Business Performance Based Credit Scoring Models: A New Approach to Credit Evaluation

Madhab R. Khoju and Peter J. Barry

Proceedings of a Seminar sponsored by
North Central Regional Project NC-207
"Regulatory, Efficiency and Management Issues Affecting Rural Financial Markets"
Federal Reserve Bank of Chicago
Chicago, Illinois
October 4-5, 1993

Department of Agricultural, Resource and Managerial Economics
College of Agricultural and Life Sciences
Cornell University, Ithaca, New York 14853

December 1993

BUSINESS PERFORMANCE BASED CREDIT SCORING MODELS: A NEW APPROACH TO CREDIT EVALUATION

Madhab R. Khoju and *Peter J. Barry*[1]

Credit evaluation basically involves making judgements about the future loan performance of a borrower. Such judgements should be contingent on future performance of the financed business, because other sources of the borrower's income (e.g., outside employment) may not be reliable. Credit scoring models, therefore, should focus on evaluating a borrower's future business performance. Similarly, the validity of these models should be based on the accuracy in predicting future business performance.

As an example, one might test to see if the credit score corresponds closely with the borrower's eventual change in net worth, repayment ability or loan performance. Such an approach requires time series data for a sample of borrowers so that the data for the preceding years are used as the estimating sample while the data for the following year constitute the hold-out sample. The estimating and hold-out samples in past credit scoring studies in both commercial and agricultural lending (Dietrich and Kaplan; Lufburrow, Barry, and Dixon; Mortensen, Watt, and Leistritz; Ohlson; Scott; Barry and Ellinger; Srinivasan and Kim; Miller and LaDue; Turvey and Brown; Turvey) consist only of cross sectional observations on different borrowers.

This study proposes and develops a performance based credit scoring model (PBCSM) for credit evaluation based on a borrowers' potential business performance. Business performance is acceptable (problem) if the repayment ability of the financed business is higher (lower) than the repayment obligation. Accordingly, the borrower attributes related to the repayment ability are the potential predictors of business performance in estimating a PBCSM.

The credit scores of borrowers computed from an estimated PBCSM serve as proxies for the repayment probabilities and are compared with an objectively predetermined standard (cut-off level) to distinguish acceptable (who would repay) from problem (who would default) borrowers. If the lender misclassifies an acceptable borrower as problem (type II error), the lender foregoes current interest earnings and potential earnings from a future relationship. On the other hand, if the lender misclassifies a problem as acceptable (type I error), the lender may lose both the principal and accrued interest. The lenders likely are more concerned about type I errors because of their immediate higher costs relative to type II errors.

Depending on the weights used for the predictor variables, even a borrower with relatively undesirable levels of one or more predictor variables (e.g., very high level of leverage) may be evaluated as acceptable from the estimated PBCSM. A lender, may not want such a borrower in the loan portfolio because of the high cost of a potential type I error. To identify such borrowers, this study also develops a financial outlook index.

The importance of the predictor variables in a credit scoring model estimated from cross sectional observations is conditional on the favorable or unfavorable economic environments that generated the estimating sample. Accordingly, the lenders must reestimate the credit scoring model each time the economic environment changes and it involves significant resources. Changes in economic environments, however, may not be a problem for the PBCSM because it is

---

[1] The authors are research associate, Department of Agricultural Economics and Rural Sociology, and professor of agricultural finance, University of Illinois, respectively. This research was funded and conducted at the Center for Farm and Rural Business Finance.

estimated from a time-series of cross-section data. Moreover, the PBCSM may be updated using a simulated data base. For this, the business performance of a representative case farm business may be subjected to a stochastic set of output and output price levels for alternative tenure and solvency positions. Replicating this approach over time results in a data base that is analogous to cross section and time series. The predictive accuracy of a PBCSM estimated from simulated data may then be compared to that of the PBCSM estimated from actual data.

The specific objectives of this paper are threefold - (1) assess the usefulness of a PBCSM in assessing the credit worthiness of agricultural borrowers, (2) use the financial outlook index to further screen borrowers evaluated as acceptable by the PBCSM, and (3) assess the usefulness of updating PBCSM using the simulated data.

