AE 88009 August 1988

Logit Analysis for Profit Maximizing Loan Classification

David L. Watt
Timothy L. Mortensen
F. Larry Leistritz*

*The authors are Assistant Professor, Research Assistant, and Professor of the Department of Agricultural Economics, North Dakota State University, respectively.

Abstract

Lending criteria and loan classification methods are developed. Rating system breaking points are analyzed to present a method to maximize loan revenues. Financial characteristics of farmers are used as determinants of delinquency in a multivariate logistic model. Results indicate that debt-to-asset and operating ratio are most indicative of default.

Key words: farm, financial, logistic regression, loan delinquency.

Selected paper presented at AAEA meetings, Knoxville, TN, August 1988.
A more detailed version of this study may also be found in the 1988 annual issue of Ag Finance Review (in press).

LOGIT ANALYSIS FOR PROFIT MAXIMIZING LOAN CLASSIFICATION

A continuing problem for banks is the determination of who to loan money to and who to refuse. The recent special issue of Agricultural Finance Review (Lines and Morehart, Melichar, and Johnson) on Financial Stress in Agriculture is a clear indication of the trauma in the credit sector. Agricultural lenders in particular have been faced with growth in a agricultural credit demand, with increased use of national credit markets (which exposed them to volatile interest rates in the 1980s), and with deteriorating credit worthiness of some farm borrowers. To classify farm borrowers as good or poor risks Luftburrow and Pederson used probit and discriminant analysis, respectively, in their research. Fiske et al. employed logistic regression to model factors influencing currentness of debt payments of Ohio farmers.

The primary cause of lender financial stress is loan delinquency. A predictor of loan default and a model of bank profitability resulting from the decision process is needed. This paper presents results for a specific set of farm borrowers and indicates a method of establishing criteria for classifying loan applications. Significant determinants of debt payment status of North Dakota farmers are identified and used to model the probability of loan default. In conjunction with a loan pricing model the breaking point which maximizes revenue to lending institutions is calculated.

DATA AND STUDY PROCEDURES

Data from statewide longitudinal, cross-sectional surveys conducted by the North Dakota Agricultural Experiment Station in 1985 and 1986 were used in this analysis (Leistritz et al.). Respondents to the surveys were initially screened to include only those who (1) were less than 65 years

old, (2) considered farming to be their primary occupation, and (3) sold at least $2500 of farm products in the past year. Comparisons with the 1982 Census of Agriculture indicate that the 763 survey respondents appear to be representative of the state's farmers who were less than 65 years old and considered farming their primary occupation. The breadth of the surveyed population is greater than the population using a single lending institution. Therefore, this analysis is heuristic and should be tailored to the population served by an individual lender.

Measures taken from farmers' responses were used as determinants of solvency and performance of the farm business. These were thought to be directly or indirectly related to farm financial stress.

Farm operators were current on annual debt obligations if all principal and interest had been paid (or debt had been renegotiated) on January 1, 1986. The dependent variable in the regression model, CURRENT, equals 0 if a farm operator was current on debt payments, and equals 1 if the operator was delinquent.

Debt payment status on January 1, 1986, was thought to be the result of the solvency position of the business on January 1, 1985, performance during 1985, and other socioeconomic measures of the operator. Most explanatory variables are ratios representing profitability and liquidity for 1985, and leverage on January 1, 1985. Absolute dollar measures and other socioeconomic variables thought to be useful as determinants of debt payment status also are tested in the analysis. Limitations are inherent in the surveys and not all measures are available for exploration. The intent of the model is to determine significant 1985 variables that would be indicators of expected debt payment status at the end of the operating year. Table 1 lists the candidate variables used in the analysis.

METHODS

Data are analyzed using the LOGIST procedure (Harrel, 1983) on Statistical Analysis System (SAS) software. We used logistic regression because the dependent variable is binary (0 or 1). In the logit model (1) the probability is bounded by 0 or 1, (2) the breaking point concept

TABLE 1. CANDIDATE VARIABLES.

