1-1-1990

Neural nets versus logistic regression: A comparison of each model's ability to predict commercial bank failures

Timothy B. Bell

Gary S. Ribar

Jennifer Verichio

Recommended Citation

Auditing Symposium X: Proceedings of the 1990 Deloitte & Touche/University of Kansas Symposium on Auditing Problems, pp. 029-053;

3

Neural Nets Versus Logistic Regression: A Comparison of Each Model's Ability to Predict Commercial Bank Failures

Timothy B. Bell
Gary S. Ribar
Jennifer Verchio
KPMG Peat Marwick

Introduction

According to SAS No. 59, *The Auditor's Consideration of an Entity's Ability to Continue as a Going Concern* [AICPA, 1988], the auditor has a responsibility to evaluate whether there is substantial doubt about the client's ability to continue as a going concern for a reasonable period of time, not to exceed one year beyond the date of the financial statements being audited. Once this evaluation is complete, if the auditor concludes there is substantial doubt, he is required to add an explanatory paragraph to the audit report reflecting his conclusion. The going concern evaluation is particularly troublesome for commercial bank clients operating in a regulated environment. For these institutions, federal and state regulators ultimately decide whether and when a particular bank will be closed, and the auditor faces the additional challenge of predicting whether regulators will take such actions within 12 months of the date of the financial statements.

This study examines the usefulness of annual financial statement data and alternative modeling methodologies for modeling regulators' decisions to close commercial banks. A bank failure prediction model could be applied at the audit planning stage (using annualized third quarter data) to aid resource allocation decisions. The model could also be applied at the review stage of the audit (using annual post-adjustment data) as an aid to the final opinion reporting decision.

We focus on two different methodologies - logistic regression and neural network computing - and compare their abilities to predict commercial bank failures over a 12-month horizon. Our preliminary results indicate that both methodologies yield similar predictive accuracy across the range of all possible model cutoff values, with the neural network performing marginally better in the "gray area" where some failing banks appear to be less financially distressed.

The remainder of the paper contains sections covering sampling methodology, selection of candidate predictor variables, modeling methodology, estimation of model fit, and prediction results. The paper concludes with a summary of our research findings.

Sample Selection Process

During the period from 1983 through 1988, there has been a dramatic increase in the number of federally insured commercial banks requiring disbursements by the Federal Deposit Insurance Corporation (FDIC). Sheshunoff & Co. of Austin, Texas reported 45 such bank failures during 1983; 79 during 1984; 117 during 1985; 138 during 1986; 164 during 1987; and 179 during 1988. These failures included institutions entering receivership, institutions that had their deposits assumed by others, and institutions merged into others under Federal assistance plans. For this study, we used an estimation sample comprised of 102 of the 117 banks that failed during 1985 (1984 annual financial statement data) and a separate holdout sample containing 131 of the 138 banks that failed during 1986 (1985 annual financial statement data). Failed banks from the 1985 and 1986 Sheshunoff lists that were not included in either sample had been closed by the regulators during the first month of each year, and as a result no prior year's financial statement data were available.

A stratified sampling design was applied to identify nonfailed banks for inclusion in both samples. Nonfailed banks were drawn from the nine different peer groups listed in TABLE 1. These peer groups are based on differing ranges of total assets. The nine strata for the holdout and estimation samples of nonfailed banks are approximately proportional to the population strata proportions as shown in TABLE 1. This stratification design was undertaken to test the general applicability of estimated models to banks of all different sizes. As shown in TABLE 1, 906 nonfailed banks were included in the 1984 estimation sample and 928 nonfailed banks were included in the 1985 holdout sample.

Selection of Candidate Predictor Variables

Candidate predictor variables were identified using the results of prior research, and researcher intuition. Altman, Avery, Eisenbeis and Sinkey [1981] summarize several prior bank failure prediction studies including studies sponsored by the FDIC, Federal Reserve Board of New York, Office of the Comptroller of the Currency (OCC), Board of Governors of the Federal Reserve System, and other studies. Our set of candidate predictor variables includes the most efficacious of the predictors tested in these studies.

During 1988, the OCC published a document entitled *Bank Failure - An Evaluation of the Factors Contributing to the Failure of National Banks* [1988]. The document reports the results of an analysis of banks that failed, became problems and recovered, or remained healthy during the period 1979 through 1987. It identifies eight broad categories where weaknesses had a significant impact on bank declines. To the extent possible, we identified ratios that capture the essence of these categories for inclusion in our set of candidate predictor variables.

APPENDIX A presents a list of 28 candidate predictor variables we identified for possible inclusion in our models. Ratio numerators and denominators are comprised of line items taken from the annual call reports of commercial banks included in our samples. We used call report financial statements in lieu of GAAP financial statements based on the presumption that

TABLE 1
SAMPLE SIZES IN RELATION TO POPULATION OF ALL COMMERCIAL BANKS

Peer		Dec - 87		Nonfailed Samples				Failed Samples			
Group	Asset Size	Population	Ppns.	1984	Ppns.	1985	Ppns.	1984	Ppns.	1985	Ppns.
1	$5 Billion +	83	0.6%	5	0.6%	5	0.5%	0	0.0%	0	0.0%
2	$1 to $5 Billion	265	1.9%	11	1.2%	15	1.6%	0	0.0%	1	0.8%
3	$500 Million to $1 Billion	218	1.6%	15	1.7%	16	1.7%	0	0.0%	2	1.5%
4	$300 to $500 Million	307	2.3%	18	2.0%	18	1.9%	0	0.0%	0	0.0%
5	$100 to $300 Million	1,876	13.8%	112	12.4%	119	12.8%	2	2.0%	12	9.2%
6	$50 to $100 Million	2,852	20.9%	158	17.4%	174	18.8%	9	8.8%	18	13.7%
7	$25 to $50 Million	3,546	26.0%	266	29.4%	264	28.4%	26	25.5%	33	25.2%
8	$10 to $25 Million	3,457	25.4%	233	25.7%	237	25.5%	40	39.2%	43	32.8%
9	$0 to $10 Million	1,018	7.5%	88	9.7%	80	8.6%	25	24.5%	22	16.8%
	Totals	13,622	100.0%	906	100.0%	928	100.0%	102	100.0%	131	100.0%

regulators focus on the information filed in these Federal documents, and because the call reports have more detail than the GAAP reports. Moreover, GAAP reports are not available for a large number of separate institutions that fall under the umbrella of a single bank holding company.

