© *The Journal of Risk and Insurance*, 2006, Vol. 73, No. 3, 397-419

A Comparison of Neural Network, Statistical Methods, and Variable Choice for Life Insurers' Financial Distress Prediction

Patrick L. Brockett
Linda L. Golden
Jaeho Jang
Chuanhou Yang

Abstract

This study examines the effect of the statistical/mathematical model selected and the variable set considered on the ability to identify financially troubled life insurers. Models considered are two artificial neural network methods (back-propagation and learning vector quantization (LVQ)) and two more standard statistical methods (multiple discriminant analysis and logistic regression analysis). The variable sets considered are the insurance regulatory information system (IRIS) variables, the financial analysis solvency tracking (FAST) variables, and Texas early warning information system (EWIS) variables, and a data set consisting of twenty-two variables selected by us in conjunction with the research staff at TDI and a review of the insolvency prediction literature. The results show that the back-propagation (BP) and LVQ outperform the traditional statistical approaches for all four variable sets with a consistent superiority across the two different evaluation criteria (total misclassification cost and resubstitution risk criteria), and that the twenty-two variables and the Texas EWIS variable sets are more efficient than the IRIS and the FAST variable sets for identification of financially troubled life insurers in most comparisons.

Introduction

Insurance company insolvency produces substantial losses to many stakeholders, and the identification of financially troubled firms is a major regulatory objective. Accordingly, there is a strong regulatory need for accurate prediction methods to

Patrick L. Brockett and Linda L. Golden are at McCombs School of Business, University of Texas at Austin. Jaeho Jang is at Samsung. Chuanhou Yang is at Dahlkemper School of Business, Gannon University. The authors can be contacted via e-mail: brockett@mail.utexas.edu.
We gratefully acknowledge the support of the Catalan government to Linda Golden and the Spanish government to Patrick Brockett during the creation of this article. The hospitality of Monserrat Guillen and the Department of Econometrics at the University of Barcelona, Spain is also appreciatively acknowledged.

signal financially impaired insurers in sufficient time to allow action to be taken to prevent insolvency or to minimize its cost.

In the context of warning of pending insurer insolvency, there are several sources of information available. These include, for example, the A. M. Best and other rating agency reports. In addition, the National Association of Insurance Commissioners (NAIC) developed the insurance regulatory information system (IRIS) and, following the extremely costly First Executive Life Insurance Company bankruptcy, the financial analysis solvency tracking (FAST) system to provide an early warning system. The NAIC also adopted risk-based capital (RBC) formula for insurance insolvency prediction. Some states have developed their own early warning systems. For example, the Texas Department of Insurance (TDI) implemented an early warning information system (EWIS) in early 1992 based upon their own model and data set.

There is substantial previous literature on insurer insolvency prediction: Barrese (1990) evaluated the adequacy of IRIS; Cummins, Harrington, and Klein (1995), and Grace, Harrington, and Klein (1998) provided evaluations concerning the accuracy of the RBC and FAST systems; and Cummins, Grace, and Phillips (1999) compared RBC and FAST using cash flow simulation.

Multivariate statistical approaches such as multiple discriminant analysis (MDA) and logistic regression (logit) have been explored in the literature. Trieschmann and Pinches (1973, 1974) reported that the six-variable MDA model outperforms all univariate models. BarNiv and Hershbarger (1990) demonstrated that logit and the nonparametric discriminant analysis outperform MDA in most situations. BarNiv and McDonald (1992) reviewed the literature and also show that qualitative response models such as Probit or logit can provide better predictions of both solvency and insolvency cases than does the MDA. Carson and Hoyt (1995) found that the logit model dominated the MDA and Recursive Partitioning (RP) models in terms of the number of correctly classified solvent insurers, and the RP model dominated the logit and the MDA model in terms of the number of correctly classified insolvent insurers. Carson and Hoyt (2000) use logistic regression to estimate the insolvency factors for the life insurers in the European Union. Baranoff, Sager, and Witt (1999) and Baranoff, Sager, and Shively (2000) constructed cascaded regression and nonlinear spline models, respectively, for insolvency prediction as well.

Nonparametric methods, such as neural network methods, have also become popular. Brockett et al. (1994) use back-propagation neural network methods to compute an estimate of the propensity toward insolvency for the property and casualty insurers. The neural network results show high predictability and generalizability, suggesting the usefulness of this method for predicting future insurer insolvency. Huang, Dorsey, and Boose (1994) used the back-propagation neural network method to forecast financial distress in life insurers. Their data, however, are limited to IRIS ratios and they do not attempt to identify an optimal set of variables for forecasting.

This study differs from previous research in five ways. First, it compares and examines the performance of two artificial neural network methods (back-propagation and learning vector quantization (LVQ)), and also makes comparisons with standard MDA and logit analysis. Second, among current available variable sets, it attempts to identify a superior choice of variables for insolvency prediction.

Third, instead of training the models using a list of public pronounced technically insolvent insurers (with receivership or liquidation declarations) to identify insolvency, it uses an expanded data set which also includes financially "troubled" insurers confidentially identified by the Financial Analysis Unit of the TDI. In this study, financially "troubled" companies are those which have been given an official "article 1.32" or "hazardous financial condition" notification by TDI. Because it is desirable (from the regulator's perspective) that troubled companies be identified early in order to have sufficient time to remedy the financial situation, and because the more severe regulatory actions such as administrative oversight, confidential supervision, supervision, conservatorship, receivership, and liquidation, often occur too late for remediation, this weaker signal (article 1.32 order) is an important earlier benchmark stage of financial hazard.[1] Indeed, one motivation for this study was to increase the ability of the regulator to adequately rehabilitate companies earlier in the regulatory process.

Fourth, this study considers two prediction time frames: a model to predict financial hazard during the current year, and a model to predict financial hazard one year out. Because this study attempts to predict based on less severe regulatory actions (1.32 actions), it is not imperative that the prediction be made as far into the future as done in the previous academic literature which (by necessity of data availability) required publicly available announcements of supervision, conservatorship, receivership, or liquidation.

Finally, this study uses out-of-sample tests which employ the weights or parameter estimates developed in the previous year to predict outcomes (hazardous financial condition or not) in the subsequent year. The out-of-sample test is practically useful in the sense that the regulator can use the *ex post* information of the past to predict the *ex ante* outcomes for the current year and future year, and this process can be updated annually by the regulator.

METHODOLOGY

This section describes the four statistical models or techniques used for predicting life insurers' financial vulnerability in this study: multivariate discriminant analysis (MDA), logistic regression analysis, and the two artificial neural network models (back-propagation and LVQ). For the models examined, we let $(x_1, x_2, \ldots, x_k)$ represent the variables collected at time t which are inputs into the financial hazard prediction models used to classify an insurer into one of two groups or populations π_1 and π_2 (financially troubled or not financially troubled)[2] at some future time $t + 1$. We assume, we have available a sample of insurers for which we know both their defining characteristics $x = (x_1, x_2, \ldots, x_k)$ and group membership, so the parameters of the models can be fit (supervised learning models in the context of artificial intelligence).

[1]The collection of companies so identified are not necessarily made public by the regulator for legal reasons. Access to this confidential data was permitted by TDI for this study, and was conducted together with Texas Department of Insurance personnel.

[2]Previous research has often used the two group dichotomy "insolvent/solvent" rather than "financially troubled/not financially troubled" because they were using publicly available solvency determinations to train their models. We have adjusted terminology to reflect the nature of the dichotomy used in our data.

