AN IMPROVED CREDIT SCORING FUNCTION FOR THE ST. PAUL BANK FOR COOPERATIVES

Martin L. Fischer & Karen Moore

Proceedings of
Regional Research Committee NC-161

FINANCING AGRICULTURE IN A CHANGING ENVIRONMENT: MACRO, MARKET, POLICY AND MANAGEMENT ISSUES

St. Paul, Minnesota
October 7-8, 1986

Department of Agribusiness Economics
Southern Illinois University at Carbondale
Carbondale, IL 62901

January 1987

AN IMPROVED CREDIT SCORING FUNCTION FOR THE ST. PAUL BANK FOR COOPERATIVES

Martin L. Fischer and Karen Moore

Introduction

Prior to 1986, a credit scoring function estimated by discriminant analysis contributed 40 percent to the loan classifying process at the St. Paul Bank for Cooperatives (BC). Loan classifications are required of all Farm Credit System banks and associations under regulations of the Farm Credit Administration (FCA). Loan classifications permit FCA to monitor the condition of Farm Credit System lenders, and are used internally by the BC for monitoring and controlling its $1.9 billion loan portfolio.

Besides contributing to the loan classification system, a credit scoring function has potential to be used for screening loan applications, diagnosing credit weaknesses, and as a criteria for pricing loans based on credit quality. For these purposes the function must be accurate, objective, and as simple as possible given the requirement of statistical validity. Accuracy is essential if the function is to contribute to sound lending decisions and valid loan classification.

The authors are respectively Senior Research Analyst, Farm Credit Services, St. Paul, and Graduate Assistant, Department of Statistics, Iowa State University.

Objectivity is desirable if the function is to be used in loan pricing.[1] Reliance on subjective factors in loan pricing has undesirable implications for customer relations and possible adverse legal consequences. Simplicity is desirable if loan officers are to be able to compute the credit score prior to the lending decision, for screening loan applications and diagnosing credit weaknesses.

The credit scoring function used prior to 1986, -- henceforth, the old function -- was no longer very accurate, in part because it was estimated five years ago and structural changes have occurred which invalidate the model. Also, the methodology used to estimate the old function -- discriminant analysis -- is known in light of recent developments in statistics to be inappropriate. Finally, the old function had 29 explanatory variables, and analysts considered it too complex for diagnosing credit weaknesses and screening loan applications. For purposes of loan pricing, actual loan classifications may be inappropriate because subjective factors are weighted 60 percent in the loan classification system. These considerations suggested the need for a new credit scoring function to replace the old function in the loan classification system, and to use directly as a basis for screening loan applications, diagnosing credit weaknesses, and as a possible criteria for loan pricing based on credit quality. A new function estimated by logistic regression analysis (logit) is reported in this paper. Compared with the old function, the new function is both simpler and more accurate.

In the following sections, we review related literature, describe the BC's loan classification system, and report and evaluate two logit credit scoring functions. The first function is based on a dichotomous (2-response) dependent variable — problem or worse loans versus acceptable weak or stronger loans. The second function uses a polychotomous (4-response) dependent variable — vulnerable or loss loans, problem loans, acceptable weak loans, and acceptable loans. The dichotomous model is nearly as accurate as the polychotomous model, even when applied to classifying cooperatives into four classes. Because the dichotomous model is also much simpler to use and understand, it is preferred to the polychotomous model.

Related Studies

Statistical methods used in credit scoring include discriminant analysis, probit, and logit. A readable exposition of the logit and probit models is found in Pindyck and Rubinfeld (Ch. 10). For comparisons between discriminant analysis and logit, see Press and Wilson, Collins and Green, and Harrell.