The 28 candidate predictor variables relate to the following general financial features: bank size, loan exposure, capital adequacy, asset quality, operating performance, non-operating performance, and liquidity. Size is measured using the natural logarithm of total assets (LGASSETS). Loan exposure variables include the proportion of total assets represented by (1) construction loans (CONSTLNS), (2) real estate loans (RLESTLNS), and (3) agricultural loans (AGLNS). Based on the OCC's finding that insider abuse leads to bank failures, we included a fourth loan exposure variable designed to capture this condition - (4) aggregate credit to officers (loans to insiders) as a proportion of net loans (LNSINSID). Although many insider abuses go unrecorded, the OCC did observe that such abuse "include[s] situations where the transactions may be technically lawful [and therefore recorded in the financial statements] but exhibit bad judgment or self-interest above the interests of the bank" [OCC, 1988, p. 33]. The completeness assertion is one of the most troublesome aspects of an external audit, and to the extent that insider transactions go unrecorded, the ability to predict financial failure is most likely decreased.

Measures designed to capture the adequacy of bank capital include (1) primary capital to adjusted assets (PRMCAPAS), (2) total capital to total loans (TOCAPLNS), and (3) the raw measure of total equity capital (EQCAPTL). It is customary to add the allowance for loan losses to equity capital when measuring primary and total capital but we found that subtracting this amount yields stronger predictions. Therefore, our measures of primary and total capital are quite conservative. Capital has actually been reduced by twice the amount of the loan loss reserve - once by the bank's accrual of loan losses and again by our subtraction of the amount. So, these measures of capital assume that actual loan losses are understated.

Asset quality predictor variables include various measures of substandard loans as a proportion of either gross loans, primary capital, or total assets. The call report includes the following separate categories of substandard loans: (1) loans past due over 90 days (used in PDLNSGRL), (2) loans for which interest accrual has been suspended (used in NONACLNS), (3) total nonperforming loans, which is the sum of past due loans and nonaccrual loans (used in NPLNSPCP and NPLNSAST), and (4) loans that have been restructured (used in RESTRLNS). Two additional asset quality predictor variables are the ratio of net charge-offs to total loans (CHRGOFFS), and the ratio of provision for loan losses to total assets (PROVLOSS).

Measures designed to capture operating performance include (1) total interest income to total assets (YIELD), (2) total interest expense to total assets (RATE), (3) net interest income to total assets (SPREAD), (4) return on total assets (RETNTA), (5) return on total equity capital (RETEQ), (6) undivided profit and capital reserves to total assets (CUMPROF), and (7) income before extraordinary items (INCOME). Non-operating performance measures include (1) total non-interest income to total assets (NONINT), (2) total overhead expense to total assets (OVRHDEXP), and (3) security gains (losses) and gross extraordinary items to total assets (SECGAINS).

Liquidity measures include (1) short-term assets less large liabilities to total assets (LIQSTAST), (2) large time deposits to total assets (TMDEPS), and net loans to total assets (NETLNS). LIQSTAST measures the gap between short-term liquid assets and large deposits and provides an indication of the bank's ability to produce cash should depositors make large withdrawals. TMDEPS measures the proportion of total assets represented by these large deposits, and NETLNS represents the proportion of total assets that is nonliquid. For brevity, we will use acronyms to represent each of the 28 predictor ratios throughout the remainder of this study. The reader is referred to APPENDIX A for detailed definitions.

Modeling Methodologies

In our effort to estimate a model for predicting bank failures, we primarily focused on two modeling methodologies: logistic regression (or the logit model), and neural network computing. Brief descriptions of each methodology are given next.

Logistic Regression

During the recent past, binary logistic regression has been applied in a number of research studies that have attempted to model specific binary decisions or the binary representation of the occurrence of an event (e.g., vote yes/vote no and bankrupt/not bankrupt). In the current study, the logistic regression model can be interpreted as follows. Suppose there exists an unobservable theoretical index, Z_i, that represents the regulators' propensity to close commercial banks. Z_i is assumed to be a continuous random variable and is determined by a linear combination of observable bank characteristics, such as asset quality, loan exposure, capital adequacy, expected future financial performance, liquidity, etc. The logit model given below allows the estimation of the weights (coefficients) for the linear combination of bank attributes, and the resultant estimation of the index Z_i:

$$P_i = \frac{1}{1 + e^{-Z_i}} \qquad (1)$$

P_i represents the conditional probability that the regulator will close the bank, and e is the base of natural logarithms.

The likelihood function for use in sample estimation of the coefficients of Z_i is given by the product of all P_is for failed banks times the product of $(1 - P_i)$ for all nonfailed banks. So, higher failure probabilities for failed banks, and lower failure probabilities for nonfailed banks, represent higher points on the likelihood function. The coefficients comprising Z_i can be estimated by finding the global maximum of the likelihood function (i.e., differentiating and setting equal to zero). Due to the nonlinearity of the partial derivatives, however, an iterative technique such as the Newton-Raphson method must be used to determine this global maximum.

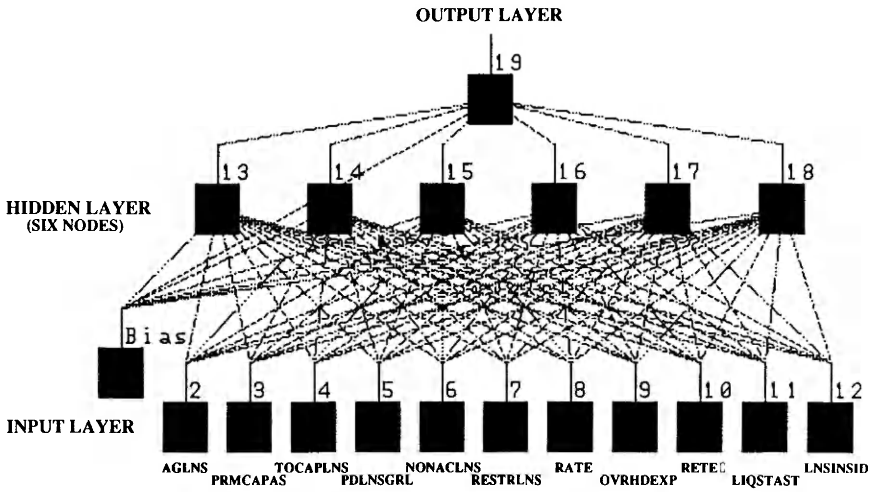
Neural Network Computing

Over the past few years a new methodology referred to as neural network computing, or connectionist modeling, has undergone rapid development. Neural nets have been applied to a variety of classification, clustering, and pattern recognition problems and in some cases have significantly outperformed standard statistical techniques such as the logit model.