R. A. Fisher introduced MDA to solve exactly the above problem in the last century under the assumption of multivariate normality of $x = (x_1, x_2, \ldots, x_k)$ with common covariance matrix within populations π_1 and π_2. Since Altman (1968) first used this method for predicting financial distress, it has become a commonly used parametric method for addressing this issue in various industries (Edmister, 1972; Sinkey, 1975).[3] The interested reader may consult any standard multivariate statistics book for formulae and further details. Common statistical packages such as SAS can be used to perform computations and provide classifications using MDA under either an assumption of equally *a priori* group membership (priors equal) or an assumption of that the likelihood of group 1 versus group 2 membership is the same as that observed in the training sample (priors proportional).

As mentioned earlier, logistic regression (or logit analysis) is another common classification technique for insolvency prediction. In this model, instead of using a linear function of $(x_1, x_2, \ldots, x_k)$ to classify into groups, the log odds ratio of group 1 membership is modeled as a linear function of $(x_1, x_2, \ldots, x_k)$. As with MDA analysis, standard multivariate texts provide formulae, and standard statistical computer packages such as SAS can perform the computations.

Both MDA and logit models suffer from potentially restrictive parametric assumptions whose violation can hinder performance. To overcome these problems nonparametric methods have been introduced. For the classification of financial distress Salchenberger, Cinar, and Lash (1992), Coats and Fant (1993), Luther (1993), Huang, Dorsey, and Boose (1994), and Brockett et al. (1994) have shown that neural network models can out-perform MDA and logit analysis. In the context of insurance fraud detection, Viaene et al. (2002) compared and contrasted several of these models as well as other expert system methods such as neural networks.

Neural network models were first developed in an attempt to simulate the processes of the brain. Just as the brain consists of a network of interconnected neurons, a mathematical neural network consists of interconnected nodes (referred to as processing elements (PE)) that receive, process, and transmit information. A processing element has many input paths and combines the values of these input paths using a weighted summative structure. The combined input is modified by a transfer function which can be a threshold function or a continuous function of the combined input. The output value of the transfer function is passed directly to the output path of the processing element. It allows for nonlinearity in the relationship between inputs and outputs.[4] Brockett et al (1994) gave more precise details and intuitive explanation.

The most popular neural network model is the single layered back-propagation feed-forward network (c.f., Bryson and Ho, 1969; Werbos, 1974; Parker, 1985; and Rumelhart, Hinton, and Williams, 1986). The algorithm iteratively updates the weights w_{pq} (associated with the connection from unit q to unit p) to allow the network to

[3]MDA has received criticism since the data used often violate the assumptions of this model. In the current data sets, for example, the assumption of normality was tested for all the independent variables, and the results demonstrated that none of these variables were normally distributed.

[4]In the area of insurer insolvency prediction, the importance of recognizing and modeling potential for nonlinearity was addressed previously by Baranoff, Sager, and Shively (2000) using another method.

"learn" the pattern exhibited in the training set. A mathematical and intuitive discussion of neural networks in the context of insurer insolvency prediction is given in Brockett et al.(1994).[5]

LVQ is another neural classification network, originally suggested by Teuvo Kohonen (c.f., Kohonen, Barna, and Chrisley, 1988; Kohonen, 1998). An LVQ network contains a "Kohonen layer" which learns and performs the classification, an input layer, and an output layer. The intuitive logic of the Kohonen layer is that if two input vectors are close together in some metric, then their output values should also be close together, so we adjust the weights in the network to make this happen as much as possible. We represent the network PEs by the weight vectors between the nodes. We proceed iteratively through the training data sample. Starting with an initial set of weights between the input neurons in the network and the output neuron, we "reinforce" the connection weights for a PE if, for the training sample input used, the closest input vector yields the correct classification for the given input. If the closest input vector is not correct, then we discourage the connection. Encouragement and discouragement are accomplished by adjusting the connection weights. See Brockett, Derrig, and Xia (1998), Kohonen, Barna, and Chrisley (1988), and Kohonen (1998) for more details and intuition on Kohonen networks. Kohonen et al. (1996) provided a program for conducting the analysis.

DATA, VARIABLES, AND RESEARCH DESIGN

The data were obtained from TDI using insurer annual statements for 1991 through 1994, as well as a list of the insurers created by TDI that were designated as being "troubled" from 1991 to 1995 (i.e., were given a so-called 1.32 action by TDI indicating a hazardous financial condition, or were given a more severe regulatory action such as administrative oversight, confidential supervision, supervision, conservatorship, receivership, or liquidation). Most insolvency studies have used a list of publicly declared technically insolvent insurers.[6] However, it is desirable that troubled companies be identified earlier so regulators have the intervention time to remedy deteriorating financial situations.

The majority of previous company failure prediction studies uses a matched-pair sampling method for training and fitting parameters in their models, which may bias the ability of their empirical error rates to generalize to that expected in practice (c.f., Palepu, 1986). We overcome this by using a full collection of companies as a sample, namely all solvent and insolvent life insurance companies whose business domiciles are in Texas and whose full set of data fields are available for the entire study period (1991–1994).[7]

This study examines four choice sets of potential explanatory variables: 22-variable set, IRIS variables, FAST variables, and Texas EWIS variables. The first three contain

[5]For this study, the actual computations were performed using the program "Predict" by Neuralwares; however, standard computer packages such as SAS can also perform back propagation neural network computations.

[6]Lists of companies designated as being in hazardous financial condition or under confidential supervision are not generally available to academic researchers.

[7]As different data sets contain different variables, and different companies have different missing variables, this implies that there are different sample sizes for the different analyses.

continuous financial ratio variables thought to be indicative of a firm's financial health. The last set consists of binary indicator variables constructed by TDI to be indicative of the insolvency propensity of an insurer. These sets are described as follows.

22-Variable Set

The 22-variable set was constructed by examining previous studies to create a master set of variables previously found to be indicative of financial distress. Some variables were eliminated because too many companies had no data for the variable. To further aid in variable choice, a stepwise regression procedure was used in conjunction with consultation with area experts at TDI. Table 1 shows these 22 variables.

IRIS Variable Set

One primary method used by state regulators to monitor the financial strength of insurance companies is the IRIS, an early warning system developed by the NAIC. The IRIS system for life insurers consists of twelve financial ratios from annual statements. If four of these twelve ratios fall outside an acceptable range of values, the company is immediately given further regulatory examination (c.f., Brockett et al., 1994).

TABLE 1
Description of 22-Variable Set

Variable Name	Description
V1	Gains/Premiums
V2	Liabilities/Surplus
V3	Net gain from operations after tax & dividends
V4	Net investment income
V5	Accident & health benefits/total benefits
V6	(Bonds+stocks+mortgages)/Cash & investment assets
V7	Cash flow/liabilities
V8	Capital & surplus/liabilities
V9	Change in capital & surplus
V10	Delinquent mortgages/capital & surplus
V11	Change in premium
V12	Insurance leverage (reserves/surplus)
V13	Financial leverage (premiums/surplus)
V14	Log of growth in assets
V15	Log of growth in premiums
V16	Log of growth in surplus
V17	Log of cash flow from operations
V18	Nonadmitted assets/admitted assets
V19	Reinsurance ceded/premium
V20	Separate account assets/assets
V21	Total benefits paid/capital & surplus
V22	Real estate/assets

FAST Variable Set

Although it has achieved some success in helping to identify financially troubled insurers, the efficacy of IRIS has been questioned, claiming that it is too dependent on the capital and surplus figures and that it treats each ratio in isolation and fails to take into account the interrelationships between the ratios. In response to these criticisms, NAIC has developed a new early warning system called the FAST system. FAST consists of seventeen financial ratios and variables based on annual and quarterly statement data for life insurers. Unlike the original IRIS ratios, it assigns different point values for different ranges of ratio results. A cumulative score is derived for each company, which is used to prioritize it for further analysis.[8] See Grace, Harrington, and Klein (1998) for more details on the FAST variables.