Theoretical Framework

A. Performance Based Credit Scoring Model:

Consider a loan of \$L for one period at rate of interest, i. The contractual revenue on this loan is $R_r = (1+i)L$. Actual revenue, however, depends on Y, the repayment ability from the financed business. Accordingly, the realized revenue as a function of Y is given as:

$$\text{Realized Revenue} = \begin{cases} R_r & \text{if } Y > R_r \\ -(R_r - Y) & \text{if } 0 < Y < R_r \\ -R_r & \text{if } Y < 0 \end{cases} \tag{1}$$

In general, the contractual revenue R_r is much smaller than M, the expected value of Y. However, because of the probability that the repayment ability may be less than R_r, the lender may not receive the contractual revenue. For simplicity, the probability distribution of the realized revenue from the loan is represented by a Bernoulli trial:

$$\text{Realized Revenue} = \begin{cases} R_r & \text{with repayment probability } P \\ R_d & \text{with default probability } (1-P) \end{cases} \tag{2}$$

where, R_d is the product of q, the write-off rate of delinquent loans, and L_d, the unpaid balance in the written-off loans.

Assuming the riskless rate of earnings, i_f, from investments like government securities as the opportunity cost of loanable funds, a risk neutral lender will approve loans to only those borrowers for whom the expected returns (represented by the left hand side of equation 3) exceed the opportunity cost of loanable funds (represented by the right hand side of equation 3) i.e.,

$$P^*L^*i - (1-P)^*L_D{}^*q > P^*L^*i_f + (1-P)^*L_D{}^*i_f \tag{3}$$

Moving the right hand side of equation 3 to the left hand side and with some algebra, the expected profit on the loan can be expressed as:

$$P^*L^*(i-i_f) - (1-P)^*L_D{}^*(q+i_f) > 0 \tag{4}$$

It is evident from (4) that the expected profit on the loan depends on both the probability of repayment and the pay-offs in the event of default or repayment. Following Boyes, Hoffman and Low, (4) can be written as:

$$P > [L_D(q+i_f)]/[(L(i-i_f)+L_D(q+i_f)] \tag{5}$$

The credit evaluation of a borrower depends on whether P is higher or lower than the cut-off level given by the right hand side term in (5). Such a cut-off level can be approximated by the lender and, therefore, the credit evaluation hinges on the lender's judgement about P. For this, the lenders utilize the estimated credit score for the ith borrower as a proxy for $\hat{P}_i$, the estimated conditional mean of the repayment probability, P_i, i.e.,:

$$P_i = \hat{P}_i + e_i \tag{6}$$

where, e_i denotes the residual associated with this process.

The repayment probability of a borrower in this study, is approximated by the credit score based on an estimated PBCSM. Formulation of a PBCSM, however, entails identification of borrower attributes affecting repayment ability of the business. Since the firm's solvency and liquidity positions at the beginning of the year and returns from the assets and operating expenses over the year determine repayment ability, these four variables serve as the potential predictors of business performance.

In this study, a logit model is used because of its merit relative to other statistical methods for credit scoring (Collins and Green). Because the business performance is binary i.e., takes a value of 1 if repayment ability is higher than the repayment obligation, and zero otherwise, a dichotomous logit model is specified as:

$$\log [P_i/(1-P_i)] = \alpha + \Sigma\beta_{ij} X_{ij} \tag{7}$$

Where, P_i is the probability of acceptable business performance for the ith borrower, the dependent variable is the logarithm of the odds ratio, α is the intercept, β_{ij} is the logistic regression coefficient, and X_{ij} is the jth attribute of ith borrower related to business performance. By taking antilogs of both sides in (7), P_i is expressed as:

$$P_i = [\exp(\alpha + \Sigma\beta_{ij} X_{ij})]/[1 + \exp(\alpha + \Sigma\beta_{ij} X_{ij})] \tag{8}$$

The estimated P_i from (8) is compared with the cut-off level to distinguish acceptable from problem borrowers. Depending on the parameter estimates such an evaluation can classify a borrower as acceptable even if the lender may not like the levels of one or more borrower attributes. To identify such borrowers, the financial outlook index is discussed next.

B. Financial Outlook Index:

Assume the utility function of a lending officer from evaluating the ith borrower is represented as:

$$U_i(.) = U_i(X_{i1},...X_{ij},...X_{ik}) \tag{9}$$

where, X_{ij} is the level of the jth predictor of business performance for the ith borrower. Each predictor, j = 1,2,...,k, is measured such that the higher the level of the jth predictor, the higher the level of utility. By invoking additive utility theory (Fishburn), (9) can be expressed as:

$$U_i(.) = a_1\, U_i(X_{i1}) + ... + a_j\, U_i(X_{ij}) + ... + a_k\, U_i(X_{ik}) \tag{10}$$

Where a_j, j = 1,...,k are the weights. Equation (10) allows numerical evaluation of the utility function (9) as a weighted average of utilities from each of the predictors. The utility from jth predictor variable, $U_i(X_{ij})$, is evaluated as:

$$U_i(X_{ij}) = U_i(X_{mj})^*F_j(X_{ij}) \tag{11}$$

where, $U_i(X_{mj})$ represents the level of utility from X_{mj} (the maximum value of X_j in the estimating sample), and $F_j(X_{ij})$ is the distribution function of X_{ij} relative to all X_j's in the estimating sample. A similar assumption was made by Shashua and Goldschmidt.