Neural network architecture is biologically inspired, involving the intricate interconnection of many nodes (also referred to as processing elements) through which inputs are transformed into outputs. Once a particular network architecture is defined, the network is repeatedly presented with training cases from an estimation sample, and the connection weights between nodes are modified to bring the network outputs closer to the actual target output values. This training process is referred to as network learning. One of the potential advantages of neural network modeling is the ability to capture inherent process nonlinearities through the specification of an intricate network architecture. Interactions can also be modeled by specifying multiple connections to individual nodes.

The basic elements of a neural network are (1) nodes, (2) layers (of nodes), (3) connections (between nodes), and (4) connection weights. FIGURE 1 contains an illustration of the specific network architecture used in the current study. The first layer found at the bottom of the illustration is comprised of a bias node (similar to a constant in a regression model, and always

FIGURE 1: NEURAL NETWORK ARCHITECTURE FOR KNOWLEDGE ACQUISITION OF REGULATORS' DECISIONS TO CLOSE COMMERCIAL BANKS



Network Attributes:

1. Hyperbolic Tangent Transfer Function
2. Normalized Cumulative Backpropagation - Error Backprogated Using Overall Error Function Instead Of Each Individual Error Function
3. Training Set = 102 Banks Closed During 1985 (1984 Year-End Data) and 102 Nonfailed Banks
4. Approximately 300,000 Epoch Iterations Performed During Learning
5. Input Variables Chosen Based On Exploratory Data Analysis, Analysis Of All Possible Regressions, And Logistic Regression Results

given a value of 1), and one input node for each predictor variable. This layer serves as an input buffer where the input nodes simply pass the given predictor-variable values for the current training case (sample observation) up the connections toward the hidden (middle) layer. The input nodes are fully connected to the six nodes in the hidden layer. Each connection path has an associated weight (similar to a regression model coefficient) that is multiplied by the input value being passed through the connection.

Each node contained in the hidden layer receives a combined signal from each connection below it. This signal is simply the sum of the products of connection weights and input values. Note that each sum of products is analogous to Z_i in the logit model described above, but each of the six nodes in the hidden layer has a separate sum of products. Upon entering the six hidden layer nodes, these sums of products are individually transformed into output signals via application of a specified transfer function. Customarily, a sigmoidal growth function is used as the network transfer function for nodes in the hidden and output layers. Two commonly used transfer functions are the sigmoid (or logistic) function given in equation (1) above, and the hyperbolic tangent function, given by:

$$\tanh x = \sinh x/\cosh x = (e^{x} - e^{-x})/(e^{x} + e^{-x}) \qquad (2)$$

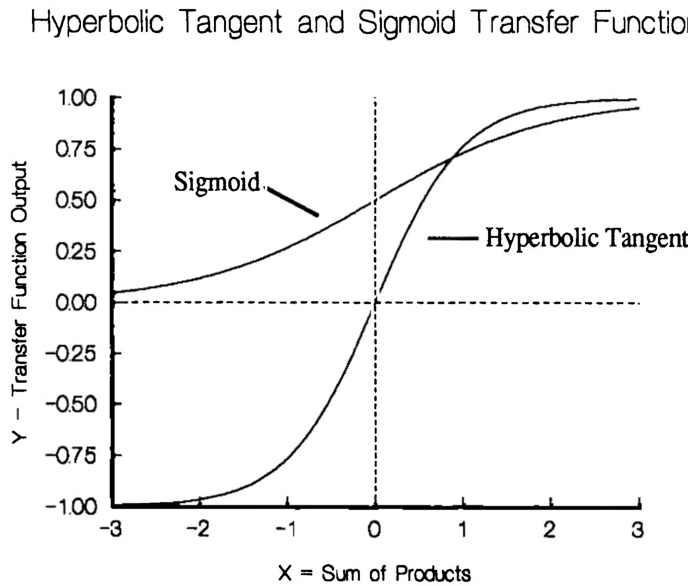
FIGURE 2 contains a comparative illustration of the behavior or these two growth functions.

Once the specified transfer function is applied to each of the six sums of products entering the hidden layer nodes, the six transformed signals are passed up through the connection paths to the output node. The bias node also passes a signal to the output node. As in the layer below, the seven signals are multiplied by connection weights and summed to form another sum of products. The transfer function is again applied to this sum of products to generate the final output signal. If the sigmoid function is used as the network transfer function, the network output value will range from 0 to 1. If the hyperbolic tangent is used, the output value will range from –1 to 1.

Our final network illustrated in FIGURE 1 contains 79 connection paths. This means 79 connection weights must be specified. During training, a neural network is repeatedly presented with sample observations (also referred to as training cases) and a learning rule is required to ensure that all connection weights are modified in a manner that will improve the network's classification ability. In this study, the particular training rule applied during network training is referred to as back-propagation.

Back-propagation is an iterative gradient-descent technique that is similar in many ways to the Newton-Raphson technique used in the maximum likelihood estimation of the logit model. The basic premise underlying the back-propagation algorithm is that each of the network connection weights is, to some degree, responsible for the final output error. Once a network is presented with a new training case, the final network output error is computed using current connection weights. This error is then propagated back through the network and applied to determine how the connection weights should be modified.

FIGURE 2: TRANSFER FUNCTIONS COMMONLY USED IN BACKPROPAGATION NETWORKS



Notes: Sigmoid Ranges Between 0 and 1
Hyperbolic Tangent Ranges Between -1 and 1
Derivative at Point of Inflection:
Sigmoid = .25
Hyperbolic Tangent = 1

The amount of output error that is back-propagated from the output node (call it back-propagated error) is computed by multiplying the derivative of the transfer function times the raw error (raw error is the actual network output value minus the desired, or target, output value). So, the rate of change in the transfer function at its current value also impacts the modification of connection weights. The amount by which the weights on connection paths between the hidden layer and the output layer are modified is determined by multiplying this back-propagated error from the output node times the current input signals that just passed through these six connections (seven connections including the bias). In addition, this amount is typically dampened by applying a learning coefficient that ranges between 0 and 1.

The amount of error to be back-propagated from a node in the hidden layer is determined by multiplying the derivative of the transfer function at its current value (different from the transfer function value at the output node) with the product of the back-propagated error coming into the hidden layer node from connection to the output node above, and the unmodified weight from this connection path. So each node in the hidden layer is assigned a different amount of back-propagated error - an amount that is dependent on the unmodified connection weight from the connection above.

Once these hidden layer back-propagated errors are assigned to each hidden layer node, they are used to modify the connection weights to the input layer in the same manner as before. That is, the delta weight (or weight modification) for a given connection to an input node is derived by computing the product of the respective hidden node's back-propagated error times the input value just passed from the input node, and multiplying this amount times the learning coefficient.