Texas EWIS Variable Set

To improve detection of potentially insolvent companies, TDI internally implemented their own EWIS in early 1992. For each company, a set of 393 binary indicator variables was calculated based upon some ratio or numerical value and some preselected threshold value. Weights were assigned to each binary indicator by the EW staff according to a subjective assessment of the importance or severity of the indicator. Each binary indicator was then multiplied by its assigned weight and the resulting values were summed across all indicators to obtain an "EWIS company score" for each company which was then used for prioritization of the insurers' examination. Some of the 393 indicators can be automatically input into the EWIS system from the company's annual statement, whereas some others require an analyst to input scores manually. Accordingly, we examine two separate Texas EWIS variable sets in this analysis: an automated indicator set (EWIS-automated) and a nonautomated indicator set (EWIS-significant).[9]

This study considers two different models, a current year model and a 1-year prior model. It is not imperative that hazardous financial condition predictions be made very far into the future since the hazardous financial condition (1.32 action) is already an early stage. Due to the slow-moving nature of the regulatory process, there can be a significant lag (usually years, and sometimes involving court action) between a company receiving a 1.32 action or confidential supervision and their actual public listing as under conservation or receivership or liquidation.

This study uses out-of-sample tests which employ the weights or parameter estimates developed in the previous year(s) to predict the outcomes in the subsequent year. The learning or training samples consist of the insurance companies in 1992 and 1993. The parameter estimates and weights from the learning samples are used to test the sample which consists of the companies in 1994. The number of companies in the training and test samples is listed in Table 2.[10]

[8]The twelve IRIS ratios and seventeen FAST variables are available from NAIC (http://www.NAIC.org).

[9]The variables used in TDI–EWIS are not publicly disclosed.

[10] Some summary statistics and tests of differences for the troubled and nontroubled companies in our sample are available from the authors but are not presented here for space reasons.

TABLE 2
Number of Financially Troubled and Not Financially Troubled Texas Insurers

Year	Variable Set	Number of Insurers		Percent of Troubled Insurers
		Troubled	Not Troubled	
1991	22-Variable	70	463	13.13
	IRIS	51	463	9.92
	FAST	66	463	12.48
	Texas EWIS	32	463	6.46
1992	22-Variable	64	463	12.14
	IRIS	51	463	9.92
	FAST	50	463	9.75
	Texas EWIS	50	463	9.75
1993	22-Variable	55	463	10.62
	IRIS	51	463	9.92
	FAST	54	463	10.44
	Texas EWIS	49	463	9.57
1994	22-Variable	49	463	9.57
	IRIS	49	463	9.57
	FAST	48	463	9.39
	Texas EWIS	46	463	9.04

EMPIRICAL RESULTS: PERFORMANCE COMPARISON AMONG PREDICTION METHODS

Prediction Efficiency in Terms of Misclassification Cost

Total misclassification costs can be calculated as $MC = C_1 n_1 + C_2 n_2$, where C_1 is the cost of type I misclassification, C_2 the cost of type II misclassification, n_1 the number of misclassified bankrupt firms, and n_2 the number of misclassified non-bankrupt firms. Because it is difficult to measure relative costs associated with type I and II errors, the estimation of the relative cost is a subjective approximation. Zavgren (1983) suggested that the prior probability of 0.1 for failure with the type I to type II cost ratio of 20 would give the best results. Different prior probabilities and misclassification cost ratios are used in this study. The type I misclassification cost takes on the values from 1 to 30 whereas the type II misclassification cost is fixed at 1. The prior probability of failure is set to equal prior probability and the proportional prior probability in the sample. An evaluation of MDA, logit, LVQ, and BP methods is performed for each of the five data sets using the current year and 1-year prior model.[11] Some of the results are presented in Tables 3–7.[12]

[11] The parameter values for the neural network approaches were obtained using the commercial software Predict by Neuralwares. Many modern software packages such as SAS have neural network options for analysis and these yield equivalent results.

[12] Due to space limitation, only the results of the 1-year prior model are presented in this article. The results of the current year model are available from the authors.

TABLE 3
Comparison of the Predictive Accuracy of the Methods Using Misclassification Cost Criterion: 22-Variable Set and 1-Year Prior Model

Year	Method	No. of Type I Errors	No. of Type II Errors	Correct Rate	Total Misclassification Cost at Ratio C_1/C_2					
					1	10	15	20	25	30
1992	MDA	18	43	0.884	61	223	313	403	493	583
	MDA-prop	32	15	0.911	47	335	495	655	815	975
	Logit	33	15	0.909	48	345	510	675	840	1,005
	Logit-prop	10	66	0.856	76	166	216	266	316	366
	LVQ	4	8	0.977	12	48	68	88	108	128
	BP	1	27	0.947	28	37	42	47	52	57
1993	MDA	17	47	0.876	64	217	302	387	472	557
	MDA-prop	32	14	0.911	46	334	494	654	814	974
	Logit	27	14	0.921	41	284	419	554	689	824
	Logit-prop	11	57	0.869	68	167	222	277	332	387
	LVQ	6	7	0.975	13	67	97	127	157	187
	BP	2	31	0.936	33	51	61	71	81	91
1994	MDA	16	48	0.875	64	208	288	368	448	528
	MDA-prop	32	16	0.906	48	336	496	656	816	976
	Logit	35	13	0.906	48	363	538	713	888	1,063
	Logit-prop	15	59	0.855	74	209	284	359	434	509
	LVQ	8	7	0.971	15	87	127	167	207	247
	BP	1	18	0.963	19	28	33	38	43	48

We can see that a consistent pattern of minimum total misclassification costs emerges.[13] For lower cost ratios (under 3), the LVQ method tends to minimize total misclassification costs. This is consistent with its minimization of type II errors. For higher cost ratios (5–30), the back-propagation method is superior in minimizing total misclassification costs because of the smallest numbers of type I errors. For these data sets, the MDA and logit methods (equal and proportional prior probability) fail consistently to minimize total misclassification costs.

The results also indicate that there is little difference in performance between the current year model and the 1-year prior model. This lack of a significant difference between the performance of the current year and 1-year prior models is probably due to the use of a broader definition of financial distress. The use of the designation "financially troubled" as the dichotomization variable in this study includes insurers at an earlier stage of financial hazard, so progression toward insolvency is slower, and some will be rehabilitated.

[13] Some summary statistics that are independent of the costs of misclassification, such as the number of type I errors, the number of type II errors, and the correct prediction rate, are also presented in the tables. The predictive accuracy of the various models can also be evaluated using the receiver operator characteristic (ROC) analysis. For details about the ROC analysis see Cummins, Grace, and Phillips (1999).

TABLE 4
Comparison of the Predictive Accuracy of the Methods Using Misclassification Cost Criterion: IRIS Variable Set and 1-Year Prior Model

Year	Method	No. of Type I Errors	No. of Type II Errors	Correct Rate	Total Misclassification Cost at Ratio C_1/C_2					
					1	10	15	20	25	30
1992	MDA	19	66	0.835	85	256	351	446	541	636
	MDA-prop	42	10	0.899	52	430	640	850	1,060	1,270
	Logit	46	9	0.893	55	469	699	929	1,159	1,389
	Logit-prop	18	94	0.782	112	274	364	454	544	634
	LVQ	10	5	0.971	15	105	155	205	255	305
	BP	4	11	0.971	15	51	71	91	111	131
1993	MDA	18	61	0.846	79	241	331	421	511	601
	MDA-prop	38	6	0.914	44	386	576	766	956	1,146
	Logit	41	5	0.911	46	415	620	825	1,030	1,235
	Logit-prop	14	76	0.825	90	216	286	356	426	496
	LVQ	10	2	0.977	12	102	152	202	252	302
	BP	4	34	0.926	38	74	94	114	134	154
1994	MDA	21	64	0.834	85	274	379	484	589	694
	MDA-prop	41	6	0.908	47	416	621	826	1,031	1,236
	Logit	42	5	0.908	47	425	635	845	1,055	1,265
	Logit-prop	16	75	0.822	91	235	315	395	475	555
	LVQ	12	2	0.973	14	122	182	242	302	362
	BP	3	11	0.973	14	41	56	71	86	101

Increasing the misclassification cost ratio tends to favor the methods which have a low type I error. While Zavgren (1983) suggested a type I to type II cost ratio of 20, researchers in previous studies have often had higher cost ratios. Hence, we focus on cost ratios of 20–30. For these cost ratios, the MDA with proportional prior probability and the logit with equal prior probability consistently yield a much higher percentage of type I errors for all data sets.