No previously published studies report the application of statistical methods to evaluating credit quality of agricultural cooperatives. However, several studies report applications involving farm loans. Johnson and Hagen, Hardy and Weed, and Dunn and Frey all used discriminant analysis on a dichotomous dependent variable — problem versus acceptable loans. Sample size ranged from 99 to 378 loans in these studies, and accuracy ranged from 62 to 81 percent correctly classified. Lufburrow, Barry and Dixon used probit analysis on a instability. Computationally, the NIM-beta could be estimated as,

$$m_{jt} = \alpha_j + \beta_j m_{pt} + e_{jt} \tag{4}$$

where m_{jt} is the "excess net interest margin" of the jth bank in period t, m_{pt} is the excess net interest margin for a market portfolio consisting of all banks, α_j and β_j are the estimated parameters for the jth bank, and e_t is the error term for period t.

The excess net interest margin for the j^{th} individual bank (m_{jt}) is defined here as the deviation of the net interest margin series from the estimated trend series of net interest margin for the portfolio of all banks.[2] Estimation of the trend in the net interest margin series for the portfolio of banks assumes that in equilibrium bank management attempts to generate net interest earnings which are growing over time in a linear fashion. The implied linear model, $NIM_t = a + bT + u_t$ (where T is serial time), yields an estimate of the equilibrium expected net interest margin over time, $E(NIM)_t = a + bT$. The excess net interest margin for the j^{th} bank in the t^{th} period is, $m_{jt} = NIM_{jt} - E(NIM)_t$. The excess net interest margin for the portfolio of all banks is defined as the estimated residual from the linear model, $m_{pt} = u_t$. When the two excess net interest margin series are defined in this way, the resulting NIM-beta estimates may be positive or negative, depending on the covariance between the detrended net interest margin series for each bank and that for the portfolio of banks.

Unlike the beta of a stock which expresses the ex post systematic risk an investor would have assumed by holding the asset, the NIM-beta has a different interpretation. A large positive, significant beta indicates that the bank's net interest margin demonstrated relatively large systematic covariation when compared to all banks in the portfolio. That is, for a given deviation of the average NIM for all banks from trend, the bank with a large beta tends to deviate proportionately more in that same direction. Therefore, the larger is beta, the greater is the systematic risk exhibited by the bank's interest margin.[3]

---

[2] It is important to note that these portfolio trend values serve as a proxy for the risk-free rate (R_{ft}) in the CAPM framework.

[3] This need not translate into a large interest margin variance, however, since only the systematic component of variability has been measured using the beta approach.

The estimated NIM-beta serves as one indicator of asset/liability management performance. A large positive beta would be associated with relatively greater asset sensitivity, and may be indicative of aggressive asset/liability management and a strategy of accepting interest rate risk. Conversely, a positive beta could also indicate a continuing inability of management to adjust the bank's portfolio of assets and liabilities to effectively reduce interest rate risk exposure. Difficulty in interpreting the source of interest margin variability could be reduced by decomposing the NIM-beta into a gross interest income beta and a gross interest expense beta. The magnitudes and signs of these component betas could then be used to interpret systematic deviations in terms of underlying management strategies.[4]

In addition to the estimated beta, the estimated alpha coefficient in Equation 4 provides information on the excess net interest margin of each bank assuming no systematic instability (i.e., $\beta = 0$). A positive and significant α-coefficient could be interpreted as a measure of the ability of bank management to consistently outperform the reference portfolio of all banks over the period being analyzed. The expectation is that $\alpha = 0$, but individual α_j values may be positive or negative and significantly different from zero.

Estimation and Results

Estimates of the alpha and beta coefficients were derived from commercial banks located in Minnesota using end-of-year Call Reports (Reports of Condition and Reports of Income and Dividends) for 1976-85. A total of 788 Minnesota banks filed reports at the end of 1985. Banks with less than 5 years of reports (out of 10 years possible) were deleted. Banks with over $100 million in total assets (according to the December 1985 report) were also deleted to focus the analysis on small banks in the state. A total of 86 banks were deleted using these criteria, leaving 702 small banks for analysis.[5] Agricultural banks were defined as those which reported an agricultural loans and leases/total loans and leases ratio greater than .1615

---

[4] Copeland and Weston suggest the use of this beta decomposition approach to compare risky cost structures when applying capital budgeting methods under uncertainty.