Assuming the network does not get trapped in a local minimum, it has been proved that iterative application of the back-propagation algorithm will improve network performance to the point where the global error is minimized. However, in addition to the potential local minima problem, networks sometimes become "paralyzed", thereby preventing further modification of connection weights. Paralysis can occur when weights become very large. In this case, signals coming into a node become very large, and the derivative of the transfer function approaches zero (see FIGURE 2).

Model Estimation Results

Logistic Regression Model

In an effort to identify a powerful logit model, several exploratory procedures were applied. Initially, we tested each candidate predictor variable for significant differences between the failed and nonfailed sub-sample means and medians using the parametric t-test and the non-parametric Mann-Whitney U-Test. These tests were applied to both the 1984 estimation sample and the 1985 holdout sample. The results are presented in TABLE 2. For 19 of the 28 candidate predictors, failed and nonfailed sub-sample means and medians were significantly different in both years. We limited our consideration of predictors for the multi-ratio model to this set of 19 significant variables.

Next, we estimated numerous multi-ratio logistic regression models using the 1984 estimation sample and assessed overall model goodness-of-fit. Also, we assessed the incremental significance of the individual predictor variables for each model. Further,we compared the signs of the estimated coefficients with expected signs during this stage of the exploratory work. Expected signs of various coefficients are given in TABLE 3. Expectations were based on evidence gathered from prior studies and, in some cases, researcher intuition. Finally, we estimated all possible pair-wise correlations to aid our specification of a final model. Pearson-product-moment correlations in excess of .5 are listed in APPENDIX B.

After much trial-and-error and fine-tuning using the exploratory procedures discussed above, we settled on the final eight-variable model given in TABLE 4. This model includes two loan exposure variables - AGLNS and LNSINSID; three asset quality variables - PDLNSGRL, NONACLNS and RESTRLNS; one capital adequacy variable - PRMCAPAS; one operating performance variable - RATE; and one non-operating performance measure - OVRHDEXP. The overall model likelihood-ratio chi-square statistic was 351.46, which is significant

TABLE 2
PAIRWISE TESTS FOR DIFFERENCES IN MEANS AND MEDIANS

		Data for Year Ended 12/31/84–Failures Occurred During 1985					Data for Year Ended 12/31/85–Failures Occurred During 1986				
					Observed Significance Levels					Observed Significance Levels	
		Mean			Parametric	Mann-Whitney	Mean			Parametric	Mann-Whitney
	Variables	Failed	Nonfailed	Difference	t-Test	U-Test	Failed	Nonfailed	Difference	t-Test	U-Test
	OBS.	102	906	N/A	N/A	N/A	131	928	N/A	N/A	N/A
1	LGASSETS	9.824	10.584	0.760	0.00*	0.00*	10.245	10.639	0.394	0.00*	0.00*
2	CONSTLNS	0.013	0.018	0.005	0.15	0.00*	0.025	0.018	–0.007	0.03*	0.07
3	RLESTLNS	0.039	0.048	0.009	0.06	0.00*	0.059	0.052	–0.007	0.14	0.70
4	AGLNS	0.216	0.084	–0.132	0.00*	0.00*	0.155	0.075	–0.080	0.00*	0.00*
5	PRMCAPAS	0.033	0.085	0.052	0.00*	0.00*	0.018	0.085	0.067	0.00*	0.00*
6	TOCAPLNS	0.053	7.574	7.521	0.65	0.00*	0.031	15.964	15.933	0.54	0.00*
7	NPLNSPCP	3.945	0.230	–3.715	0.00*	0.00*	–1.732	0.248	1.980	0.01*	0.00*
8	NPLNSAST	0.063	0.015	–0.048	0.00*	0.00*	0.084	0.016	–0.068	0.00*	0.00*
9	PDLNSGRL	0.035	0.013	–0.022	0.00*	0.00*	0.041	0.012	–0.029	0.00*	0.00*
10	NONACLNS	0.058	0.014	–0.044	0.00*	0.00*	0.086	0.018	–0.068	0.00*	0.00*
11	RESTRLNS	0.005	0.001	–0.004	0.00*	0.00*	0.012	0.002	–0.010	0.00*	0.00*
12	CHRGOFFS	0.043	0.009	–0.034	0.00*	0.00*	0.062	0.013	–0.049	0.00*	0.00*
13	YIELD	0.113	0.104	–0.009	0.00*	0.00*	0.110	0.098	–0.012	0.00*	0.00*
14	RATE	0.074	0.063	–0.011	0.00*	0.00*	0.072	0.056	–0.016	0.00*	0.00*
15	SPREAD	0.039	0.041	0.002	0.14	0.38	0.038	0.041	0.003	0.00*	0.00*
16	NONINT	0.009	0.008	–0.001	0.38	0.01*	0.011	0.008	–0.003	0.00*	0.00*
17	OVRHDEXP	0.043	0.033	–0.010	0.00*	0.00*	0.050	0.034	–0.016	0.00*	0.00*
18	PROVLOSS	0.036	0.006	–0.030	0.00*	0.00*	0.043	0.008	–0.035	0.00*	0.00*
19	SECGAINS	0.000	0.000	0.000	0.81	0.77	0.002	0.001	–0.001	0.00*	0.00*
20	RETNTA	–0.028	0.008	0.036	0.00*	0.00*	–0.039	0.006	0.045	0.00*	0.00*
21	RETEQ	–0.866	0.080	0.946	0.00*	0.00*	–5.486	0.051	5.537	0.08	0.00*
22	LIQSTAST	–0.059	0.066	0.125	0.00*	0.00*	–0.070	0.068	0.138	0.00*	0.00*
23	TMDEPS	0.162	0.112	–0.050	0.00*	0.00*	0.186	0.112	–0.074	0.00*	0.00*
24	NETLNS	0.646	0.518	–0.128	0.00*	0.00*	0.631	0.512	–0.119	0.00*	0.00*
25	LNSINSID	0.017	0.010	–0.007	0.00*	0.00*	–0.022	0.011	–0.011	0.00*	0.00*
26	CUMPROF	–0.007	0.039	0.046	0.00*	0.00*	–0.028	0.037	0.065	0.00*	0.00*
27	INCOME	–709.882	1451.146	2161.028	0.09	0.00*	–2311.771	635.068	2946.839	0.03*	0.00
28	EQCAPTL	797.500	14128.488	13330.988	0.45	0.00*	1083.802	13938.208	12854.406	0.34	0.00

$\boxed{\text{Highlighted}}$ Variables Significant For Both Years Using Both Tests.