Since X_{mj} is the maximum value of X_j, $U_i(X_{mj})$ is assumed constant and is represented by b_j. The utility function in (10) then can be expressed:

$$U_i(.) = W_1\ F_1(X_{i1}) + \ldots + W_j\ F_j(X_{ij}) + \ldots + W_k\ F_k(X_{ik}) \tag{12}$$

where, $W_j = a_j{*}b_j$

Equation (12) allows numerical computation of the utility function (9) as a linear combination of the distribution function of each predictor variable. However, because it is based on the assumption of additive utility theory, it is valid only if the arguments of the utility function (9) are independent.

The distribution function for a particular predictor may be approximated from the cumulative percent of the predictor in the estimating sample. The weights, on the other hand, may be chosen based on the relative importance of the predictors. By choosing the weights such that $\Sigma W_j = 1$, equation (12) results in a financial outlook index, I_i, whose maximum possible value is 100.

The financial outlook index of a borrower is sensitive to the cumulative percent of each of the predictor variables in the estimating sample. Accordingly, the financial outlook index will be lowered if the cumulative percents of one or more predictor variables are low relative to others in the estimating sample. Accordingly, the financial outlook index resulting from (12) has the potential of identifying the borrowers who have relatively low levels of one or more predictors and thus further screen borrowers evaluated as acceptable from the estimated PBCSM.

Data and Model Specifications

Data for the study came from the Illinois Farm Business Farm Management (FBFM) Association records. Although FBFM started maintaining records of cooperating farmers in Illinois in the 1940's, certification of balance sheets was initiated only in 1985. The majority of certified farm businesses are located in central Illinois and are predominantly grain farms followed by hog and dairy farms. Because the coefficients in (8) and weights in (12) are likely to differ between grain farms and other types of farming enterprises e.g., livestock or dairy, the PBCSM in this study is estimated for grain farms only.

The following two criteria are used in selecting the sample for this study - (1) the grain farms must have certified financial statements for each of the six years (1985 to 1990), and (2) at least three grain farms must be in a given county to reduce heterogeneity in the sample data. In checking the FBFM records, only 74 grain farms satisfied the criteria.

A grain farm is classified as having acceptable (problem) business performance if the repayment ability (sum of depreciation and net farm income net of income tax, social security tax and family living withdrawal) of the business in year t is higher (lower) than the repayment obligations in year t. Because the repayment ability of a business at the end of year t depends on the solvency and liquidity position at the beginning of year t, as well as the realized profitability and operating efficiency during the year t, these criteria are identified as potential predictors of business performance.

In this study, solvency is represented by the equity to asset ratio (EAR). Higher EAR indicates lower debt financing which in turn implies less repayment obligation. The liquidity position is represented by the current ratio (CR) which is measured as the ratio of current assets to current liabilities. Higher liquidity is associated with higher repayment ability. Profitability is represented by the rate of return on equity (ROE) because it is sensitive to the financial structure of the business. ROE is measured as the ratio of net farm income from operations less unpaid labor charge for operator and family to the farm equity. Higher profitability implies a stronger repayment ability of the business. Efficiency is measured by the operating efficiency ratio (OER) which is computed as the ratio of gross returns plus depreciation net of operating expenses to the gross returns. This ratio reflects the efficiency of operating expense management. Hence, higher the OER the higher is the profitability and repayment ability.

The dichotomous logit model is specified as :

$$\log[P_i/(1-P_i)] = \alpha + \beta_{1i}X_{1i} + \beta_{2i}X_{2i} + \beta_{3i}X_{3i} + \beta_{4i}X_{4i} \tag{13}$$

where, X_{1i} is the observed ROE of the business in year t, X_{2i} is the EAR of the business at the beginning of year t, X_{3i} is the CR of the business at the beginning of year t, and X_{4i} is the observed OER of the business in year t. Other variables have been defined above.

Model Estimation and Validation Results

The end of the year balance sheet for the preceding year is the same as the beginning of the year balance sheet for the following year. Accordingly, the beginning of the year balance sheet and the income statement for year t were used to compute business performance for year t and the levels of the predictor variables for year t. For example, the end of the year balance sheet for 1985 and the income statement for 1986 were used to compute the business performance for 1986. This resulted in a pooled time series of 370 observations on business performance and predictor variables (74 for each of the years 1986 to 1990). Of this, a pooled time series of 296 observations (74 observations from 1986 to 1989) were used as an estimating sample and the observations for 1990 as a hold-out sample. In the estimating sample 142 observations (i.e., 47.97 percent) had acceptable business performance and the remaining 154 (i.e., 52.03 percent) had problem business performance. Since the businesses are chosen and then their performances are observed later, the sampling design used is considered exogenous (Manski and Lerman).