[5] Out of the 702 small banks remaining, 53 banks had from 5 to 9 years of reports and were retained. All 53 banks were nonag banks. Small banks with 10 years of complete data totaled 649 of which 433 were ag banks and 216 were nonag banks.

polychotomous (3-response) dependent variable -- low risk, intermediate risk, and high risk loans. Their model correctly classified 71 percent of loans in the 3-class system.

Credit scoring functions estimated by statistical methods are widely used by corporations which issue credit cards (The Fair, Isaacs Companies). Discriminant analysis has been used in assessing risk of bankruptcy (Altman, et.al.). A variety of applications of discriminant analysis, probit, and logit in economics, social sciences and medicine are cited in Amemiya.

Strong arguments favoring logit over discriminant analysis are advanced by Press and Wilson, Collins and Green, and Harrell. Discriminant analysis assumes multivariate normality of explanatory variables with identical covariance matrices for each state of the dependent variable. Logit is much more robust with respect to assumptions concerning underlying probability distributions. When normality is violated, discriminant analysis estimators are neither consistent nor asymptotically efficient.[2] It is unlikely that the financial ratios used in most credit scoring studies are multivariate normal. Therefore, discriminant analysis is inappropriate for these applications. In contrast, logit models estimated by maximum likelihood methods yield consistent, asymptotically efficient estimators. Comparing logit with probit, Amemiya reports that the logistic probability distribution is similar to the normal distribution (on which probit is based) so that in dichotomous univariate applications, the choice can be based on convenience. We used logistic regression because of its superiority over discriminant analysis and because an excellent computer program for estimating logit models using maximum likelihood methods was available in the SAS software system (Harrell).

The BC Loan Classification System

Five loan classifications are used by the BC:

Acceptable (A):	Highest quality loans with minimal credit risk.
Acceptable Weak (A-):	Acceptable loans with some significant credit weaknesses.
Problem (P):	Loans having serious credit weaknesses requiring more than normal supervision, but believed to be collectible in full.
Vulnerable (V):	High risk loans still considered collectible, but involving probability of loss in the event repayment from available sources does not materialize.
Loss (L):	Loans on which all or any portion is deemed uncollectible.

In classifying loans, five elements are considered. The first element is the value of the credit scoring function.[3] This score is rescaled to an integer score (1 = acceptable, 2 = acceptable weak, ... 5 = loss), and weighted (multiplied) by .4. The four remaining elements are loan officers' subjective appraisals of management, collateral, financial reporting, and the competitive environment of the borrower. For each element an integer score is assigned (1 = acceptable ... 5 = loss) by the loan officer, and the integer scores are weighted respectively by: .2, .2, .1 and .1. The products of integer scores and weights are summed to determine the classification. In some instances loan officers report that the old credit score had to be overriden by manipulating the subjective elements so that the "proper" classification resulted. As one loan officer stated, the credit score should not require correction by subjective elements; rather, subjective elements should generally reinforce the credit score, and assist in determining the proper classification in cases where the score is a borderline value.

Despite reservations about the complexity and accuracy of the old credit scoring function, the overall loan classification system is highly regarded by loan officers, management, and FCA regulators. These experts agree that the resulting loan classifications are generally "correct", and that loan classifications are a valid indicator of credit quality. Management's request of the authors was to develop a simpler and more accurate credit scoring function to replace the old function in the loan classification system. The loan classification system itself would be retained, but the new credit scoring function would replace the old function.

Because the new function predicts actual loan classifications more accurately than the old function, there will be less need to override the credit score in arriving at the loan classification, and the importance of subjective elements in determining the loan classification will be reduced. Because the new function is also less complex, it is more useful as a tool for screening loan applications and diagnosing credit weaknesses.