TABLE 3
EXPECTED COEFFICIENT SIGNS FOR SIGNIFICANT PREDICTOR VARIABLES

Feature/Ratio	Expected Sign	Feature/Ratio	Expected Sign
SIZE:		**PERFORMANCE RATIOS:**	
LGASSETS	Minus (–)	YIELD	Plus (+)
LOAN EXPOSURE:		RATE	Plus (+)
AGLNS	Plus (+)	OVRHDEXP	Plus (+)
LNSINSID	Plus (+)	PROVLOSS	Plus (+)
CAPITAL ADEQUACY:		RETNTA	Minus (–)
PRMCAPAS	Minus (–)	CUMPROF	Minus (–)
ASSET QUALITY:		**LIQUIDITY:**	
NPLNSPCP	Plus (+)	LIQSTAST	Minus (–)
NPLNSAST	Plus (+)		
PDLNSGRL	Plus (+)		
NONACLNS	Plus (+)		
RESTRLNS	Plus (+)		
CHRGOFFS	Plus (+)		
NETLNS	Plus (+)		

at the .0000 level. All estimated model coefficients are incrementally significant at the .05 level, and estimated signs agreed with expected signs. The ratio with the greatest incremental explanatory power was PRMCAPAS, while the weakest ratio was NONACLNS.

In order to test for parameter stability, we estimated the same 8-variable model using the 1985 holdout sample. Estimation results for this sample are also given in TABLE 4. As with the estimation sample, the model based on the holdout sample had consistent signs, significant overall model goodness-of-fit, and incrementally significant model coefficients.

Neural Network Model

The process of specifying an appropriate neural net model is even less structured than the exploratory process related to specifying a statistical model. In addition to facing the problem of identifying the appropriate predictor variables for inclusion in the model, one must make additional ad hoc choices about network architecture and training. For example, should you include only one hidden layer? How many nodes should the hidden layer(s) contain? What should be the value of the learning coefficient? Which transfer function should be applied? Should the nodes be fully interconnected or should some connections be disabled or held constant?

TABLE 4
ESTIMATED LOGISTIC REGRESSION EQUATIONS

Variables	Signed Asymptotic t-Statistics 1984	1985
Constant	–3.46	–3.79
AGLNS	5.85	4.29
PRMCAPAS	–5.88	–5.55
PDLNSGRL	4.59	2.51
NONACLNS	1.85	4.19
RESTRLNS	1.91	1.49
RATE	2.03	2.07
OVRHDEXP	5.50	3.97
LNSINSID	3.30	3.68
–2 Times Log Likelihood Ratio [Chi-Sq (8 df)]	351.46	476.97
Sample Sizes		
Failed Banks	102	131
Nonfailed Banks	906	928
Total	1008	1059

Note: Both unweighted and weighted (using the WESML technique) estimations were made for each year. Only the unweighted results are reported. The weighted results were not significantly different.

Due to our lack of experience in the area of neural network modeling, we consulted with NeuralWare, Inc. of Pittsburg, PA and obtained a great deal of helpful advice about network architecture and network training. NeuralWare develops and markets neural network software, and they also provide consulting services in the area of network design, and application-oriented training. They have provided neural net consulting services to many large corporations and Federal agencies, and have established an impressive record of many successful neural net applications.

In an effort to identify an appropriate network architecture for the bank failure prediction process, additional exploratory analyses were undertaken. An all-possible-regressions routine was applied to the estimation sample as a means of identifying additional candidate predictors from our set of 28 ratios. Scatter diagrams were generated for each variable, and outliers were identified. A search was undertaken to identify sample observations with more than one outlier ratio value, but none were found. After exhaustive exploratory data analysis, we decided to include 11 predictor variables in the input layer of the neural network. These ratios included the eight predictors from the final logit model, and TOCAPLNS, RETEQ, and LIQSTAST.

Initially, eight nodes were included in the network's one hidden layer. The network was fully interconnected, and the hyperbolic tangent transfer function was chosen for the purpose of generating all hidden layer outputs and the output layer output. In an effort to avoid network paralysis, the target outputs for failed and nonfailed banks were .9 and -.9 respectively. Mapping sums of products to values inside of the transfer function extremes has resulted in successful avoidance of network paralysis in other applications.

The network training set was comprised of the 102 failed banks from the 1984 estimation sample, and a randomly drawn sample of 102 of the 1984 nonfailed banks. Normalized cumulative back-propagation was chosen as the method for updating network weights. Approximately 300,000 epoch iterations were carried out during the network training phase, and the network root mean square error was monitored throughout the training period. Adjustments were made to the learning coefficient at times when the network error increased significantly.

About halfway through the training process, we decided to disable two nodes within the hidden layer. This decision was made after viewing a Hinton [1987] diagram of the network. The Hinton diagram pictorially portrays the significance of inputs and hidden layer outputs, and at this time it became clear that two of the hidden layer nodes were not making significant contributions to the output layer. At the completion of the training period, the network mean square error was approximately .45.

Prediction Results

Once the final logit and network models were identified, we performed a comparative analysis of the predictive abilities of both models when applied to the full holdout sample of 131 failed and 928 nonfailed banks. Prediction results from applying the logit model are given in TABLE 5. Both the upper and lower tails of the distribution of predicted values contain accurate predictions. For example, at a cutoff value of .01, the model accurately predicts almost 50 percent of the nonfailed sample, and over 99 percent of the failed sample. At a cutoff of .05, the model accurately predicts 95 percent (124 of 131) of the failed banks and 81 percent of the nonfailed banks. Moving to a cutoff of .1, the model accurately predicts over 90 percent of both sub-samples.

The model's predictive ability is also impressive at the top of the distribution. For example, at a cutoff value of .8, the model accurately predicts 99 percent of the nonfailed banks and 52 percent of the failed banks. The predictive strength of the model at the tails indicates that a multi-cutoff decision approach may be beneficial.