CURRENT	= 0 if current on debts; 1 otherwise (on January 1, 1986).
AGE	= age of the operator.
LONGFARM	= number of years the farmer has operated the farm.
RATACRES	= the ratio of acres rented to total acres operated.
GROSIN	= gross cash farm income.
NETIN	= net cash farm income.
GPCR	= gross profit on cash revenue.
OPERAT	= operating ratio.
NETCAP	= net cash farm income divided by persons in the household.
VIAB	= the viability ratio.
NETAST	= net cash farm income divided by total assets.
INTGRS	= interest paid divided by gross cash farm income.
CURAT	= a ratio of current debts divided by current assets.
OFFDBT	= a ratio of non-farm income to total debt.
DBTAST	= debt to asset ratio.

is maintained, and (3) coefficients can be ranked by their relative effect on the dependent variable (Press and Wilson, 1978). Although direct continuous relationships between various independent variables and behavior of a farm operator are possible, survey data lack continuous measurement for all attributes associated with farm loan delinquency. The purpose is to predict the likelihood of a farm operator's being current or delinquent given attributes of that particular operator.

A logit model was used to estimate the bivariate events. The cumulative logistic probability function is

(1) $P_i = F(Z_i) = F(a + bX_i) = 1/1+e^{-z_i}$.

Where: P_i = probability of $Y_i = 1$
$Z_i = a + bX_i$
X_i = attribute of individual i
e = base of the natural logarithm
a = intercept parameter
b = parameter associated with attribute X_i.

The logit model is further derived as follows:

(2) $(1 + e^{-Z_i})P_i = 1$,

(3) $\ln (P_i/1-P_i) = Z_i = a + bX_i$.

Z_i expanded to a multivariate model is used as a breaking point (Z_i*) to predict the repayment status of the loan at the end of the year. When Z_i is less than Z_i* the operator is predicted as current (0) on debt obligations and when Z_i is greater than Z_i* the operator is predicted as delinquent (CURRENT=1). Z_i is the natural log of the probability of being delinquent.

RESULTS

Criteria used to select the best model were (1) overall chi-square significance at .10, (2) rho-square (pseudo-R square), and (3) correct classification of observations.

The model selected included only debt-to-asset ratio in 1985 [DBTAST] and the operating ratio [OPRAT] to predict probability of loan default and is

$$\text{(5)} \qquad \ln P/1-P = \underset{(52.4)}{-5.67} + \underset{(31.4)}{3.86}\ \text{DBTAST} + \underset{(10.2)}{1.99}\ \text{OPRAT}.$$

Attempts to specify a model with more than two variables were not successful at the 10 percent level of significance. Models containing other variables also were significant, but chi-square statistics were lower than for the one discussed above. Viability [VIAB] and return on assets [NETAST], when combined separately with debt-to-asset ratio, revealed model chi-square statistics of 38.2 and 41.1, respectively. This model had a chi-square statistic of 51.5 and was significant at the 1 percent level. Individual chi-square statistics are shown in parentheses and all are statistically significant at the one percent level. The model correctly classified 88.2 percent of the observations (Figure 1).

Optimization of the Classification Table.

While the actual percentage of correct classifications is high, the payoffs for the four cells are not equal. Because delinquent loans are expensive the percent of correctly classified predictions is not the best criterion for profit maximization.

Fiske et al. reported that their two model specifications correctly classified 78 and 76 percent of the observations, respectively. If a lender

were to use the example model to classify potential borrowers, an 88.2 percent accuracy rate would cause financial stress for the bank.

LOGIST software defines Z_i^* as PPROB and gives it a default setting of .500 for classifying observations. Observations having a probability greater than PPROB are classified delinquent and observations having a probability of less than PPROB are classified current.