SAS LOGIST procedure was used for model estimation. Because the coefficient of current ratio had a p-value of 0.69 it was omitted from the final model. The estimates of the final model are presented in Table 1. All coefficients are significant at the 0.02 level. The R statistic for the model is 0.575 which is relatively high. R statistic is similar to the multiple correlation coefficient in the normal setting after a correction is made to penalize for the number of parameters to estimated (Harrell). To highlight the relative importance of predictor variables in indicating acceptable business performance, their estimated elasticities (evaluated at the means) are also presented in Table 1. These elasticities represent the percent increase in the probability of acceptable business performance for one percent increase in the corresponding predictor variable. Based on estimated elasticities the OER has the greatest impact on the probability of acceptable business performance followed by EAR and ROE, respectively.

Validation of the Estimated PBCSM

The validity of the estimated PBCSM was examined by its accuracy in predicting known 1990 business performances of 74 grain farms in the hold-out sample. Of the 74 businesses in the hold out sample, 32 had acceptable and the remaining 42 had problem business performances. The predicted business performance for each observation is based on whether the estimated probability of acceptable business performance is higher or lower than the predetermined cut-off probability level. The estimated model is validated using cut-off levels of 50 and 60 percent in order to examine the sensitivity of the model validation to changes in the cut-off probability. The predicted business performance for each observation is then compared with its actual performance.

The results are presented in Table 2 for the cut-off probability levels of 50 and 60 percent, respectively.

Table 1. Estimated Performance Based Credit Scoring Model

Variable	Coefficient	Std. Error	Chi-square	P Value	Elasticities
Intercept	-6.04	0.88	46.67	0.00	
ROE	14.95	2.57	33.79	0.00	0.536
EAR	2.91	0.93	9.65	0.02	1.060
OER	5.82	1.34	18.71	0.00	1.690

Model Chi-square with 3 degrees of freedom = 141.71.
P Value = 0.00.
R Statistic = 0.575.

Using the information in Table 2, different prediction measures of the estimated PBCSM for the hold-out sample are calculated and presented in Table 3. Table 3 also presents the prediction accuracy of the estimated model in the estimating sample. The estimated PBCSM correctly predicted 74.32 and 78.38 percent of the business performances in the hold-out sample for the cut-off levels of 50 and 60 percent, respectively. These predictive accuracies are associated with type I errors of 30.96 and 19.05 percent, respectively. The estimated model also correctly predicted the business performances of 80.40 and 77.03 percent of the estimating sample for the cut-off levels of 50 and 60 percent, respectively. The corresponding type I errors are 20.13 and 12.99 percent, respectively. These results do not differ significantly from the prediction accuracies of reported credit scoring models in the literature e.g., 79.7 percent for Lufburrow, Barry and Dixon; 85.7 percent for Miller and LaDue, and 69.7 percent for the Logit model in Turvey. However, because the economic environment generating the observations in the hold-out sample may be different from the economic environment that generated the estimating sample, the validity of the PBCSM is subjected to a more stringent test than these credit scoring models.

Table 2. Classification Table for Alternative Cut-off Levels, Actual Data

	Predicted Classification		
	Acceptable	Problem	Total
50 Percent Level			
Actual Acceptable	26	6	32
Classification Problem	13	29	42
Total	39	35	74
60 Percent Level			
Actual Acceptable	24	8	32
Classification Problem	8	34	42
Total	32	42	74

As expected, the prediction results in Table 3 indicate a modest degree of sensitivity to changes in the cut-off level. Raising the cut-off level from 50 to 60 percent reduces the likelihood of type I error; however, the likelihood of type II error also rises.

Table 3. **Prediction Measures**

Measures	Cut-off Level of 50 Percent	Cut-off Level of 60 Percent
Hold-out Sample		
Total Predictive Accuracy	74.32	78.38
Type I Error	30.96	19.05
Type II Error	18.75	25.00
Estimating Sample		
Total Predictive Accuracy	80.40	77.03
Type I Error	20.13	12.99
Type II Error	19.01	33.80