After assessing the logit model's predictive ability on the holdout sample, the final step in the research project was to assess whether the neural network could achieve equal or superior predictive performance. To the best of our knowledge, no unambiguous method exists for comparing alternative model predictions. Measuring and comparing both models' "hit rates" at a

TABLE 5
LOGISTIC REGRESSION MODEL HIT RATES AT VARIOUS CUTOFFS
Model Estimated on 1984 Data

	Column Headings = Probability Cutoffs											
	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
1984 Classifications												
Failed Banks	0.990	0.912	0.863	0.775	0.686	0.618	0.559	0.520	0.490	0.412	0.324	0.225
Nonfailed Banks	0.392	0.804	0.896	0.946	0.967	0.977	0.982	0.989	0.993	0.996	0.998	0.999
Totals	0.452	0.814	0.893	0.929	0.938	0.940	0.939	0.941	0.942	0.937	0.930	0.921
1985 Predictions												
Failed Banks	0.992	0.947	0.908	0.863	0.779	0.740	0.695	0.649	0.595	0.519	0.466	0.374
Nonfailed Banks	0.488	0.805	0.900	0.941	0.961	0.968	0.973	0.976	0.983	0.988	0.991	0.996
Totals	0.551	0.822	0.901	0.931	0.939	0.940	0.939	0.936	0.935	0.930	0.926	0.919

Descriptive Statistics for Predicted Probabilities:

	1984 Classification Results		1985 Prediction Results	
	Failed	Non-Failed	Failed	Non-Failed
Minimum Probability	.008	.000	.010	.000
Median Probability	.673	.014	.825	.010
Mean Probability	.578	.048	.684	.052
Maximum Probability	1.000	.966	1.000	.987

particular cutoff value would not be appropriate unless the distributions of predicted values from applying both models are identical. Obviously, this is not true for our models since the logit model maps a single sum of products to a point on the sigmoid function, and the neural net model maps its output-node sum of products to a point on the hyperbolic tangent function. Even if the sigmoid function was used in the neural network, generating a distribution of predictions identical to the logit model is highly unlikely.

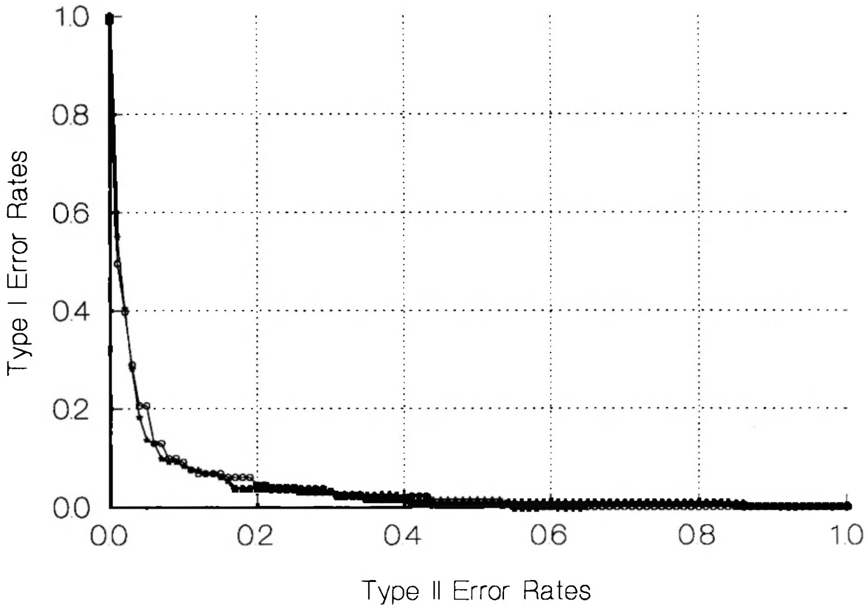
One approach that would allow for comparability across models involves estimating the relative costs of Type I and Type II errors, and then determining the optimal cutoff associated with each model's minimum misclassification costs. Due to the high degree of uncertainty involved in the identification of relative misclassification costs of Type I and Type II errors, researchers typically assume several alternative relative cost ratios, and identify the optimal cutoffs under each assumption. The idea is to determine if either model dominates the other in terms of minimum misclassification costs across a frontier of optimal cutoffs associated with assumed cost ratios.

We decided to measure the entire range of tradeoffs between the failed and nonfailed sub-sample error rates for both models and then visually inspect the relative positions of both tradeoff functions using graphical analysis. By using tradeoff functions we avoid the inappropriate use of specific cutoff values. Instead, we can compare the predictive abilities of both models across the entire frontier of all possible cutoffs. If one model produces a tradeoff function that falls below the second model's tradeoff function in at least one spot, and does not fall above the second model's function at any point, the first model can be judged superior to the second.

FIGURE 3 contains overlaid graphs of the logit and neural net models' tradeoff functions. Visual inspection of FIGURE 3 reveals that neither the logit model nor the neural net model dominates in terms of predictive ability. The tradeoff functions for both models cross one another at several points. It should be noted that one additional Type I error shifts the Type I error rate upward by .0076 (1/131), or approximately .01. Therefore, most of the differences between the logit model and the neural network are not greater than two holdout sample failed banks.

Specific portions of the graph in FIGURE 3 were magnified and are presented in FIGURES 4 through 6. FIGURE 4 focuses on the top tail of the tradeoff functions where high failed and low nonfailed errors rate are found. Model cutoffs related to these error rates would be appropriate if the cost of misclassifying a nonfailed bank is greater than the cost of misclassifying a failed bank. The logit tradeoff function remains slightly below the neural net function over this region of the frontier where misclassifications of nonfailed banks remain below two percent. FIGURE 5 focuses on the central portion of the tradeoff functions where Type II error rates range between two percent and ten percent. The neural network's tradeoff function is below the logit model's over most of this region. FIGURE 6 focuses on the bottom portion of the tradeoff functions where high TYPE II and low Type I error rates are given. The neural net continues to outperform the logit model up to the point where the Type II error rate is 20 percent, and then the models' performances reverse.

FIGURE 3: TRADEOFFS BETWEEN TYPE I AND TYPE II ERROR RATES
(Circles Denote the Logit Model; Stars Denote the Neural Network)



The largest difference between the two models is found at the point where the Type II error rate is .05 (see FIGURE 5). At this point, the neural net correctly predicts nine more bank failures than the logit model. Across the entire tradeoff frontier, only three points are found where the difference between the two models is greater than 3 mispredictions.

Summary

The preliminary results indicate that neither modeling approach dominates the other in terms of predictive ability across the entire frontier of all possible model cutoffs. On average, the neural network model does appear to perform equally as well as the logistic regression model. According to the neural network literature, the back-propagation network may be desirable when a decision process is inherently nonlinear, with many interactions among the input cues, and/or when a cascaded approach to data processing is used by the decision maker. In the case of regulators' decisions to close commercial banks, the preliminary evidence implies that these process attributes do not exist.

