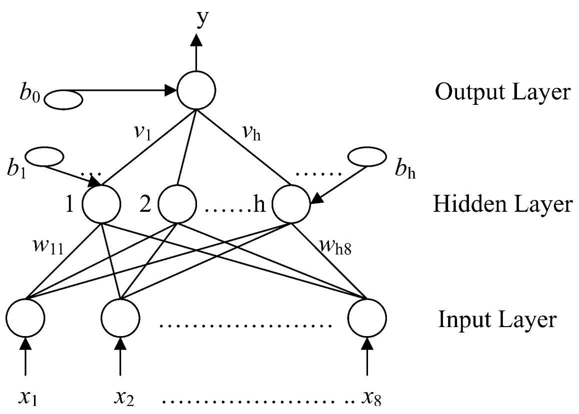

Fig. 1 Neural network model

node then uses a sigmoid transfer function to generate an output, $Z_t = [1 + \exp(-\mathrm{NET}_t)]^{-1} = f(\mathrm{NET}_t)$, between 0 and 1. We then sent the outputs from each of the hidden nodes, along with the bias b_0 on the output node, to the output node and again calculated a weighted sum $\mathrm{NET} = \sum_{t=1}^{h} v_t Z_t + b_0$. The weighted sum becomes the input to the sigmoid transfer function of the output node. We then scaled the resulting output, $\hat{Y} = f(\mathrm{NET}) = [1 + \exp(-\mathrm{NET})]^{-1}$, to provide the predicted output value. At this point, the second phase of the BP algorithm, adjustment of the connection weights, begins. The parameters of the neural network can be determined by minimizing the following objective function of SSE in the training process: $\mathrm{SSE} = \sum_{j=1}^{n} \left(y_j - \hat{Y}_j\right)^2$ where $\hat{Y}_j$is the output of the network for jth observation.

Assume the relationship of Y and X is monotone, then calculate the sensitivity S_i of the outputs to each of the ith inputs as a partial derivative of the output with respect to the input [21].

$$\mathrm{S_i} = \frac{\partial \hat{\mathrm{Y}}}{\partial \mathrm{X_i}} = \sum_{\mathrm{t}=1}^{h} \frac{\partial \hat{\mathrm{Y}}}{\partial \mathrm{NET}} \frac{\partial \mathrm{NET}}{\partial Z_t} \frac{\partial Z_t}{\partial \mathrm{NET_t}} \frac{\partial \mathrm{NET}_t}{\partial X_i}$$
$$= \sum_{t=1}^{h} [f'(\mathrm{NET}) v_t f'(\mathrm{NET_t}) w_{ti}]$$

Assume $f'(\mathrm{NET})$ and $f'(\mathrm{NET_t})$ are constants and we ignore them. Then the relative sensitivity is $\hat{S}_i = \sum_{t=1}^{h} v_t w_{ti}$. The independent variable with higher relative positive (negative) sensitivity has the bigger positive (negative) impact on the dependent variable.

Looking at the degree to which the NN output matches the actual value for the corresponding input values measures performance. In this study, we varied the number of hidden nodes for the neural network from one to ten. It is noteworthy that the resulting neural network models performed relatively better with four to six hidden nodes. However, the predictive accuracy of the model improved with the in-sample data set and declined with the out-of-sample data set when using more than six hidden nodes. Hence, we used five hidden nodes in the resulting NN. In general, the need for less hidden nodes indicates little interaction of the inputs, and a diminished ability for the neural networks to outperform other statistical models. A small number of hidden nodes provides assurance of the robustness of the NN.

While NNs have some limitations, several researchers have demonstrated that NNs are excellent at developing overall models. A wide variety of applications have documented the accuracy of NN in predicting outcomes. This study is the first attempt to examine the usefulness of NNs in predicting capital structure and to compare these NNs with TSCSREG models.