Prediction Efficiency in Terms of Resubstitution Risk

According to Fryman, Altman, and Kao (1985), resubstitution risk is defined as $RR = C_1 p_1 (n_1/N_1) + C_2 p_2 (n_2/N_2)$, where p_1 is the proportion of the bankrupt group in the population, p_2 the proportion of the nonbankrupt group in the population, C_1 the cost of misclassifying a bankrupt firm as a nonbankrupt firm, C_2 the cost of misclassifying a nonbankrupt firm as a bankrupt firm, n_1 the number of misclassified bankrupt firms, N_1 the number of bankrupt firms in the sample, n_2 the number of misclassified nonbankrupt firms, and N_2 the number of nonbankrupt firms in the sample.

Minimization of resubstitution risk was used as the criterion for evaluating the MDA, logit, LVQ, and BP methods. Resubstitution risk criterion is different from the misclassification cost criterion in considering the prior probabilities of being a troubled insurer, and hence may be considered as the expected misclassification cost. For each

TABLE 5
Comparison of the Predictive Accuracy of the Methods Using Misclassification Cost Criterion: FAST Variable Set and 1-Year Prior Model

Year	Method	No. of Type I Errors	No. of Type II Errors	Correct Rate	Total Misclassification Cost at Ratio C_1/C_2					
					1	10	15	20	25	30
1992	MDA	26	45	0.862	71	305	435	565	695	825
	MDA-prop	38	6	0.914	44	386	576	766	956	1,146
	Logit	48	6	0.895	54	486	726	966	1,206	1,446
	Logit-prop	22	77	0.807	99	297	407	517	627	737
	LVQ	21	3	0.953	24	213	318	423	528	633
	BP	3	17	0.961	20	47	62	77	92	107
1993	MDA	25	35	0.884	60	285	410	535	660	785
	MDA-prop	36	9	0.913	45	369	549	729	909	1,089
	Logit	42	7	0.905	49	427	637	847	1,057	1,267
	Logit-prop	5	99	0.799	104	149	174	199	224	249
	LVQ	20	1	0.959	21	201	301	401	501	601
	BP	4	19	0.956	23	59	79	99	119	139
1994	MDA	25	34	0.885	59	284	409	534	659	784
	MDA-prop	34	6	0.922	40	346	516	686	856	1,026
	Logit	45	5	0.902	50	455	680	905	1,130	1,355
	Logit-prop	21	73	0.816	94	283	388	493	598	703
	LVQ	19	6	0.951	25	196	291	386	481	576
	BP	4	11	0.971	15	51	71	91	111	131

of the five data sets and using both current year and 1-year prior models, resubstitution risks for each method are calculated over a wide range of cost ratios (1–30). The results on the 1-year prior model are not shown in this article because of the lack of a significant difference between the performance of the current year and 1-year prior models. Some of the results for the current year model are presented in Tables 8–12.

The results indicate that a consistent pattern of minimum resubstitution risk emerges for the 22-variable, IRIS, and FAST data sets. Assuming equal costs for type I and type II errors (cost ratio = 1), LVQ is the best for both the current year and 1-year prior models. For higher cost ratios, the back-propagation works best. For all data sets, the MDA and logit (equal and proportional priors) are almost always inferior to LVQ and BP methods. These results are consistent with those obtained using the misclassification cost criterion. The MDA with equal prior probability and the logit with equal prior probability give unacceptably high levels of resubstitution risks for all data sets. The increased resubstitution risk is due to the additional weights placed on type I errors by using equal prior probability.

EMPIRICAL RESULTS: PERFORMANCE COMPARISON AMONG DIFFERENT VARIABLE SETS

In the previous section, we evaluated all four prediction methods. Results were obtained that consistently favored the BP and LVQ methods in reducing expected

TABLE 6
Comparison of the Predictive Accuracy of the Methods Using Misclassification Cost Criterion: Texas EWIS-Automated Variable Set and 1-Year Prior Model

Year	Method	No. of Type I Errors	No. of Type II Errors	Correct Rate	Total Misclassification Cost at Ratio C_1/C_2					
					1	10	15	20	25	30
1992	MDA	5	23	0.945	28	73	98	123	148	173
	MDA-prop	9	10	0.963	19	100	145	190	235	280
	Logit	17	10	0.947	27	180	265	350	435	520
	Logit-prop	3	48	0.901	51	78	93	108	123	138
	LVQ	3	1	0.992	4	31	46	61	76	91
	BP	3	5	0.984	8	35	50	65	80	95
1993	MDA	5	20	0.951	25	70	95	120	145	170
	MDA-prop	5	8	0.975	13	58	83	108	133	158
	Logit	15	9	0.953	24	159	234	309	384	459
	Logit-prop	4	46	0.902	50	86	106	126	146	166
	LVQ	2	1	0.994	3	21	31	41	51	61
	BP	3	7	0.980	10	37	52	67	82	97
1994	MDA	2	14	0.969	16	34	44	54	64	74
	MDA-prop	4	8	0.976	12	48	68	88	108	128
	Logit	23	6	0.943	29	236	351	466	581	696
	Logit-prop	5	70	0.853	75	120	145	170	195	220
	LVQ	3	0	0.994	3	30	45	60	75	90
	BP	2	9	0.978	11	29	39	49	59	69

bankruptcy costs. Accordingly, we now evaluate the usefulness of each data set choice using BP and LVQ. Based on the results of the previous section, it can be suggested that both of the two evaluation criteria are essentially equivalent. So we will present results using the misclassification cost criterion.

Performance of BP Method Under Different Variable Sets

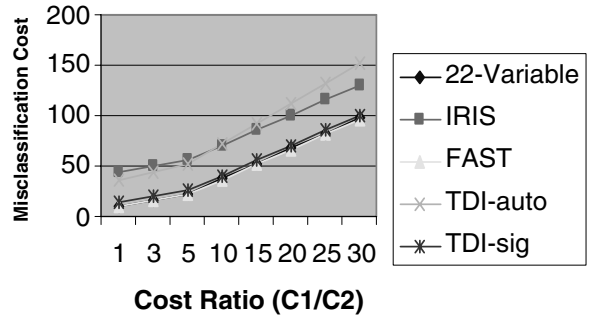
The performance of the BP method under different variable sets for the current year model is illustrated in Figures 1–3. With the back-propagation method and using total misclassification cost as the evaluation criterion, 22-variable and Texas EWIS-significant variables yield the best results for the current year model. FAST model is a close third. For one of the three years (1993), IRIS yields slightly better results than FAST, but only for the lowest two cost ratios (1–3).