Logit Models

In describing the estimating methodology and results, we shall first report on the dichotomous model, then the polychotomous model, and then compare the two.

Given agreement among the experts that the overall loan classifications are valid, and lacking any better measure of credit quality, actual loan classifications were used to define the dependent variables for our models. The accuracy of our models was judged primarily on their ability to predict actual loan classifications.

Dichotomous Model

The dependent variable in the dichotomous model was defined as follows:

$$Y = \begin{cases} 1 & \text{if actual loan classification is acceptable weak or acceptable.} \\ 0 & \text{if actual loan classification is problem, vulnerable or loss.} \end{cases}$$

Letting P denote the probability that Y = 1, the assumption of the dichotomous logit model is:

$$P = \frac{1}{1 + e^{-[A + B_1 X_1 + \ldots + B_n X_n]}}$$

where e = base of natural logarithm system, A = intercept, $B_1 \ldots B_n$ = coefficients, and $X_1 \ldots X_n$ are explanatory variables.

The estimating sample contained 448 cooperatives. Thirty-five variables were tested for statistical significance in the estimating procedure. Included were measures of liquidity, solvency, operating efficiency, coverage, profitability, and earnings stability, as well as dummy variables for different types of cooperatives. Criteria for inclusion of variables were (1) the variable had to be significant at the .85 probability level, and (2) the sign of the estimated coefficients had to be consistent with prior expectations. Several statistically acceptable models were reported to BC loan analysts, who then chose the "best" model based on statistical criteria and usefulness in diagnosing credit weaknesses. The variables included in the model chosen by loan analysts were:

ROABAR = Mean rate of return on local assets, where the return on local assets (ROA) = local earnings before interest/local assets. Local assets = total assets - investments in other cooperatives. ROA is a measure of the efficiency with which the cooperative employs its local assets. ROABAR is the average of ROA, computed over 2-4 years, depending on data availability.

DETAS = Total debt/local assets, a measure of solvency.

PROD = (Gross margin + operating revenue)/operating expense, a measure of operating efficiency.

DUM 3 = A dummy variable for type 3 (grain) cooperatives.

It is of interest that the mean value of ROA — ROABAR — proved more significant than either the most recent ROA value or a distributed lag specification for ROA. Use of ROABAR instead of ROA dampens the effect of year-to-year fluctuations in ROA on the score.

It is also of interest that the "best" model from a purely statistical viewpoint included the interest expense/sales ratio in place of PROD, but was rejected by loan officers for the following reasons: (1) If the model with interest expense/sales was used to differentially price loans, a borrower could argue that his score was too low to receive a preferential rate because the BC's interest rate was too high, and hence, his interest expense/sales ratio was too high; (2) loan officers preferred the "productivity ratio" PROD because it is a general indicator of operating efficiency, and can be used to counsel borrowers to control all categories of operating expense — not just interest expense.

Results for the dichotomous model are reported in Table 1. The estimated coefficients, standard errors, and chi-square statistics for individual coefficients are shown, as well as several statistics useful for evaluating overall performance. Each explanatory variable is significant at the .85 probability level, and when these variables are included, no other variables are significant at the .85 level. The R statistic (R = .851) is a measure of the predictive ability of the model and is one criteria for choosing the "best" model out of several competing models [Harrell, P.183]. The fraction of concordant pairs — another measure of predictive ability — is the fraction of cooperatives having different true classifications for which the predicted probability (score) is larger for the cooperative whose true classification is higher [Harrell, P. 184]. When cooperatives are classified into two groups (problem, vulnerable, or loss) versus (acceptable weak or acceptable), the model correctly classified 93.8 percent of the cooperatives.