Construction of Financial Outlook Index

Two requirements must be satisfied for the construction of financial outlook index - (1) each predictor should be measured such that the higher the level of the predictor, the higher is the lender's utility level, (2) the predictor variables must be independent. The measures of ROE, EAR and OER (significant predictors of business performance) satisfy the first requirement. However, the estimated sample correlation coefficients among these variables are significantly different from zero and, hence the second requirement is not satisfied. The predictor variables are, therefore, needed to be represented by their transformations such that the transformed variables are independent of one another. For this, the three predictor variables are substituted by three principal components, PC_p; p = 1,2,3. Accordingly, equation (12) is expressed as

$$U_i(.) = W_1\ F_1(PC_{i1}) + W_2\ F_2(PC_{i2}) + W_3\ F_3(PC_{i3}) \qquad (14)$$

The computation of financial outlook index, therefore, requires - (a) the construction of three principal components for each observation in the hold-out sample, (b) the cumulative percent of each of these principal components in the estimating sample, and (c) the weights. The principal components for each observation in the hold-out sample are constructed in two steps - (i) the predictor variables are first standardized using their means and standard deviations in the estimating sample, and (ii) the standardized predictors are then multiplied by the eigne vector corresponding to each principal component in the estimating sample. Moreover, since the eigne values represent the amount of variance in the data accounted for by the principal components, the proportion of the variance explained by each principal component is used as the weights in computing the financial outlook index.

The resulting financial outlook index is used to further evaluate the borrowers in the hold-out sample that are acceptable by PBCSM. For the cut-off levels of 50 and 60 percents, the estimated PBCSM respectively predicted 39 and 32 observations in the hold-out sample to have acceptable business performances (Table 2). However, of these only 26 and 24 actually had acceptable business performance. Hence if the credit evaluation was based on PBCSM alone, there was a potential of committing type I error in 13 and 8 cases respectively. This is where the financial outlook index may be helpful.

Of the 39 observations in the hold-out sample that are acceptable by the PBCSM at 50 percent cut-off level, only 32 have a financial outlook index greater than 50. The remaining 7 with the financial outlook index less than 50 have in fact problem business performances. Similarly, of the 32 observations in the hold-out sample that are acceptable by the PBCSM at 60 percent cut-off level, only 28 had a financial outlook index greater than 50. The remaining 4 with less than 50 have in fact problem business performance. Hence, if the financial outlook index is used to further screen acceptable borrowers from PBCSM, the potential type I error may be reduced by 53.84 (=(7/13)*100) and 50 (=(4/8)*100) percent, respectively.

Simulated Data for Updating PBCSM

Two approaches to generate a data base for updating PBCSM were explored. First, a stochastic multi-period model of a case grain farm business was formulated. The attempts to establish the validity of such a multi-period model were not successful because of the lack of detailed information about the case farm in FBFM records. As a result, the data base is generated using separate simulation models for each year (with no linkages over the years) of the case grain farm. This approach utilizes the available certified balance sheet and income statements for the case farm. These statements in a given year reflect actual business performance for the chosen asset structure, operating expenses, financing terms, tenure status, solvency position, production, price levels, etc.

A data base on business performance and its predictors is generated by subjecting the case farm to a set of stochastic outputs and their prices under alternative solvency and tenure scenarios. The solvency position reflects the repayment obligation due and interest expenses. The tenure status reflects the operator's share in the revenue and operating expenses as well as the liability structure (because of differences in the level of real estate loan to finance real estate across tenure status). Hence for the same volumes and prices of outputs, the income statement for the same asset structure will differ across solvency and tenure scenarios. The data base generated this way may be interpreted as the potential consequences if the farm operator were to choose the same asset structure across alternative solvency and tenure positions.

A farm business in Champaign county was randomly selected as a case farm from a list of grain farms in FBFM records that represented majority of grain farms in central Illinois. The case farm's balance sheet information for 1985 to 1990, and actual crop acreages, crop yields, prices, tenure status, operating expenses, operator's share in revenue and expenses, other farm income including government payments, interest expenses, depreciation and unpaid family and operator for 1986 to 1990 are reported in Khoju.

The crop revenue is computed as the sum of the returns from corn and soybeans. The operator's revenue includes 100 percent of the revenue on the land the operator owns plus one-half of the revenue on the leased land. The operator's crop revenue is then subtracted from operator's reported revenue (in FBFM records) to calculate the amount of other farm income including government payment.

Alternative solvency positions affect the repayment obligations and interest expenses, and, hence, the repayment ability. For data generation purpose, four debt-to- asset ratios - 0.2, 0.4, 0.6 and 0.8 - were used to represent solvency scenarios. The liabilities under alternative debt-to-asset ratios are adjusted using the debt adjustment factor (the ratio of assumed debt-to-asset ratio to the actual debt-to-asset ratio). This factor is used to proportionately adjust the short, intermediate and long term liabilities. The interest expenses on the liabilities are calculated using the actual interest rates prevailing in each year (Agricultural Statistics, 1990).