For 1992 data, the FAST and 22-variable data sets give identical results and dominate over the higher cost ratio range. The TDI-significant data sets results are only slightly worse than these two. Using the IRIS data set yields higher misclassification costs than FAST due to the higher number of type I errors. For 1993 data, the TDI-significant clearly gives the best results. The 22-variable and TDI-automated data sets are distant seconds. For these data, results for IRIS and FAST data sets are almost indistinguishable and are the worst of the five data sets considered. Using

TABLE 7
Comparison of the Predictive Accuracy of the Methods Using Misclassification Cost Criterion: Texas EWIS-Significant Variable Set and 1-Year Prior Model

Year	Method	No. of Type I Errors	No. of Type II Errors	Correct Rate	Total Misclassification Cost at Ratio C_1/C_2					
					1	10	15	20	25	30
1992	MDA	8	4	0.977	12	84	124	164	204	244
	MDA-prop	8	1	0.982	9	81	121	161	201	241
	Logit	27	8	0.932	35	278	413	548	683	818
	Logit-prop	6	65	0.862	71	125	155	185	215	245
	LVQ	0	0	1.000	0	0	0	0	0	0
	BP	1	16	0.967	17	26	31	36	41	46
1993	MDA	6	0	0.988	6	60	90	120	150	180
	MDA-prop	6	0	0.988	6	60	90	120	150	180
	Logit	32	0	0.938	32	320	480	640	800	960
	Logit-prop	32	0	0.938	32	320	480	640	800	960
	LVQ	1	0	0.998	1	10	15	20	25	30
	BP	4	2	0.988	6	42	62	82	102	122
1994	MDA	4	0	0.992	4	40	60	80	100	120
	MDA-prop	4	0	0.992	4	40	60	80	100	120
	Logit	32	0	0.937	32	320	480	640	800	960
	Logit-prop	32	0	0.937	32	320	480	640	800	960
	LVQ	0	0	1.000	0	0	0	0	0	0
	BP	2	13	0.971	15	33	43	53	63	73

FIGURE 1
Performance of the BP Method Using Misclassification Cost Criterion: Current Year Model, 1992



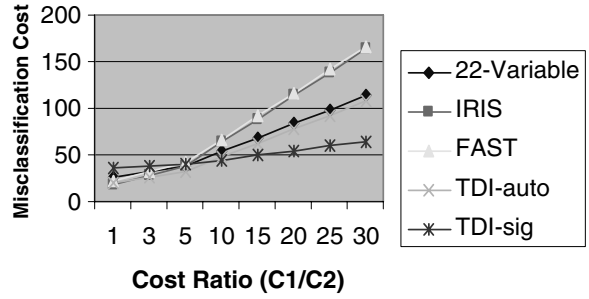
1994 data, the TDI-significant data set again dominates clearly. The 22-variable data set is a distant second. Using these data, IRIS results are somewhat better than FAST results.

Looking at all three years, it is difficult to see any consistent improvement in performance using FAST rather than IRIS. Both data sets generally perform worse than the TDI or the 22-variable sets.

TABLE 8
Comparison of the Predictive Accuracy of the Methods Using Resubstitution Cost Criterion: 22-Variable Set and Current Year Model

Year	Method	Type I Error Rate	Prior prob. of Type I	Type II Error Rate	Prior Prob. of Type II	Resubstitution Risk				
						1	5	10	20	30
1992	MDA	0.250	0.5	0.112	0.5	0.18	0.68	1.31	2.56	3.81
	MDA-prop	0.469	0.121	0.037	0.879	0.09	0.32	0.60	1.17	1.73
	Logit	0.610	0.5	0.035	0.5	0.32	1.54	3.07	6.12	9.17
	Logit-prop	0.170	0.121	0.143	0.879	0.15	0.23	0.33	0.54	0.74
	LVQ	0.063	0.121	0.017	0.879	0.02	0.05	0.09	0.17	0.24
	BP	0047	0.121	0.017	0.879	0.02	0.04	0.07	0.13	0.19
1993	MDA	0.291	0.5	0.145	0.5	0.22	0.80	1.53	2.98	4.44
	MDA-prop	0.546	0.106	0.022	0.894	0.08	0.31	0.60	1.18	1.76
	Logit	0.745	0.5	0.015	0.5	0.38	1.87	3.73	7.46	11.2
	Logit-prop	0.255	0.106	0.21	0.894	0.21	0.32	0.46	0.73	1.00
	LVQ	0.145	0.106	0.015	0.894	0.03	0.09	0.17	0.32	0.47
	BP	0.055	0.106	0.052	0.894	0.05	0.08	0.10	0.16	0.22
1994	MDA	0.367	0.5	0.112	0.5	0.24	0.97	1.89	3.73	5.56
	MDA-prop	0.612	0.096	0.026	0.904	0.08	0.32	0.61	1.20	1.79
	Logit	0.816	0.5	0.013	0.5	0.41	2.05	4.09	8.17	12.2
	Logit-prop	0.265	0.096	0.235	0.904	0.24	0.34	0.47	0.72	0.98
	LVQ	0.143	0.096	0.004	0.904	0.02	0.07	0.14	0.28	0.42
	BP	0.082	0.096	0.015	0.904	0.02	0.05	0.09	0.17	0.25

FIGURE 2
Performance of the BP Method Using Misclassification Cost Criterion: Current Year Model, 1993

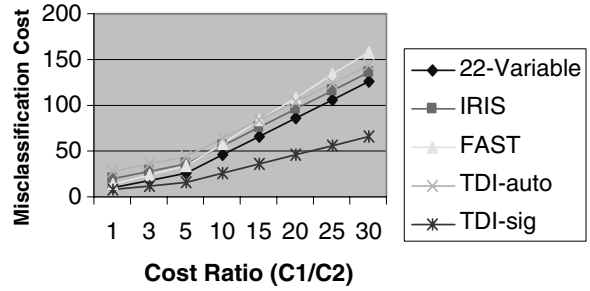


The performance of the BP method under different variable sets for the 1-year prior model is illustrated in Figures 4–6. For the 1-year prior model, both TDI and the 22-variable data sets yield the best results. The two TDI sets provide similar results, as is expected given the overlap of certain variables in these two sets. For 1992 data, the TDI-significant performs best, with the 22-variable set a reasonably close second. FAST and IRIS are the worst, with FAST slightly better. Using 1993 data, the 22-variable data set is the best and TDI-auto performs equally well. TDI-significant is a fairly close third.

TABLE 9
Comparison of the Predictive Accuracy of the Methods Using Resubstitution Cost Criterion: IRIS Variable Set and Current Year Model

Year	Method	Type I Error Rate	Prior prob. of Type I	Type II Error Rate	Prior Prob. of Type II	Resubstitution Risk				
						1	5	10	20	30
1992	MDA	0.392	0.5	0.130	0.5	0.26	1.05	2.03	3.99	5.95
	MDA-prop	0.843	0.099	0.011	0.901	0.09	0.43	0.84	1.68	2.51
	Logit	0.941	0.5	0.006	0.5	0.47	*2.36*	4.71	9.41	14.1
	Logit-prop	0.294	0.099	0.246	0.901	0.25	0.37	0.51	0.80	1.09
	LVQ	0.177	0.099	0.006	0.901	0.02	0,09	0.18	0.36	0.53
	BP	0.059	0.099	0.089	0.901	0.09	0.11	0.14	0.20	0.26
1993	MDA	0.333	0.5	0.138	0.5	0.24	0.90	1.73	3.40	5.06
	MDA-prop	0.686	0.099	0.032	0.901	0.10	0.37	0.71	1.39	2.07
	Logit	0804	0.5	0.017	0.5	0.41	2.02	4.03	8.05	12.1
	Logit-prop	0.294	0.099	0.171	0.901	0.18	0.30	0.45	0.74	1.03
	LVQ	0.177	0.099	0.004	0.901	0.02	0.09	0.18	0.35	0.53
	BP	0.098	0.099	0.030	0.901	0.04	0.08	0.12	0.22	0.32
1994	MDA	0.408	0.5	0.106	0.5	0.26	1.07	2.09	4.13	6.17
	MDA-prop	0.612	0.096	0.022	0.904	0.08	0.31	0.61	1.19	1.78
	Logit	0.800	0.5	0.011	0.5	0.41	2.01	4.01	8.01	12.0
	Logit-prop	0.388	0.096	0.173	0.904	0.19	0.34	0.53	0.90	1.27
	LVQ	0.225	0096	0.011	0.904	0.03	0.12	0.23	0.44	0.66
	BP	0.082	0.096	0.035	0.904	0.04	0.07	0.11	0.19	0.27