The dichotomous model was also used to classify cooperatives into four classes. By examining the predicted probabilities ($\hat{P}$) from the model, it was determined that a critical score of $\hat{P} = 0.99$ would minimize misclassifications at the borderline between acceptable weak and acceptable, and a critical score of $\hat{P} = .005$ would minimize misclassifications at the borderline between problem and vulnerable or loss.[4] Thus, the dichotomous model was used to predict loan classifications in a 4-class system according to the following formula:

$$\text{Predicted Classification} = \begin{cases} \text{Acceptable if } \hat{P} \geq 0.99 \\ \text{Acceptable weak if } 0.99 > \hat{P} \geq 0.5 \\ \text{Problem if } 0.5 > \hat{P} \geq .005 \\ \text{Vulnerable or loss if } .005 \geq \hat{P} \end{cases}$$

When used in this way to classify cooperatives into four classes -- acceptable, acceptable weak, problem, or (vulnerable or loss) -- the dichotomous model correctly classifies 73.9 percent of cooperatives in the estimating sample.

Polychotomous Model

The dependent variable in the polychotomous application is an ordinarily ranked variable defined as:

$$Y = \begin{cases} 3 \text{ if classification is acceptable.} \\ 2 \text{ if classification is acceptable weak.} \\ 1 \text{ if classification is problem.} \\ 0 \text{ if classification is vulnerable or loss.} \end{cases}$$

In the 4-response application, the assumption of logit is:

$$P_j = \frac{1}{1 + e^{-[A_j + B_1 X_1 + \ldots + B_n X_n]}}$$

where P_j is the probability that $Y \geq j$, $j = 1 \ldots 3$ [Harrell]. This model yields three intercepts and three probabilities, $P_1 \ldots P_3$: P_1 = the probability that the cooperative is problem or stronger, P_2 = the probability that the cooperative is acceptable weak or stronger, and P_3 = the probability that the cooperative is acceptable. The cooperative would then be classified according to these probabilities.

In stepwise estimation of the polychotomous model, ROABAR and DETAS were included, and five new variables met the criteria for inclusion:

INTEXS = Interest expense/sales.

OENA = Owners equity/net assets, where net assets = total assets - current liabilities.

TLEV = Term debt/local net worth, where local net worth = owners equity - investments in other cooperatives.

LNWNLA = Local net worth/net local assets, where net local assets = local assets - current liabilities.

SDROA = Standard deviation of ROA.

Results of the polychotomous model are reported in Table 2. The variables PROD and DUM 3, which were included in the dichotomous model, did not meet the inclusion criteria in the polychotomous model.

OENA, TLEV and LNWLNA are measures of solvency, while SDROA measures variability of earnings before interest.

Inclusion of SDROA in the polychotomous model caused the number of observations to decline from 448 to 438, because there were ten observations with only one year of data, and at least two years are required to compute SDROA. Although LNWLNA is significant at only the .89 probability level, all other variables in Table 2 are significant at the .97 level or higher.

Note that by three measures of predictive ability — R, fraction of concordant pairs, and percent correctly classified in a 2-class system — the dichotomous model performs better than the polychotomous model. The polychotomous model is slightly better in classifying 4-ways, attaining 74.4 percent correct versus 73.9 percent correctly classified by the dichotomous model (4-class system). The dichotomous model is preferable because it is simpler to use, nearly as accurate as the polychotomous model in classifying cooperatives 4-ways, and more accurate in classifying 2-ways.

Validation of Dichotomous Model

The dichotomous model was validated on a test sample containing 260 cooperatives which were not included in the estimating sample. Errors were examined by type of cooperative to determine for which types the model is valid. Table 3 reports number of observations and errors by type of cooperative in the combined estimating and test samples. For purposes of classifying cooperatives in a 2-class system, the model is valid for all types shown in Table 3. In classifying 4-ways, the model has difficulty with fruit and vegetable, sugar, service, and transport associations. The model is biased against sugar cooperatives, i.e., tends to predict a lower classification than the current true classification. There is no bias toward over or underprediction for other types.