FIGURE 4: TRADEOFF FUNCTIONS AT LOW TYPE II ERROR RATES (LESS THAN .02)
(Circles Denote the Logit Model; Stars Denote the Neural Network)

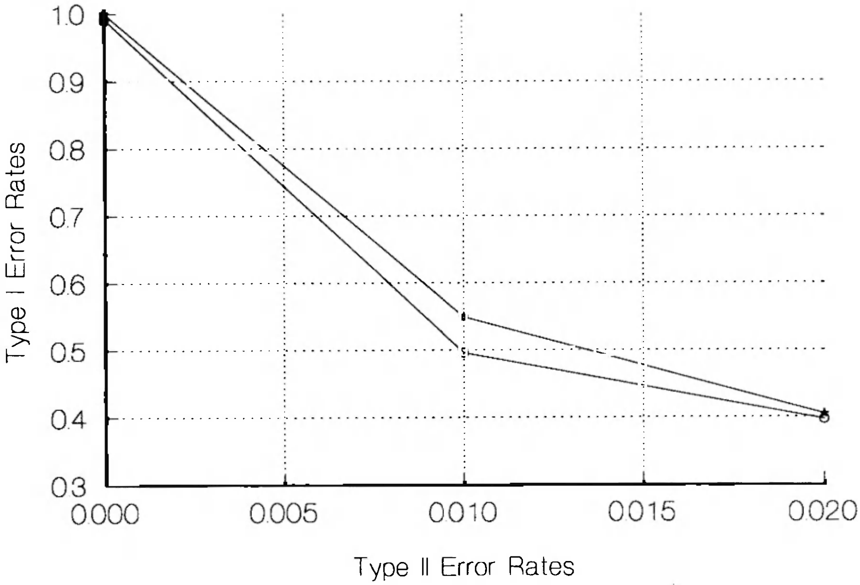


FIGURE 5: TRADEOFF FUNCTIONS AT MID TYPE II ERROR RATES (BETWEEN .02 AND .1)
(Circles Denote the Logit Model; Stars Denote the Neural Network)

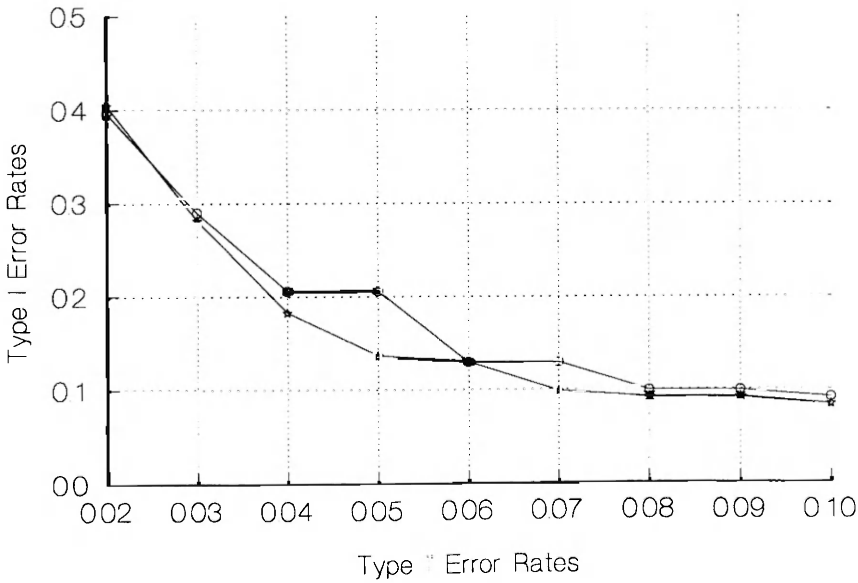
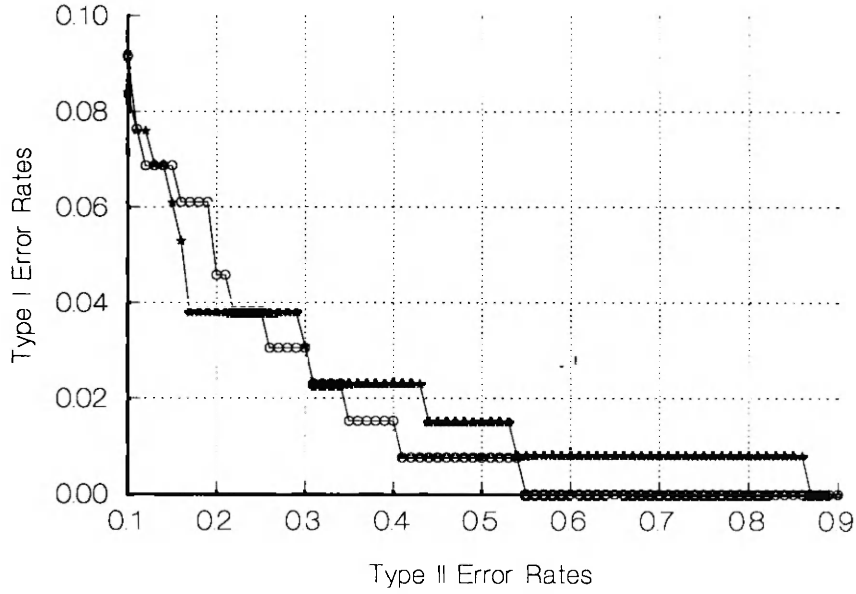


FIGURE 6: TRADEOFF FUNCTIONS AT HIGH TYPE II ERROR RATES (GREATER THAN .1)
(Circles Denote the Logit Model; Stars Denote the Neural Network)



References

American Institute of Certified Public Accountants, *AICPA Professional Standards,* Volume 1 (Commerce Clearing House, 1988).

Altman, E. I., R. B. Avery, R. A. Eisenbeis, and J. F. Sinkey, *Application of Classification Techniques in Business, Banking and Finance* (JAI Press, 1981).

Hinton, G. E., and T. J. Sejnowski, "Neural Network Architectures for AI," *Tutorial Number MP2, National Conference on Artificial Intelligence (AAAI-87),* 6th, Proceedings (July 13-17, 1987), pp. 9-14.

OCC, *Bank Failure: An Evaluation of the Factors Contributing to the Failure of National Banks* (Washington, D. C., June 1988).

Sheshunoff Information Services, Inc., Austin, Texas 78704.

APPENDIX A
CANDIDATE PREDICTOR VARIABLE DEFINITIONS

Variable Acronym	Variable Name and Formula	
1. LGASSETS	***Natural Logarithm of Total Assets***	
	Natural Logarithm of Total Assets	
2. CONSTLNS	***Construction Loans to Total Assets***	
	Numerator:	Construction & Land Development Loans
	Denominator:	Total Assets
3. RLESTLNS	***Commercial Real Estate Loans to Total Assets***	
	Numerator:	Loans Sec: Nonfarm + Loans Secured by 5+Res
	Denominator:	Total Assets
4. AGLNS	***Agricultural Loans to Total Assets***	
	Numerator:	Ag Prod & Farm Loans + Loans Secured by Farm
	Denominator:	Total Assets
5. PRMCAPAS	***Primary Capital to Adjusted Assets***	
	Numerator:	Total Equity Capital + Minority Interest + Total Mand Conv. in Cap – Allowance for Losses
	Denominator:	Total Assets – Allowance for Losses
6. TOCAPLNS	***Total Capital to Total Loans***	
	Numerator:	Total Equity Capital + Minority Interest + Total Man Conv. in Cap + Subordinated Notes & Deb + Ltd Life Pref Stock – Allowance for Losses
	Denominator:	Loans, Net: Unearn Inc.