FIGURE 3
Performance of the BP Method Using Misclassification Cost Criterion: Current Year Model, 1994



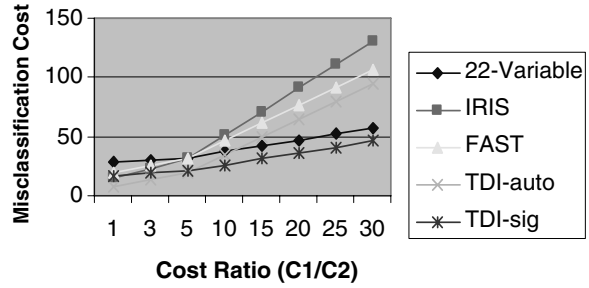
Once again, FAST and IRIS give the worst results, with FAST dominating slightly. For 1994 data, the 22-variable data set dominates. TDI-automated and TDI-significant give similar results and are second and third, respectively. In this year, IRIS performs better than FAST, due to a lower number of type I errors.

Overall, for the current year and 1-year prior models, the TDI and 22-variable data sets perform the best. The FAST and IRIS data sets are clearly inferior. Also, it is important

TABLE 10
Comparison of the Predictive Accuracy of the Methods Using Resubstitution Cost Criterion: FAST Variable Set and Current Model

Year	Method	Type I Error Rate	Prior prob. of Type I	Type II Error Rate	Prior Prob. of Type II	Resubstitution Risk				
						1	5	10	20	30
1992	MDA	0.380	0.5	0.082	0.5	0.23	0.99	1.94	3.84	5.74
	MDA-prop	0.560	0.097	0.024	0.903	0.08	0.29	0.56	1.11	1.65
	Logit	0.700	0.5	0.012	0.5	0.36	1.76	3.51	7.01	10.5
	Logit-prop	0.320	0.097	0.201	0.903	0.21	0.34	0.49	0.80	1.11
	LVQ	0.180	0.097	0.002	0.903	0.02	0.09	0.18	0.35	0.53
	BP	0.060	0.097	0.015	0.903	0.02	0.04	0.07	0.13	0.19
1993	MDA	0.389	0.5	0.127	0.5	0.26	1.04	2.01	3.95	5.90
	MDA-prop	0.648	0.104	0.015	0.896	0.08	0.35	0.69	1.36	2.04
	Logit	0.852	0.5	0.015	0.5	0.43	2.14	4.27	8.53	12.8
	Logit-prop	0.333	0.104	0.194	0.896	0.21	0.35	0.52	0.87	1.21
	LVQ	0.259	0.104	0.009	0.896	0.04	0.14	0.28	0.55	0.82
	BP	0.093	0.104	0.037	0.896	0.04	0.08	0.13	0.23	0.32
1994	MDA	0.396	0.5	0.125	0.5	0.26	1.05	2.04	4.02	6.00
	MDA-prop	0.646	0.094	0.011	0.906	0.07	0.31	0.62	1.22	1.83
	Logit	0.896	0.5	0.002	0.5	0.45	2.24	4.48	8.96	13.4
	Logit-prop	0.438	0.094	0.160	0.906	0.19	0.35	0.56	0.97	1.38
	LVQ	0.292	0.094	0.002	0.906	0.03	0.14	0.28	0.55	0.83
	BP	0.104	0.094	0.019	0.906	0.03	0.07	0.11	0.21	0.31

FIGURE 4
Performance of the BP Method Using Misclassification Cost Criterion: 1-Year Prior Model, 1992



to note that no consistent improvement in performance is obtained using FAST rather than IRIS. Results using the resubstitution risk criterion are not included since a similar evaluation of the data sets to the misclassification cost criterion is obtained.

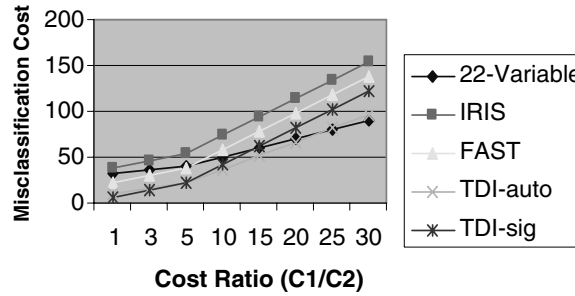
Performance of LVQ Method Under Different Variable Sets

The performance of the LVQ method under different variable sets for the current year model is illustrated in Figures 7–9. With the LVQ method and using the

TABLE 11
Comparison of the Predictive Accuracy of the Methods Using Resubstitution Cost Criterion: Texas EWIS-Automated Variable Set and Current Year Model

Year	Method	Type I Error Rate	Prior prob. of Type I	Type II Error Rate	Prior Prob. of Type II	Resubstitution Risk				
						1	5	10	20	30
1992	MDA	0.240	0.5	0.050	0.5	0.15	0.63	1.23	2.43	3.63
	MDA-prop	0.320	0.097	0.013	0.903	0.04	0.17	0.32	0.63	0.94
	Logit	0.580	0.5	0.024	0.5	0.30	1.46	2.91	5.81	8.71
	Logit-prop	0.180	0.097	0.138	0.903	0.14	0.21	0.30	0.47	0.65
	LVQ	0.080	0.097	0.002	0.903	0.01	0.04	0.08	0.16	0.23
	BP	0.080	0.097	0.069	0.903	0.07	0.10	0.14	0.22	0.30
1993	MDA	0.245	0.5	0.045	0.5	0.15	0.64	1.25	2.47	3.70
	MDA-prop	0.306	0.096	0.011	0.904	0.04	0.16	0.30	0.60	0.89
	Logit	0.490	0.5	0.026	0.5	0.26	1.24	2.46	4.91	7.36
	Logit-prop	0.204	0.096	0.114	0.904	0.12	0.20	0.30	0.49	0.69
	LVQ	0.082	0.096	0.006	0.904	0.01	0.04	0.08	0.16	0.24
	BP	0.061	0.096	0.039	0.904	0.04	0.06	0.09	0.15	0.21
1994	MDA	0.304	0.5	0.035	0.5	0.17	0.78	1.54	3.06	4.58
	MDA-prop	0.413	0.09	0.015	0.91	0.05	0.20	0.39	0.76	1.13
	Logit	0.630	0.5	0.017	0.5	0.32	1.58	3.16	6.31	9.46
	Logit-prop	0.261	0.09	0.149	0.91	0.16	0.25	0.37	0.61	0.84
	LVQ	0.087	0.09	0.004	0.91	0.01	0.04	0.08	0.16	0.24
	BP	0.087	0.09	0.052	0.91	0.06	0.09	0.13	0.20	0.28

FIGURE 5
Performance of the BP Method Using Misclassification Cost Criterion: 1-Year Prior Model, 1993



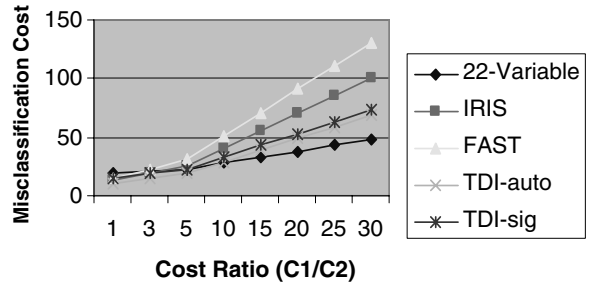
total misclassification cost criterion, TDI-significant and TDI-automated data sets consistently tend to minimize costs for the current year model. For each of the three years of data examined, the 22-variable data set is a close third. Once again, IRIS and FAST data sets produce the worst results, with IRIS dominating in two of the three years.