4 Forecasting evaluation methods

Yokum and Armstrong [22] conducted an expert opinion survey to select evaluation criteria for forecasting techniques. Accuracy was the most important criterion, followed by the savings in cost as a result of the improved decisions. In particular, execution issues, such as ease of interpretation and ease of use, were also highly rated. In this study, we used three criteria to evaluate forecasting models.

The first measurement criterion is root mean square error (RMSE):

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{n} (P_i - A_i)^2 \Big/ n}$$

The second measurement criterion is mean absolute error (MAE):

$$\mathrm{MAE} = \sum_{i=1}^{n} |P_i - A_i| \Big/ n$$

The third criterion is absolute percentage error for ith observation, APE_i:

$$\mathrm{APE_i} = |(P_i - A_i)/A_i| * 100,$$

where P_i and A_i are the ith predicted and actual value, respectively, and n is the total number of predictions. The absolute percentage error (APE_i) classifies firms into three categories: (1) those with an APE_i of less than 5%; (2) those with an APE_i between 5–15%; and (3) those with an APE_i greater than 15%. We chose these forecasting errors based on the understanding that 5% is acceptable to most financial managers, 5 to 15% is fuzzy area and is a somewhat unreliable indicator, while more than 15% is unacceptable. It follows that the superior model is the model with the

Table 1 TSCSREG results for debt ratios

Variables	Total debt-ratio			Long-term marker-debt ratio		
	VC	AR	VCMA	VC	AR	VCMA
	Coefficients (*t*-statistics)			Coefficients (*t*-statistics)		
Intercept	0.1014 (1.622)	0.1815 (2.124)**	0.1801 (2.100)**	–0.1749 (–1.903)**	–0.0102 (–0.255)	
TAXR	0.0460 (1.435)	0.0599 (1.801)*	0.0486 (1.613)	–0.0185 (–1.028)	0.0780 (1.791)*	0.0841 (2.315)**
SIZE	0.0237 (6.144)***	0.0188 (5.304)***	0.0201 (5.627)	0.0109 (2.335)**	0.0139 (2.940)***	0.0142 (3.177)***
PR	–0.3581 (–7.254)***	–0.4118 (–9.012)***	–0.3012 (–6.501)***	–0.1558 (–4.616)***	–0.1902 (–5.301)***	–0.1591 (–4.753)***
GR	–0.0042 (–1.030)	–0.0061 (–1.962)**	–0.0061 (–1.971)**	–0.0035 (–1.753)*	–0.0043 (–2.311)**	–0.0049 (–3.012)***
BR	–0.1049 (–0.901)	–0.3019 (–2.101)**	–0.2017 (–1.607)	0.1044 (1.332)	0.0541 (0.2763)	–0.2908 (–1.709)*
CVA	0.1107 (3.490)***	0.1016 (3.325)***	0.0448 (2.322)**	0.1055 (2.226)**	0.1301 (2.544)***	0.1282 (2.391)***
UNI	–0.2536 (–1.762)*	–0.3126 (–2.102)**	–0.3001 (–1.984)**	–0.1006 (–1.352)	–0.1714 (–2.201)**	–0.1801 (–2.254)**
INV	–0.2039 (–4.126)***	–0.2201 (–4.514)***	–0.1768 (–3.829)***	–0.1268 (–2.962)***	–0.1135 (–2.780)***	–0.1146 (–2.833)***
RMSE	0.1170	0.0992	0.1028	0.0960	0.0833	0.0912

*, **, and *** significant at the 10, 5, and 1% level

higher percentage of most accurately predicted properties.

5 Experimental results

Table 1 shows the TSCSREG results on TDR and LDR. Results indicated that the tax rate and business risk were not significant on TDR and LDR, respectively. Table 2 lists the relative sensitivities of each independent variable to debt ratio. Comparing the results in both tables shows that the signs of each independent variable in Table 1 are the same as those in Table 2. In addition, the unimportant determinants in TSCS models also have low-relative sensitivities in NN models. Furthermore, the NN models show smaller RMSE. These results reveal that debt ratio does have a monotone relationship with the eight independent variables, but their relationships are not linear.