From a lender's perspective, an institution is exposed to the greatest potential loss by farm operators who were predicted to remain current on debt obligations that actually defaulted. Following Saxowsky et al., earnings for building reserves is considered the difference between return on funds invested in securities and the cost of funds. A premium is added to the securities rate to allow for the risk of loaning to individuals or businesses.

CLASSIFICATION TABLE

		PREDICTED	
		CURRENT	DELINQUENT
ACTUAL	CURRENT	QUADRANT 1 338 (280)	QUADRANT 2 4 (62)
	DELINQUENT	QUADRANT 3 42 (16)	QUADRANT 4 7 (33)

PERCENT CORRECT = 88.2 (80.1)

Figure 1. Classification Table Results with PPROB at .5.

NOTE: Results when PPROB is .18 are in parenthesis.

Observations in quadrant one (Figure 1) are borrowers that were correctly predicted as current on payments. For heuristic purposes, revenue derived from loans (loan rate - cost of funds) in this quadrant is 1.5 percent. A bank using this system should use their own payoff values.

In quadrant two the model incorrectly classified potential borrowers as delinquent. In this example funds not loaned to individuals or businesses are assumed invested in securities and earn a return of .5 percent over cost of funds.

Quadrant four contains delinquent borrowers predicted to be delinquent. They are correctly classified, and a lender also would invest funds in securities, earning .5 percent above the cost of funds.

Quadrant three contains the borrowers who were predicted current that actually defaulted. If collateral is sufficient to cover principal and recovery costs, a lender could potentially lose interest income from borrowers who defaulted and never paid. For illustration, losing interest (not principal) on loans is the cost of loan default. Delinquent loan losses vary among institutions, and the average loss is unknown. An institution using this model could use either their historic values or an estimate of future impacts of loan default. In this example, if the interest rate charged on a loan is 11.0 percent and overall cost of funds is 9.5 percent, the lender is penalized and will lose 9.5 percent interest income on misclassified borrowers in quadrant three.

The PPROB that maximizes lender earnings is derived using the tabular method. As PPROB is reduced, potential borrowers classified in quadrants one and three are reclassified into quadrants two and four, respectively. The optimum PPROB is reached when the sum of revenue from the four quadrants is maximized.

Maximizing Lender Revenue.

The model selected was evaluated for sensitivity of the percentage correctly classified and the number of loans written when PPROB was varied from .095 to .5 (Figure 2).

Net income across PPROB values are shown in Figure 3. Maximum net revenue occurred at a PPROB of .18 (equation 5) where slightly more than 80 percent of the observations were correctly classified.

Solution of Model.

The specified model may be used to calculate the probability of a farm operator being delinquent on debt payments. Evaluating both variables at the sample means (.354 for DBTAST and .940 for OPRAT) yields the following solution

(6) $\ln P/1-P = -5.67 + 3.86\ (.354) + 1.99\ (.940)$

$\ln P/1-P = -2.433.$

Solving this equation, P is equal to .081. This indicates that for the sample average, a debt-to-asset ratio of .354 and an operating ratio of .940 the probability of being delinquent is 8.1 percent. The decision rule is to accept up to a probability of 18 percent (PPROB=.18).

SUMMARY

The results of using logistic regression as a tool for classifying farm loan applications as current or delinquent appear promising. The model correctly classified 88.2 percent of the borrowers at a breaking point of .5. The classification of borrowers can be further enhanced by applying the