There are 55 errors in the 2-way classification, resulting in a 92.2 percent correct classification rate. This compares with an 85.5 percent correct rate for the old 29-variable credit scoring function applied to a 2-class system with the same data.[5] Of the 55 errors, 27 were cases where either the value of the predicted score was a borderline value or there was a discrepancy between the internal BC classification and the classification assigned by the Farm Credit Administration.[6]

Application of the dichotomous model to a 4-class system is reported in Table 4, which shows a cross-tabulation of actual classifications versus predicted classifications for all cooperatives in the combined sample. Also shown are intervals of the new score corresponding to each predicted classification. Elements on the diagonal of Table 4 represent correct classifications, off-diagonal elements are misclassified. The percent correctly classified in the combined sample 4-class system is 70. This compares with a 63 percent correct classification rate for the old function.

Conclusions

An improved credit scoring function for the St. Paul Bank for Cooperatives was estimated by logistic regression. The new function replaced an earlier function in the BC's loan classification system. This should improve the classification system by reducing the importance of subjective factors in determining classifications. The new function may also be used for loan pricing based on credit quality, screening loan applications, and diagnosing credit weaknesses. Similar models could presumably be developed for BCs in other districts of the Farm Credit System. Subject to further validation, the model reported here might even be adopted by other districts. Similar models are being developed for Production Credit Associations in the St. Paul Farm Credit District.

The validity of the model for different types of cooperatives was tested, and the model was judged valid for all types when used to distinguish problem or worse from acceptable weak or stronger loans. For purposes of classifying cooperatives 4-ways, the model is valid for all but four types. In a combined estimating and test sample of 708 cooperatives, the model classified 92 percent correctly in a 2-class system and 70 percent correctly in a 4-class system. These are substantially better correct classification rates than were achieved by the old credit scoring function, or by studies cited in the literature review which used disciminant analysis.

End Notes

[1]Differential pricing of loans based on credit quality is not currently practiced by the BC, but is being considered (along with cost of servicing) as a possible basis for differential pricing.

[2]An estimator of a parameter is consistent if the probability that the estimator differs from the parameter approaches zero as sample size approaches infinity. An estimator is asymptotically efficient if, for a given sample size, the asymptotic variance of the estimator is smaller than that of any other consistent estimator. Lack of consistency and asymptotic efficiency for discriminant analysis estimators (under nonnormality) is a strong indictment of this methodology.

[3]Prior to January 1986, the old function was used. Since January 1986, the new function has been used.

[4]Although vulnerable and loss are distinct classes, we treated them as one class for the following reasons: (a) The chief distinction between vulnerable and loss lies in collateral, and an objective collateral variable was not available for estimating purposes, and (b) there were too few cooperatives in the vulnerable and loss classes to separate these classes for estimating purposes, so they were combined into one class when the polychotomous function was estimated. To facilitate comparisons between the dichotomous and polychotomous models, we combined vulnerable and loss throughout our analysis.

[5] Loan officers considered it quite remarkable that the new function, with only four explanatory variables, was better able to predict actual loan classifications than the old 29 variable function -- even though the old function had contributed 40 percent to the determination of the actual loan classifications!

[6] FCA annually validates the loan classifications of the BC by assigning its own classifications to a sample of loans, and comparing with the classification assigned by the BC. There are typically very few discrepancies between classifications assigned by the FCA and the BC.

Table 1. Dichotomous logit model — results of estimation.

Coefficients

Variable	Coefficient	Standard Error	Chi-square
INTERCEPT	7.6560	2.1008	13.28
ROABAR	26.5873	6.6964	15.76
DETAS	-15.6961	2.1381	53.89
PROD	1.6177	1.1014	2.16
DUM 3	1.1306	0.4856	5.42

Overall Performance

R = .851

Fraction of concordant pairs = .976

Percent correctly classified:	2-class-system:	93.8
	4-class-system:	73.9

Number of observations: 448

Table 2. Polychotomous logit model — results of estimation.