Tables 3 and 4 summarize the comparative forecasting performance of TSCS and NN models using 207 testing data. Table 3 demonstrates the outcome when TDR is the dependent variable and Table 4 exhibits the outcome when LDR is the dependent variable. As seen in Table 3, TSCS models perform best under first-order autoregressive (AR) error structure when using all evaluation criteria, i.e., generally RMSE and MAE are the lowest, APE 5% is the highest, and APE over 15% is the lowest for this model. The variance components (VC) error structure has the poorest performance among TSCS models. In addition, NN models perform better, especially for NN with one-year lag, than all of the TSCS models when using all evaluation criteria. Table 4 also reveals the same conclusions. The TSCS and NN models generally demonstrate that there is a close correlation between the firm's current debt ratio with the debt ratio in the previous period.

We sort 207 testing data by debt ratio in descending order. By using TSCS and NN models, we can forecast the first 52, the central 103, and the last 52 debt ratios of firms. According to Tables 3 and 4, the NN models have the lowest RMSE among all of TSCS models when firms have high- or low-debt ratios. In general, the NNs performance is somewhat constant as the debt ratio varies, while the TSCSs performance deteriorates significantly when the debt ratio increases or decreases.

6 Conclusion

Empirical studies on capital structure have examined the determinant of debt ratio using statistical models, but

Table 2 The NN results for debt ratio

	Total debt-ratio		Long-term marker-debt ratio	
	NN	NN with 1-year lag	NN	NN with 1-year lag
Variables			Sensitivities	
TAXR	0.118	0.106	0.095	0.112
SIZE	0.688	0.652	0.582	0.602
PR	–0.970	–1.012	–1.139	–0.869
GR	–0.332	–0.262	–0.301	–0.401
BR	0.052	–0.133	0.200	–0.093
CVA	0.501	0.426	0.336	0.522
UNI	–0.527	–0.325	–0.432	–0.361
INV	–0.714	–0.601	–0.544	–0.461
LAGDEBT		3.301		5.417
RMSE	0.0778	0.0623	0.0861	0.0740

Table 3 Comparative forecasting performance of total-debt ratio

	Size	TSCS-VC	TSCS-AR	TSCS-VCMA	NN	NN with 1-year lag
RMSE of training data	621	0.1170	0.0992	0.1028	0.0778	0.0623
RMSE of testing data	207	0.1328	0.1121	0.1160	0.0849	0.0706
RMSE of high debt ratio	52	0.1801	0.1629	0.1662	0.1210	0.1004
RMSE of moderate debt ratio	103	0.0775	0.0720	0.0698	0.0691	0.0652
RMSE of low debt ratio	52	0.1672	0.1206	0.1334	0.0903	0.0871
MAE of testing data	207	0.1418	0.1301	0.1288	0.0914	0.0769
APE 5% of testing data (%)		9.61	11.32	12.24	22.33	36.24
APE 5% $\sim$ 15% of testing data (%)		19.53	33.78	30.43	30.07	58.23
APE Over 15% of testing data (%)		70.86	54.90	57.33	47.60	5.53
Superior model	NN with 1-year lag					

Table 4 Comparative forecasting performance of long-term market-debt ratio

	Size	TSCS-VC	TSCS-AR	TSCS-VCMA	NN	NN with 1-year lag
RMSE of training data	621	0.0960	0.0833	0.0912	0.0861	0.0740
RMSE of testing data	207	0.1130	0.0964	0.1052	0.0902	0.0827
RMSE of high debt ratio	52	0.1501	0.1322	0.1302	0.1215	0.1202
RMSE of moderate debt ratio	103	0.0846	0.0724	0.0795	0.0822	0.0714
RMSE of low debt ratio	52	0.1972	0.2011	0.2052	0.1120	0.1103
MAE of testing data	207	0.1248	0.1009	0.1097	0.0962	0.0836
APE 5% of testing data (%)		11.05	14.21	10.20	14.00	28.85
APE 5% $\sim$ 15% of testing data (%)		26.52	35.73	36.08	36.12	50.12
APE Over 15% of testing data (%)		62.43	50.06	53.72	49.88	21.03
Superior model	NN with 1-year lag					