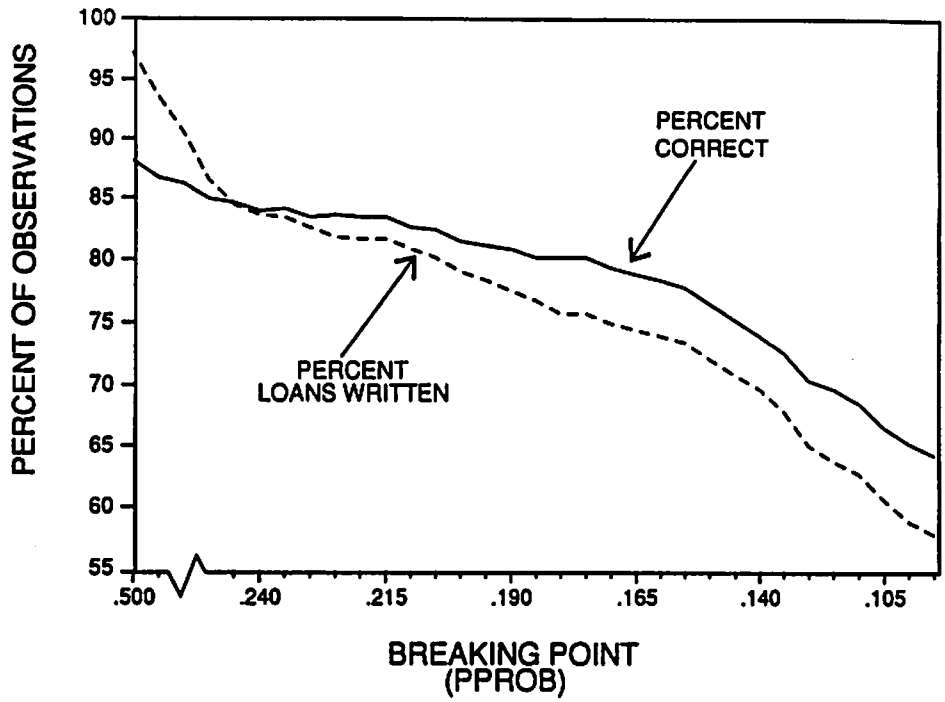

Figure 2. Percent Observations Classified Correct and Percent Loans Written as PPROB is Adjusted

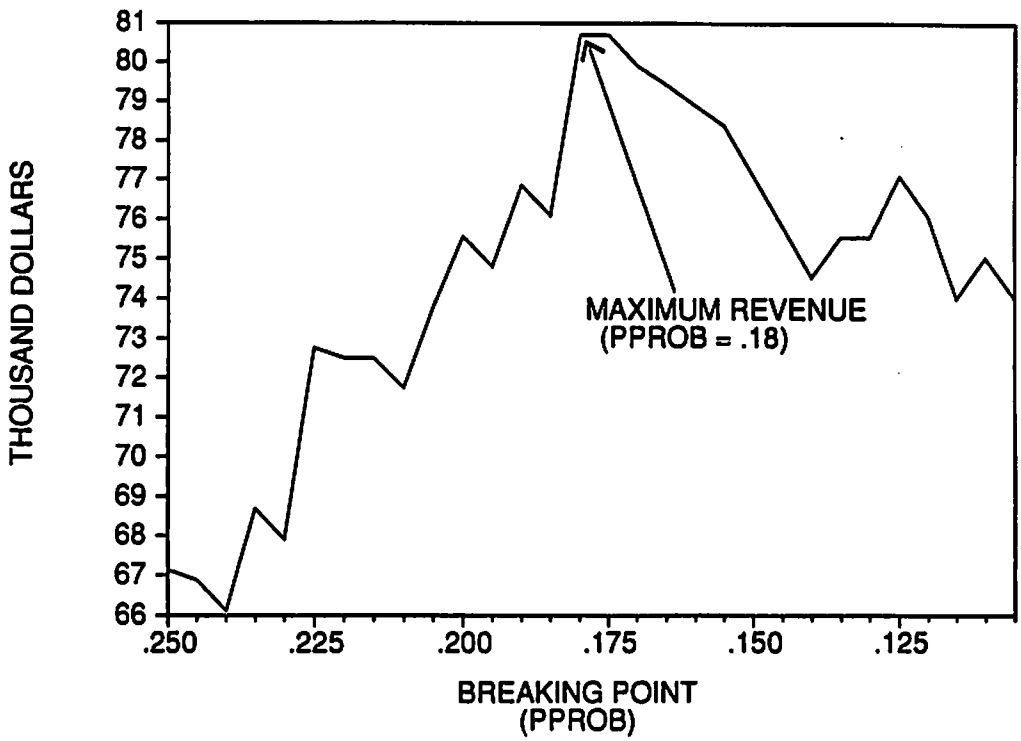

Figure 3. Interest Revenue to a Lender as PPROB is Adjusted

loan pricing model and adjusting the breaking point to maximize profit to a lending institution.