C o e f f i c i e n t s

Variable	Coefficient	Standard Error	Chi-square
A_1	10.0941	1.0137	99.16
A_2	5.0814	0.7611	44.58
A_3	2.0071	0.7089	8.02
ROABAR	15.7412	2.8834	29.80
DETAS	-7.6164	0.8519	79.93
INTEXS	-49.6796	12.4412	15.95
OENA	1.8666	0.4660	16.05
TLEV	-0.1739	0.0690	6.35
LNWLNA	0.0570	0.0358	2.54
SDROA	-6.9225	3.1553	4.81

O v e r a l l   P e r f o r m a n c e

R = .685

Fraction of concordant pairs = .922

Percent correctly classified: 2-class-system: 93.6

4-class-system: 74.4

Number of observations: 438

Table 3. Analysis of errors by type of cooperative, combined estimating and test samples, dichotomous model.

Type of Cooperative	Number in Combined Sample	2-Class System Errors	2-Class System % Correct	4-Class System Errors	4-Class System % Correct
Fruit & Vegetable	9	1	88.9	5	44.4
Dairy	25	3	88.0	7	72.0
Grain	248	29	88.3	91	63.3
Farm Supply General	282	15	94.7	68	75.9
Farm Supply Petroleum	120	7	94.2	31	74.2
Livestock	4	0	100.0	1	75.0
Wool	1	0	100.0	0	100.0
Sugar	4	0	100.0	3	25.0
Wood Products	4	0	100.0	1	75.0
Service	6	0	100.0	3	50.0
Transportation	2	0	100.0	1	50.0
Other	3	0	100.0	1	66.7
Total	708	55	92.2	212	70.1

Table 4. Cross-tabulation of actual versus predicted classifications, combined sample.

		Predicted Classification and Score				
		V or L	P	A-	A	
		0-.005	.005-.50	.50-.99	.99-1.0	Total
Actual Classification	V or L	18	9	1	0	28
	P	24	66	28	2	120
	A-	1	21	149	57	228
	A	0	2	67	263	332
						708

Key: V = Vulnerable

L = Loss

P = Problem

A- = Acceptable weak

A = Acceptable

REFERENCES

Altman, E. I., Robert Haldeman and P. Narayan. "ZETA Analysis: A New Model for Identifying Bankruptcy Risk." Journal of Banking and Finance 1 (June 1977): 29-54.

Amemiya, Takeshi. "Qualitative Response Models: A Survey," Journal of Econ. Lit. 19 (December 1981): 1483-1536.

Collins, R. A. and R. D. Green. "Statistical Methods for Bankruptcy Forecasting," Journal of Econ. and Business 34 (1982): 349-354.

Dunn, D. and T. L. Frey. "Discriminant Analysis of Loans for Cash Grain Farmers," Ag. Fin. Review 36 (1976): 60-66.

Fair, Isaacs Companies, Articles on Credit Scoring. Fair, Isaacs Companies (November 1980).

Hardy, W. E. and J. B. Weed. "Objective Evaluation for Agricultural Lending." Southern Journal of Agr. Econ. 12 (1980): 159-164.

Harrel, Frank E., Jr., "The Logist Procedure," SAS Applications Guide. SAS Institute Inc. (1980): Ch. 19.

Johnson, R. Bruce and Albert R. Hagen, "Agricultural Loan Evaluation with Discriminant Analysis," Southern Journal of Agr. Econ. 5 (1973): 57-62.

Lufburrow, Jean, Peter J. Barry and Bruce L. Dixon. "Credit Scoring for Farm Loan Pricing," Agr. Finance Review 44 (1984): 8-14.

Pindyck, Robert S. and Daniel L. Rubinfeld, Econometric Models and Economic Forecasts, second ed. (McGraw-Hill Book Co., 1981).

Press, S. James and Sandra Wilson, "Choosing Between Logistic Regression and Discriminant Analysis," Journal of Amer. Stat. Assoc. 73 (1978): 699-705.