have made little effort to compare linear and nonlinear models or to examine the prediction capabilities of these models. This paper addresses the following two issues. First, the signs of each determinant in TSCS linear regression models are the same as those of the relative sensitivities of these determinants in neural network nonlinear models. In addition, the insignificant determinants in TSCS models have low relative sensitivities in NN models. It seems that these two models show consistent results for capital structure determinants. Researchers and practitioners can employ either neural networks or traditional statistical models to analyze the important determinants of capital structure.

Second, we conducted a comparative analysis to determine a superior prediction model for panel data. We found that neural network models with one-year lag do have the competitive ability for panel data modeling and forecasting. In addition, NN models perform much better than TSCSREG models when firms have either high- or low-debt ratios. One possible reason that neural network models with one-year lag can outperform three TSCSREG models is that the two panel data sets of debt ratios we investigated show nonlinear relationships with the eight determinants selected, although their relationships are monotone. We concluded that neural network models can solve panel data analysis and forecast problems.

References

1. Judge G, Griffiths W, Hill R, Lee T (1980) The theory and practice of econometrics. John Willey and Sons, New York
2. Dielman T (1983) Pooled cross-sectional and time-series data: a survey of current statistical methodology. The Am Statistician 37:111–122
3. Hsiao C (1986) Analysis of panel data. Cambridge University Press, Cambridge
4. Booth L, Aivazian V, Demirguc-Kunt A, Maksimovic V (2001) Capital structure in developing countries. J Financ 56(1):87–130
5. Myers SC, Majluf NS (1984) Corporate finance and investment decisions when firms have information that investors do not have. J Financ Econ 13:187–221
6. Shyam-Sunder L, Myers S (1999) Testing static tradeoff against pecking order models of capital structure. J Financ Econ 51:219–244
7. Haugen R, Senbet L (1986) Corporate finance and taxes: a review. Financial Manage. 15:5–22
8. Zimmerman J (1983) Taxes and firm size. J Account Econom 5:119–149
9. Titman S, Wessels R (1988) The determinants of capital structure choice. J Financ 43:1–20
10. Myers SC (1984) The capital structure puzzle. J Financ 39:575–592
11. Myers SC (1977) Determinants of corporate borrowing. J Financ Econ 5:147–175
12. Miller Merton, Modigliani F (1961) Dividend policy, growth and the valuation of shares. J Bus 34:411–433
13. Kim WS, Sorensen EH (1986) Evidence on the impact of the agency cost of debt on corporate debt policy. J Financ Quant Anal 21:131–144
14. Mehran H (1992) Executive incentive plans, corporate control, and capital structure. J Financ Quant Anal 27:539–560
15. Fuller WA, Battese GE (1974) Estimation of linear models with crossed-error structure. J Econometrics 2:67–78
16. Parks RW (1967) Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated. J Am Stat Assoc 62:500–509
17. Da Silva JGC (1975) The analysis of cross-sectional time series data, PhD dissertation, Department of Statistics, North Carolina State University

18. Hill T, O'Connor WRM (1996) Neural network models for time series forecasts Manage Sci 42(7):1082–1092
19. Chiang WC, Urban TL, Baldridge GW (1996) A neural network approach to mutual fund net asset value forecasting, Omega. Int J Manage Sci 24(2):205–215
20. Liu MC, Kuo W, Sastri T (1995) An exploratory study of a neural network approach for reliability data analysis, Qual Reliab Eng Int 11:107–112
21. Hwang JN, Choi JJ, Oh S, Marks RJ (1991) Query based learning applied to partially trained multilayer perceptron. IEEE T-NN 2(1):131–136
22. Yokum JT, Armstrong JS (1995) Beyond accuracy: comparison of criteria used to select forecasting methods. Int J Forecast 11(4):591–597