APPENDIX A (CONTINUED)
CANDIDATE PREDICTOR VARIABLE DEFINITIONS

Variable Acronym	Variable Name and Formula	
7. NPLNSPCP	***Nonperforming Loans to Primary Capital***	
	Numerator:	Total Past Due Loans + Total Nonaccrual Loans
	Denominator:	Total Equity Capital + Minority Interest + Total Man Conv. in Cap – Allowance for Losses
8. NPLNSAST	***Nonperforming Loans to Total Assets***	
	Numerator:	Total Past Due Loans + Total Nonaccrual Loans
	Denominator:	Total Assets
9. PDLNSGRL	***Past Due Loans to Gross Loans***	
	Numerator:	Total Past Due Loans
	Denominator:	Loans & Leases + Unearned Income
10. NONACLNS	***Nonaccrual Loans to Gross Loans***	
	Numerator:	Total Nonaccrual Loans
	Denominator:	Loans & Leases + Unearned Income
11. RESTRLNS	***Restructed Loans to Gross Loans***	
	Numerator:	Total Restructured Loans
	Denominator:	Loans & Leases + Unearned Income
12. CHRGOFFS	***Net Chargeoffs to Total Loans***	
	Numerator:	Total Chargeoffs – Total Recoveries
	Denominator:	Loans & Leases

APPENDIX A (CONTINUED)
CANDIDATE PREDICTOR VARIABLE DEFINITIONS

Variable Acronym	Variable Name and Formula	
13. YIELD	***Yield on Total Assets***	
	Numerator:	Total Interest Income
	Denominator:	Total Assets
14. RATE	***Rate Paid on Total Assets***	
	Numerator:	Total Interest Income
	Denominator:	Total Assets
15. SPREAD	***Net Interest Income to Total Assets***	
	Numerator:	Total Interest Income – Total Interest Expense
	Denominator:	Total Assets
16. NONINT	***Noninterest Income to Total Assets***	
	Numerator:	Total Noninterest Income
	Denominator:	Total Assets
17. OVRHDEXP	***Total Overhead Expense to Total Assets***	
	Numerator:	Total Noninterest Expense + Interest on Mtge Indebtedness
	Denominator:	Total Assets
18. PROVLOSS	***Provision for Loan Loss to Total Assets***	
	Numerator:	Prov: Loan & Lease Loss + Prov: All Transfer Risk
	Denominator:	Total Assets
19. SECGAINS	***Security Gains (Losses) & Extra. Items to Total Assets***	
	Numerator:	Gains (Losses) on Sec + Extra Items, Gross
	Denominator:	Total Assets

APPENDIX A (CONTINUED)
CANDIDATE PREDICTOR VARIABLE DEFINITIONS

Variable Acronym	Variable Name and Formula	
20. RETNTA	***Return on Total Assets***	
	Numerator:	Inc. before Extra. Items
	Denominator:	Total Assets
21. RETEQ	***Return on Equity***	
	Numerator:	Inc. before Extra. Items
	Denominator:	Total Equity Capital
22. LIQSTAST	***Short-Term Assets Less Large Liabs. to Total Assets***	
	Numerator:	Due: Int. Bearing + Federal Funds Sold + Assets in Trading Accts. + Debt Sec. Reprc <1 Yr. – Time CDs >$100M – Open Acct. Time >$100M – Dep: For Nonint. Bearing – Dep: For Int. Bearing – Federal Funds Purchased – Notes Issued to U.S. Treas. – Liab. for Borrowed $
	Denominator:	Total Assets
23. TMDEPS	***Large Time Deposits to Total Assets***	
	Numerator:	Time CDs >$100M + Open Acct. Time >$100 M
	Denominator:	Total Assets
24. NETLNS	***Net Loans to Total Assets***	
	Numerator:	Loans & Leases – Allowance for Losses
	Denominator:	Total Assets

APPENDIX A (CONTINUED)
CANDIDATE PREDICTOR VARIABLE DEFINITIONS

Variable Acronym	Variable Name and Formula	
25. LNSINSID	***Loans to Insiders over Net Loans***	
	Numerator:	**Credit to Officers Agg. Amt.**
	Denominator:	**Loans & Leases – Allowance for Losses**
26. CUMPROF	***Undivided Profit & Cap. Reserve to Total Assets***	
	Numerator:	**Undivided Profit & Cap. Reserve**
	Denominator:	**Total Assets**
27. INCOME	***Income before Extra. Items***	
28. EQCAPTL	***Total Equity Capital***	
	Numerator from TOCAPLNS	

APPENDIX B
SIGNIFICANT PAIRWISE CORRELATIONS

CORRELATIONS OVER .5 – 1984 DATA

	NPLNSAST	CHRGOFFS	PROVLOSS	LIQSTAST	YIELD	RETNTA	PRMCAPAS	NONACLNS	TOCAPLNS	OVRHDEXP
PDLNSGRL	.641									
NONACLNS	.838					–.575				
CHRGOFFS	.603							.591		
PROVLOSS	.662	.858						.595		
RETNTA	–.633	–.741	–.858							
NETLNS				–.648						
TMDEPS				–.730						
RATE					.699		–.547			
CUMPROF						.616				
NONINT									.593	.713
SPREAD										.500

APPENDIX B (CONTINUED)
SIGNIFICANT PAIRWISE CORRELATIONS

CORRELATIONS OVER .5 – 1985 DATA

	NPLNSAST	CHRGOFFS	PROVLOSS	LIQSTAST	YIELD	RETNTA	PRMCAPAS	NONACLNS	OVRHDEXP	TOCAPLNS	INCOME
PDLNSGRL	.702										
NONACLNS	.840										
CHRGOFFS	.518		.888					.579			
PROVLOSS	.578							.582			
RETNTA	–.596	–.768	–.874					–.576			
NETLNS				–.648							
TMDEPS				–.717							
RATE					.730		–.557				
CUMPROF			–.545			.752			–.524		
NONINT									.648	.530	
SPREAD					.606						
EQCAPTL											–.788