The performance of the LVQ method under different variable sets for the 1-year prior model is illustrated in Figures 10–12. For the 1-year prior model, we again observe

TABLE 12
Comparison of the Predictive Accuracy of the Methods Using Resubstitution Cost Criterion: Texas EWIS-Significant Variable Set and Current Year Model

Year	Method	Type I Error Rate	Prior prob. of Type I	Type II Error Rate	Prior Prob. of Type II	Resubstitution Risk				
						1	5	10	20	30
1992	MDA	0.22	0.5	0.017	0.5	0.12	0.56	1.11	2.21	3.31
	MDA-prop	0.30	0.097	0.004	0.903	0.03	0.15	0.29	0.59	0.88
	Logit	0.52	0.5	0.022	0.5	0.27	1.31	2.61	5.21	7.81
	Logit-prop	0.16	0.097	0.121	0.903	0.12	0.19	0.26	0.42	0.57
	LVQ	0.04	0.097	0.004	0.903	0.01	0.02	0.04	0.08	0.12
	BP	0.06	0.097	0.024	0.903	0.03	0.05	0.08	0.14	0.20
1993	MDA	0.225	0.5	0.022	0.5	0.12	0.57	1.14	2.26	3.39
	MDA-prop	0.286	0.096	0.013	0.904	0.04	0.15	0.29	0.56	0.84
	Logit	0.755	0.5	0.019	0.5	0.39	1.90	3.78	7.55	11.3
	Logit-prop	0.245	0.096	0.145	0.904	0.15	0.25	0.37	0.60	0.84
	LVQ	0.061	0.096	0	0.904	0.01	0.03	0.06	0.12	0.18
	BP	0.020	0.096	0.076	0.904	0.07	0.08	0.09	0.11	0.13
1994	MDA	0.217	0.5	0.013	0.5	0.12	0.55	1.09	2.18	3.26
	MDA-prop	0.283	0.09	0.009	0.91	0.03	0.14	0.26	0.52	0.77
	Logit	0.804	0.5	0.011	0.5	0.41	2.02	4.03	8.05	12.1
	Logit-prop	0.391	0.09	0.188	0.91	0.21	0.35	0.52	0.87	1.23
	LVQ	0.13	0.09	0	0.91	0.01	0.06	0.12	0.23	0.35
	BP	0.043	0.09	0.013	0.91	0.02	0.03	0.05	0.09	0.13

FIGURE 6
Performance of the BP Method Using Misclassification Cost Criterion: 1-Year Prior Model, 1994



that the TDI-significant data sets performed best. TDI-automated comes in second and the 22-variable data set is a somewhat distant third. It is important to note that the results for both the current year and 1-year prior models are invariant to the level of cost ratios. Like the current year model, we observe consistently better performance from the IRIS data set than the FAST data set in each of the three years.

Overall, for both the current year and 1-year prior models, the TDI and 22-variable data sets give the best performance using misclassification cost criterion. It is interesting

FIGURE 7
Performance of the LVQ Method Using Misclassification Cost Criterion: Current Year Model, 1992

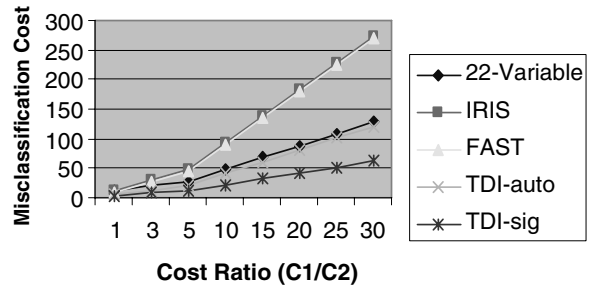


FIGURE 8
Performance of the LVQ Method Using Misclassification Cost Criterion: Current Year Model, 1993

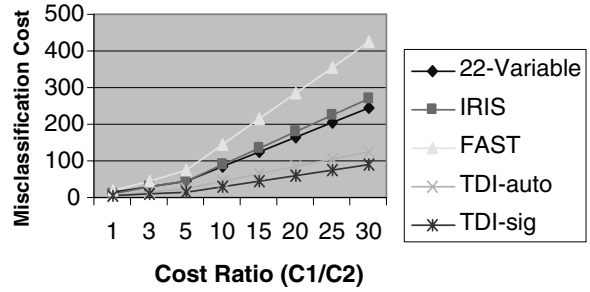
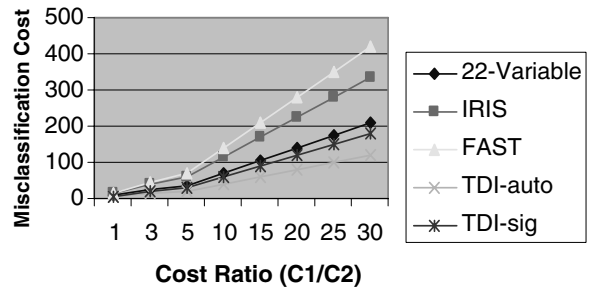


FIGURE 9
Performance of the LVQ Method Using Misclassification Cost Criterion: Current Year Model, 1994



to note the superiority of the IRIS data set over FAST, which was designed as an improvement to IRIS. Results using the resubstitution risk criterion are not included since a similar evaluation of the data sets is obtained.

FIGURE 10
Performance of the LVQ Method Using Misclassification Cost Criterion: 1-Year Prior Model, 1992

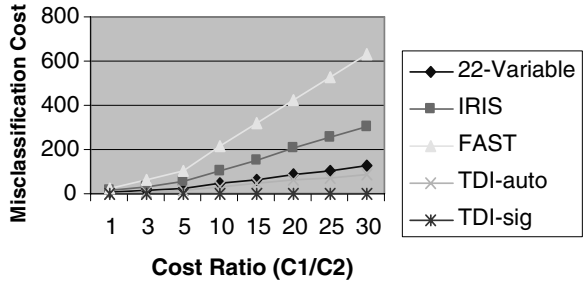


FIGURE 11
Performance of the LVQ Method Using Misclassification Cost Criterion: 1-Year Prior Model, 1993

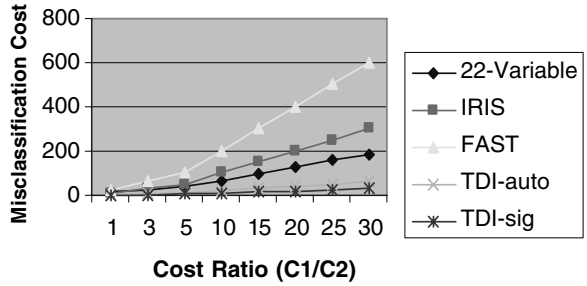
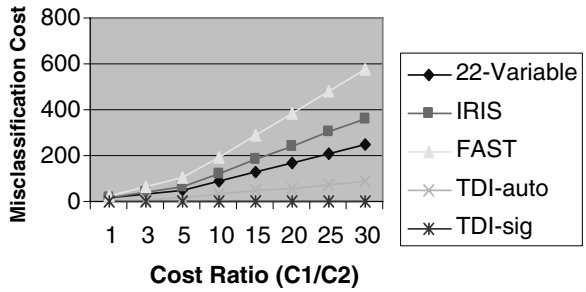


FIGURE 12
Performance of the LVQ Method Using Misclassification Cost Criterion: 1-Year Prior Model, 1994



CONCLUSION

This study employs and compares two statistical methods and two artificial neural network methods for prediction of financial hazard in life insurers. It also investigates the usefulness of the IRIS, FAST, and Texas EWIS variable sets for the two neural network methods.