Alternative tenure status is associated with different operator's share in the revenue and operating expenses as well as the liability structure. Since these variables have direct effects on the repayment ability, for data generation purpose, five tenure levels (proportion of tillable acres owned by the operator) are considered - 0.05, 0.20, 0.40, 0.60 and 0.80. The farm size across tenure status, however, is assumed fixed to maintain similar machinery and other required inputs. The operator's share of land ownership under alternative tenure scenarios is adjusted by the land adjustment factor (the ratio of assumed tenure status to actual tenure status).

The operator's ownership of land is assumed to be financed by the same ratio of debt financing as the actual ratio of real estate loan to the farm real estate, henceforth called the land-debt factor. For the changes in the real estate loans, the principal amount due each year is also adjusted using the ratio of actual annualized principal amount due to the actual real estate loan. These adjustments alter the liability structure and, therefore, the solvency positions. Since four alternative solvency positions have already been considered, the reported median debt to asset ratios for the group of farms within tenure status of 0-10, 11-25, 26-50, 51-75 and above 75 (Financial Characteristics of Illinois Farms) are used as the solvency position for assumed tenure scenario of 0.05, 0.20, 0.40, 0.60 and 0.80, respectively. The liability structure is then adjusted using the ratio of median debt to asset ratio to the debt-to-asset ratio for the given tenure status resulting from the adjustments in land ownership and its financing. Detailed discussions about the adjustment factors can be found in Khoju.

For the given solvency and tenure scenarios, the business performance for the asset structure depends on the realized outputs and their prices. Prediction of these stochastic variables for each of the years (1986 to 1990) is discussed next.

Prediction of Crop Yields and Prices

The farm level yields and prices for each crop (corn and soybeans) are assumed to be uncorrelated for two reasons. First, the production level at a particular farm is not large enough to affect its price (assuming a perfectly competitive market). Secondly, crop prices are not expected to affect current production levels because acreage decisions are based on future expected prices. Corn and soybean yields, however, are assumed to be correlated because they are grown under the same conditions (weather, pest damage etc). Similarly, since corn and soybeans are substitutes to each other to some degree their prices are also assumed to be correlated. Accordingly, the crop yields and prices can be predicted separately.

Farm level crop yield reflects two sources of risk: 1) aggregate factors such as weather and pest infestations in a county, and 2) idiosyncratic factors such as localized weather, soil fertility and management. The variability attributable to these factors represents, respectively, the systematic and nonsystematic risks of an individual farm yields. Because county average yield (acreage weighted average of individual farm yields) diversifies away the nonsystematic risk, it may be used to represent aggregate risk factors. As a result, just as the return of an individual security is related to the level of an index (Sharpe's single index model), the crop yields of an individual farm are linearly related to the county average yields and random elements. Estimation of such a relationship is used to predict farm level crop yields given the predicted levels of county average yields (systematic risk) and random elements (idiosyncratic risk). The county average yields are predicted from their estimated density function using their available long series, and the random elements are predicted from the distribution of the residuals of the estimated single index models. This approach of predicting farm level crop yields is discussed in detail in Khoju, Nelson and Barry. Following this approach a total of 225 corn and soybean yields (25 for each of 9 scenarios i.e., 4 solvency and 5 tenure positions) were predicted for each of the years 1986 to 1990.

Time-series on corn and soybean nominal prices are available only at the county level (Illinois Agricultural Statistics). Because crop prices received by farmers in a county are essentially equal, the historical data are used to estimate the parameters of the density function, the random draws from which are used to represent farm level prices. Because negative crop prices are not possible, corn and soybean prices are assumed to have log-normal distributions. Since the most recent prices contain the most information about future prices, the parameters of the density function were estimated using the price series from 1970 onward only. The predicted corn and soybean prices are represented by the random draws from their estimated bivariate log normal distribution. To predict prices for 1986, the bivariate log normal distribution was estimated using the price series from 1970 to 1985. Similarly, the prices for 1987 to 1990 were predicted. As with the crop yields, 225 prices were drawn randomly for each year.

Model Estimation Based on Generated Data Base:

Using the beginning of the year balance sheet and the predicted price and production levels, the income statement for each year is computed for assumed solvency and tenure scenarios. The business performance and the predictor variables for the logit model are computed as in the case of pooled time series. This approach resulted in a sample of 900 observations (225 for each of the years 1986 to 1989) on business performance and predictors.