When the breaking point was adjusted downward, borrowers were reclassified among quadrants in the classification table. The resulting 80 percent accuracy rate maximized profit to a lending institution. Some borrowers (16 percent) were predicted delinquent even though they did not eventually default on their obligations. The rejection of "good" borrowers is unavoidable in most situations, even using subjective analysis of loan applicants. The intent was to combine significant variables in an objective analysis of loan applications and use this as a partial guide for basing loan decisions.

Perhaps the most serious limitation of this study was that farms are not homogenous. Farm types, climate, and soil characteristics vary widely across the sample from farm operations exclusively growing small grains or row crops to various livestock operations. Ideally, to determine appropriate indicators at the time a loan decision is made, data from a lending institution are necessary.

Profitability, management ability, and efficiency vary among farm operators, farm types, and geographic regions. The model confirms that profitability and solvency remain as important determinants of successful farm borrowing.

Potential users of this model should use their own historic data from individual borrowers. This would localize weather, soil properties, and local farming practices. An institution can tailor a model to its individual situation. Under the controlled local environment, a greater number of borrower attributes would probably become apparent, and should be combined with the judgement of the loan officer.

LITERATURE CITED

Fiske, John R., Marvin T. Batte and Steven L. Richenbacker. Factors Influencing the Currentness of Debt Payments for Ohio Commercial Farmers. ESO Number 1291. Columbus: Ohio State University, Department of Agricultural Economics and Rural Sociology, 1986.

Harrel, Frank E., Jr. "The LOGIST Procedure." SUGI-Supplemental Library Users Guide. Version 5 Edition. Cary, North Carolina: SAS Institute Inc., 1983.

Johnson, James D., Mitchell J. Morehart, and Kenneth Erickson. "Financial Conditions of the Farm Sector and Farm Operators." Agricultural Finance Review 47(1987):1-18.

Leistritz, F. Larry, Arlen G. Leholm, Brenda Ekstrom, and Harvey G. Vreugdenhil. "North Dakota Farmer's Views on Financial Assistance Policies." North Dakota Farm Research 44, 1(July-Aug., 1986:26-30)

Lines, Allan E., and Mitchell Morehart. "Financial Health of U.S. Farm Businesses in 1984: A Region, Type, and Size Analysis." Agricultural Finance Review 47 (1987):43-52.

Luftburrow, Jean, Peter J. Barry, and Bruce Dixon. "Credit Scoring for Farm Loan Pricing." Agricultural Finance Review 44 (1987):8-14.

Melichar, Emanuel. "Financial Conditions of Agricultural Banks." Agricultural Finance Review 47(1987):23- 39.

Pederson, Glenn D. and Douglas G. Duncan. Performance Pricing of Farm Loans Under Risk: Methods and Implementation. Agr. Econ. Rpt. No. 171. Fargo: North Dakota State University, Department of Agricultural Economics, 1983.

Press, James, and Sandra Wilson. "Choosing Between Logistic and Discriminant Analysis." Journal of the American Statistical Association 73(1978):699-705.

Saxowsky, David M., Cole R. Gustafson, Mir B. Ali, Joan M. Braaten, and Jeff Rotering. Economic Impacts of North Dakota Laws that Permit Delayed or Partial Repayment of Agricultural Debt. Agr. Econ. Rpt. No. 216. Fargo: North Dakota State University, Department of Agricultural Economics, 1987.

TM/mla
7/26/88