This study shows that the neural network models back-propagation (BP) and LVQ outperform the traditional statistical approaches for all four data sets with a consistent superiority across the two different evaluation criteria, total misclassification cost, and resubstitution risk criteria. BP is the most efficient and stable predictive tool in the 22-variable model, IRIS, and FAST data sets whereas LVQ is the most efficient in Texas EWIS data set with accuracy as high as 100 percent correct rate. Neural network models hence show promise for early warning systems.

Finally, the overall performance of the neural network methods appear robust over time. Regulators can use these models for more than one year, allowing rapid identification of potentially troubled companies without needing to constantly update the prediction model. The out-of-sample test used in this study is practically useful in the sense that we can use the *ex post* information of the past to predict the *ex ante* outcomes for the current year.

Further work on insurer insolvency prediction could attempt to incorporate other significant variables. Browne, Carson, and Hoyt (1999) indicated that life-health insurer insolvencies are positively related to increases in long-term interest rates, personal income, unemployment, the stock market, and to the number of insurers, and negatively related to real estate returns. Business environment changes every year, the individual firm's financial ratios should be combined with macroeconomic variables obtained externally which represent business environment changes.

REFERENCES

Altman, E. I., 1968, Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *Journal of Finance*, 23: 589-609.

A.M. Best Company, 1992, *Best's Insolvency Study, Life-Health Insurers* (A.M. Best Company, Oldwick, NJ).

Baranoff, Etti G., Thomas W. Sager, and Robert C. Witt, 1999, Industry Segmentation and Predictor Motifs for Solvency Analysis of the Life/Health Insurance Industry, *Journal of Risk and Insurance*, 66(1): 99-123.

Baranoff, Etti G., Thomas W. Sager, and Thomas S. Shively, 2000, A Semiparametric Stochastic Spline Model as a Managerial Tool for Potential Insolvency, *Journal of Risk and Insurance*, 67(3): 369-396.

BarNiv, R., and J. B. McDonald, 1992, Identifying Financial Distress in the Insurance Industry: A Synthesis of Methodological and Empirical Issues, *Journal of Risk and Insurance*, 59: 543-573.

BarNiv, R., and R. A. Hershbarger, 1990, Classifying Financial Distress in the Life Insurance Industry, *Journal of Risk and Insurance*, 57: 110-136.

Barrese, James, 1990, Assessing the Financial Condition of Insurers, *CPCU Journal*, 43(1): 37-46.

Brockett, Patrick L., Richard Derrig, and Xiaohua Xia, 1998, Using Khonen's Self Organizing Feature Map to Uncover Automobile Bodily Injury Claim Fraud, *Journal of Risk and Insurance*, 65(2): 245-274.

Brockett, P. L., W. W. Cooper, L. L. Golden, and U. Pitaktong, 1994, A Neural Network Method for Obtaining and Early Warning of Insurer Insolvency, *Journal of Risk and Insurance*, 61: 402-424.

Browne, M. J., J. M. Carson, and R. E. Hoyt, 1999, Economic and Market Predictors of Insolvencies in the Life-Health Insurance Industry, *Journal of Risk and Insurance*, 66(4): 643-659.

Bryson, A. E., and Y. C. Ho, 1969, *Applied Optimal Control* (Blaisdell, New York).

Carson, J. M., and Robert E. Hoyt, 1995, Life Insurer Financial Distress: Classification Models and Empirical Evidence, *Journal of Risk and Insurance*, 62: 764-775.

Carson, J. M., and R. E. Hoyt, 2000, Evaluating the Risk of Life Insurer Insolvency: Implications from the US for the European Union, *Journal of Multinational Financial Management*, 10: 297-314.

Coats, P. K., and L. F. Fant, 1993, Recognizing Financial Distress Patterns Using a Neural Network Tool, *Financial Management*, 22

Cummins, J. David, Scott E. Harrington, and Robert Klein, 1995, Insolvency Experience, Risk-Based Capital, and Prompt Corrective Action in Property-Liability Insurance, *Journal of Banking and Finance*, 19: 511-527.

Cummins, J. D., M. E. Grace, and R. D. Phillips, 1999, Regulatory Solvency Prediction in Property-Liability Insurance: Risk-Based Capital, Audit Ratios, and Cash Flow Simulation, *Journal of Risk and Insurance*, 66(3): 417-458.

Edmister, R. O., 1972, An Empirical Test of Financial Ratio Analysis for Small Business Failure Predictions, *Journal of Financial and Quantitative Analysis*, 7: 1477-1493.

Fryman, H., E. I. Altman, and D. L. Kao, 1985, Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress, *Journal of Finance*, 40: 269-291.

Grace, Martin, Scott Harrington, and Robert Klein, 1998, Risk-Based Capital and Solvency Screening in Property-Liability Insurance: Hypotheses and Empirical Tests, *Journal of Risk and Insurance*, 65(2): 213-243.

Huang, C. S., R. E. Dorsey, and M. A. Boose, 1994, Life Insurer Financial Distress Prediction: A Neural Network Model, *Journal of Insurance Regulation*, 13: 133-167.

Kohonen, T. G. Barna, and R. Chrisley, 1988, Statistical Pattern Recognition with Neural Networks: Benchmark Studies, *Proceedings of the Second Annual IEEE International Conference on Neural Networks*, 1: 61-68.

Kohonen, T., 1998, Learning Vector Quantization, in M. A. Arbib, ed., *The Handbook of Brain Theory and Neural Networks* (MIT Press, Cambridge, MA, 537-540).

Kohonen, T., J. Hynninen, J. Kangas, J. Laakosonen, and K. Torkkola, 1996, *LVQ_PAK: The Learning Vector Quantization Program Package* (Helsinki University of Technology, Laboratory of Computer and Information Science Technical Report A30).

Luther, R. K., 1993, Predicting the Outcome of Chapter 11 Bankruptcy: An Artificial Neural Network Approach, *Ph.D. Dissertation*, University of Mississippi.

Palepu, K. G., 1986, Predicting Takeover Targets: A Methodological and Empirical Analysis, *Journal of Accounting and Economics*, 8: 3-35.

Parker, D. B., 1985, Learning Logic, Technical Report, TR-47, *Center for Computational Research in Economics and Management Science* (MIT, Cambridge, MA).

Pinches, G. E., and J. S. Trieschmann, 1974, The Efficiency of Alternative Models for Solvency Surveillance in the Insurance Industry, *Journal of Risk and Insurance*, 41: 563-577.

Rumelhart, D. E., G. E. Hinton, and R. J. Williams, 1986, Learning Representations by Back-Propagating Errors, *Nature*, 323(9): 523-536.

Salchenberger, L. M., E. M. Cinar, and N. A. Lash, 1992, Neural Networks: A New Tool for Predicting Thrift Failures, *Decision Science*, 23(4): 899-915.

Sinkey, J. F., Jr., 1975, Multivariate Statistical Analysis of the Characteristics of Problem Banks, *Journal of Finance*, 30: 21-36.

Trieschmann, J. S., and G. E. Pinches, 1973, A Multivariate Model for Predicting Financially Distressed Property-Liability Insurance Companies, *Journal of Risk and Insurance*, 40: 327-338.

Viaene, Stijn, Richard A. Derrig, Bart Baesens, and Guido Dedene, 2002, A Comparison of State of the Art Classification Techniques for Expert Automobile Insurance Claim Fraud Detection, *Journal of Risk and Insurance*, 69(3): 373-421.

Werbos, P., 1974, Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences, Ph.D. Dissertation, Harvard University.

Zavgren, C. V., 1983, The Prediction of Corporate Failure: The State of the Art, *Journal of Accounting Literature*, 2: 1-38.

Zavgren, C. V., 1985, Assessing the Vulnerability of Failure of American Industrial Firms: A Logistic Analysis, *Journal of Business Finance and Accounting*, 12: 19-45.