The generated sample of 900 observations was used to examine the usefulness of PBCSM estimated from the simulated data base. Of these, 638 (i.e., 70.88 percent) had acceptable business performance and the remaining 262 (i.e., 29.12 percent) had problem business performance. These were used to estimate the PBCSM. Since the liquidity measure was not a significant predictor of business performance in the pooled time series, it was dropped as a potential predictor of business performance in the generated data base. The estimates of the model are presented in Table 4. All regressors are highly significant and have the expected signs. The R statistic for the model is 0.88 which is higher than 0.575 for the model estimated from the observed pooled time series data base. The estimated elasticities of the predictor variables (computed at the means) are also presented in Table 4. These elasticities indicate that the EAR has the greatest impact on the probability of acceptable business performance followed by OER and ROE respectively.

Table 4. Performance Based Credit Scoring Model Estimated from Generated Data Base

Variable	Coefficient	Std. Error	Chi-square	P Value	Elasticities
Intercept	-27.00	3.03	79.04	0.00	
ROE	60.06	6.66	81.34	0.00	0.159
EAR	25.42	2.47	105.17	0.00	0.395
OER	12.88	3.18	16.40	0.00	0.211

Model Chi-square with 3 degrees of freedom = 847.08.
P Value = 0.00.
R Statistic = 0.88.

Validity of PBCSM Based on Generated Data

To examine the usefulness of the PBCSM estimated from the generated data, the estimated model was also used to predict business performance in the hold-out sample of the pooled time series (i.e., actual business performances of 74 grain farms in 1990). For this, the coefficients estimated from the simulated data were used to compute the estimated probability of acceptable business performance for each observation in the hold-out sample. Parallel to the PBCSM estimated from the observed pooled time-series, this model was also validated using cut-off levels of 50 and 60 percent. The predicted business performances are compared with their actual performances and the classification results are presented in Table 5.

Using the information in Table 5, different prediction measures of the estimated models are computed. For comparison, these prediction measures for the PBCSMs estimated from actual and generated time series are presented in Table 6. While both models have respectable degrees of accuracy, the prediction accuracy of the PBCSM estimated from pooled time series was higher and it has lower type I error. The pooled time-series represents the experiences of a group of farms over time, rather than the single farm in the case of the generated data. The pooled time series for a sample of grain farms represent the actual relationship of predictor variables to business performance in each of the years 1986 to 1989. The generated data base was created by subjecting the case grain farm to a set of stochastic production and prices under alternative solvency and tenure positions. Moreover, the hold-out sample consisting of 74 grain farms are heterogenous in terms of farm size. The predictive accuracy of the PBCSM estimated from generated data, therefore, may be increased if the model is estimated from generated data for three case grain farms each representing small, medium and large farms, respectively.

Table 5. Classification Table for Alternative Cut-off Levels, Generated Data

	Predicted Classification		
	Acceptable	**Problem**	**Total**
50 Percent Level			
Actual Acceptable	28	4	32
Classification Problem	23	19	42
Total	51	23	74
60 Percent Level			
Actual Acceptable	28	4	32
Classification Problem	21	21	42
Total	39	25	74

Of all the credit scoring studies related to agricultural lending, Turvey and Brown (1990) examined the prediction accuracy of the estimated credit scoring model using the observations for the following years as the hold-out samples. They estimated the credit scoring model using the observations for 1981 and examined the prediction accuracy using the observations for 1982. The estimated model had a total predictive accuracy of 61.29 percent with 71.70 and 8.64 percent type I and type II errors. Based on the type I error for one year ahead projection, the PBCSM estimated from the simulated data for a single case grain farm yielded a lower type I error than Turvey and Brown credit scoring model. Such results are expected because the PBCSM is estimated from the observations over four years while Turvey and Brown model was estimated from the observations of only one year.

Table 6. Prediction Measures of Estimated PBCSM Based on Simulated Data

	PBCSM from Pooled Data		PBCSM from Simulated Data	
Measures	**50 Percent**	**60 Percent**	**50 Percent**	**60 Percent**
Total Predictive Accuracy	74.32	78.38	63.51	66.22
Type I Error	30.96	19.05	54.76	50.00
Type II Error	18.75	25.00	12.50	12.50

Concluding Comments

This study is designed to assess the usefulness of PBCSM as a new approach to credit evaluation of agricultural firms. The results of the study indicate that the PBCSM has respectable degree of prediction accuracy and is close to the prediction accuracies of existing credit scoring models even though the PBCSM is subject to a more stringent test. The credit evaluation based on PBCSM should be appealing to lenders because it is based on the potential performance of the financed business, rather than on subjective classification of lenders. Moreover, the financial health of the lender also depends on the business performance of the borrowers in terms of both timely loan repayment and increased loan demands in the future.

The financial outlook index developed in this study allows the lender to further investigate the levels of the predictor variables of the borrowers that are evaluated as acceptable from estimated PBCSM. This index helps in identifying and excluding the borrowers with relatively undesirable levels of one or more predictor variables from the loan portfolio. The results of this study indicate that adoption of this approach may reduce the type I error by as much as 50 percent. Hence, the financial outlook index should be of interest to the lending institutions.

This study also evaluated the usefulness of one cost effective technique of updating the PBCSM -- a farm level simulation approach of data generation. The validation results of the PBCSM estimated from the simulated data indicate respectable degrees of prediction accuracy. Accordingly, this approach of updating PBCSM has potential usefulness and warrants further research.

References

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

Barry, P.J., J.A. Hopkin, and C.B. Baker. *Financial Management in Agriculture.* 4th ed. Danville, IL: Interstate Printers and Publishers Inc., 1988.

Barry, P.J. and P.N. Ellinger. "Credit Scoring, Loan Pricing and Farm Business Performance." *Western Journal of Agricultural Economics* 14(1989):45-55.

Boyes, W.J., D.L. Hoffman, and S.A. Low. "An Econometric Analysis of the Bank Credit Scoring Problem." *Journal of Econometrics* 40(1989):3-14.

Chhikara, R. "The State of the Art in Credit Evaluation." Invited Paper presented at 1989 annual meeting of American Agricultural Economics Association in Baton Rouge, Louisiana, July 30-August 2, 1989.

Collins, R.A. and R.D. Green. "Statistical Methods for Bankruptcy Forecasting." *Journal of Economics and Business* 34(1982):349-54.

Dietrich, R. and R. Kaplan. "Empirical Analysis of the Commercial Loan Classification Decision." *Accounting Review* 57(1982):18-38.

Fishburn, P.C. "Methods of Estimating Additive Utility." *Management Science* (March 1967): 435-53.

Fishburn, P.C. *Utility Theory for Decision Making.* Wiley: New York, 1970.

Gustafson, C.R. "Promising Future Research Related to Credit Scoring." Paper presented at Annual Meeting of NC-161, Denver, October 6 and 7, 1987.

Harrell, Jr., F.E. *The Logist Procedure.* SAS Institute Inc., Cary, North Carolina.

Khoju, M.R., C. Nelson, and P.J. Barry. "Application of Single-Index Model to Predict Farm-Level Yield Data." Paper Presented at the AAEA Annual Meeting, August 8-12, 1992, Baltimore, Maryland.

Khoju, M.R. "Performance Based Credit Scoring Model." Unpublished Ph.D. Dissertation, University of Illinois at Urbana-Champaign, 1993.

Lufburrow, J., P.J. Barry, and B. Dixon. "Credit Scoring for Farm Loan Pricing." *Agricultural Finance Review* 44(1984):8-14.

Maddala, G.S. *Limited Dependent and Qualitative Variables in Econometrics.* Cambridge University Press, 1987.

Manski, C.F. and S.R. Lerman. "The Estimation of Choice Probabilities from Choice-Based Samples." *Econometrica* 45(1977):1977-88.

Miller, L.H. and E.L. LaDue. "Credit Assessment Models for Farm Borrowers: A Logit Analysis." *Agricultural Finance Review* 49(1989):22-36.

Mortensen, T., D.L. Watt, and F.L. Leistritz. "Predicting Probability of Loan Default." *Agricultural Finance Review* 48(1988):60-67.

Neter, J. "Financial Ratios as Predictors of Failure: Discussant." *Journal of Accounting Research Supplement* 4(1966):112-18.

Ohlson, J.A. "Financial Ratios and the Probabilistic Prediction of Bankruptcy." *Journal of Accounting Research* 18(1980):109-31.

Pyndick, R.S., and D.L. Rubinfeld. *Econometric Models and Economic Forecasts.* 2nd ed., New York: McGraw Hill Book Co., 1981.

Scott, J. "The Probability of Bankruptcy: A Comparison of Empirical Predictions and Theoretical Models." *Journal of Banking and Finance* 5(1981):317-44.

Shashua, L., and Y. Goldschmidt. "An Index for Evaluating Financial Performance." *The Journal of Finance* 34(1974):797-814.

Srinivasan, V. and Y.H. Kim. "Credit Granting: A Comparative Analysis of Classification Procedures." *Journal of Finance* (July 1987):665-81.

Turvey, C.G. and R. Brown. "Credit Scoring for a Federal Lending Institution: The Case of Canada's Farm Credit Corporation." *Agricultural Finance Review* 50(1990):47-57.

Turvey, C.G. "Credit Scoring for Agricultural Loans: A Review with Applications." *Agricultural Finance Review* 51(1991):43-54.