Review of Agricultural Economics—Volume 28, Number 1—Pages 4–23
DOI:10.1111/j.1467-9353.2006.00270.x

Determining the Probability of Default and Risk-Rating Class for Loans in the Seventh Farm Credit District Portfolio

Allen M. Featherstone, Laura M. Roessler, and Peter J. Barry

Credit risk is the primary risk facing financial institutions. With the proposed guidelines under the New Basel Accord, financial institutions will benefit from better assessing their risks. The probability of default (PD) and risk-rating class is studied for 157,853 loans in the Seventh Farm Credit District Portfolio. Repayment capacity, owner equity, and working capital origination loans are important determinants of the PD. Standard & Poor's (S&P) reported probabilities of default were used to classify each of the loans into a risk-rating class. The average predicted PD is 1.61%, which would fall into the BB− S&P class.

Capital management is a significant topic for financial institutions as they prepare to meet new guidelines under the proposed New Basel Accord (Basel Committee on Banking Supervision). The advanced options of the Accord will require financial institutions to assess their credit, market, and operational risk, to link these risks to capital management, and to comply with safety and soundness regulations "for uniform application throughout" the institution (Barry 2001). The goals of the Accord are to tailor risk measurement and capital management to the characteristics of diverse financial institutions and to more finely distinguish segments of their loan portfolios.

Credit risk is the primary source of risk to financial institutions. As part of the credit risk assessment, the Accord suggests more granularity in risk-rating classes than currently exists, along with other improvements to existing risk-rating systems. Many banks are using more advanced rating systems under the

■ *Allen M. Featherstone is a professor of Agricultural Economics at Kansas State University.*
■ *Laura M. Roessler is a master of Agribusiness graduate student at Kansas State University.*
■ *Peter J. Barry is a professor of Agricultural Economics at the University of Illinois.*

assumption that they will lead to decreased capital requirements compared to those institutions that do not adhere as closely to the Basel Accord (Yu, Garside, and Stoker). The advanced rating systems generally include dual ratings—one for the probability of default (PD) and one for the amount of potential loss given that default occurs.

Virtually all banks use some type of risk-rating systems (Siddiqi and Klein). Ratings serve multiple purposes, including contributing to the loan origination process; aiding in monitoring the safety and soundness of loan portfolios and in management reporting; facilitating adequacy of loan loss reserves; and providing components of loan pricing and profitability analysis systems (Barry 2003). Improved risk ratings can also heighten returns for financial institutions by lowering processing costs (Yu, Garside, and Stoker).

Financial institutions that serve agriculture also have a need to review their existing risk-rating systems. Different institutions have their own definitions of risk and risk ratings, making it difficult to uniformly assess the risk held by agriculture. The objective of this article is to develop a consistent risk-rating system with increased granularity using data from the Seventh Farm Credit District (AgriBank). First, the AgriBank's historical loan portfolio (comprised of eleven states and $22.2 billion of loans outstanding on 31 December 2002) will be used to model loan default. These results will be used to determine probabilities of default for each of the loans in the portfolio, and to map these loans into a Standard & Poor's (S&P) probability of bond default for publicly rated companies. The resulting predicted probabilities of default will be compared to actual default experienced by the loans in those classes to determine the accuracy of the model.

Related Literature

The "fundamental goal" of a credit risk-rating system is to accurately estimate the credit risk of a specific transaction or portfolio of transactions/assets. The "ultimate goal" is to measure the expected and unexpected loss from investing in an asset and the capital required to support it. The risk-rating system must be able to produce consistent results. Different results for loans that appear to be similar damages the credibility of the risk-rating system (Yu, Garside, and Stoker). If these systems are applied accurately and consistently, portfolio management will be successful (Crouhy, Galai, and Mark). According to Yu, Garside, and Stoker, "it is essential that credit risk is measured in a more standardized, accurate, and timely fashion" (p. 39).

Most risk-rating systems consider both quantitative and qualitative elements of a loan. Analysis is completed on many attributes of the borrower including financial, managerial, earnings and cash flow, quality of assets, and liquidity of the firm (Crouhy, Galai, and Mark). Lenders tend to rely more heavily on repayment capacity, solvency, and loan security than on the borrower's profitability and financial efficiency. A weakness of many risk-rating systems is that they are based on historical financial information generated under conditions that may not be applicable in the future. Additionally, the information utilized in creating risk-rating systems is often based on accounting data that are updated infrequently and at discrete points in time and do not fully reflect the dynamics of the firm (Crouhy, Galai, and Mark).

Default mode and mark-to-market[1] are the two main approaches to measuring credit risks identified in the literature (Barry 2001). The default mode approach "focuses directly on the possibilities of loan loss, including PD and the severity of loss given that default has occurred" (Barry 2001). The mark-to-market approach attempts to measure how future changes in the credit risk characteristics of a loan or a group of loans will affect the loan(s) market value, including potential losses in value (Barry 2001).

The default mode, used in this article, is more traditional and widely applied. Historical data are an issue with the default mode; there must be enough loans that have experienced default and loss to derive statistically sound measures of expected loss (EL), unexpected loss, and recovery rates. Within the default mode approach, there are several ways of analyzing the data. One method is based on default rates from the year of issue, another focuses on the incidence of default based on the actual year a loan was originated, and the last is based on the exposure years. The method utilizing default rates from the year of issue is an "explicit depiction" of the historical time pattern of losses, and projections of default rates can thus reflect these time patterns. The year of origination method focuses on the incidence of default for loans originated in specific years (Barry 2001). This method was utilized in developing regulatory capital requirements for certain government sponsored enterprises such as Fannie Mae and Freddie Mac, and is applicable when the borrowers' underwriting data are only available for the year of origination (Barry 2001, Barry 2003).

While a goal of risk-rating systems is to produce accurate and consistent ratings, professional judgment and experience must also be part of the rating process (Crouhy, Galai, and Mark). Many financial institutions utilize some type of statistical modeling to aid in the rating process; however, most believe that judgmental rating systems provide "greater accuracy, confidence, and flexibility." Judgmental rating systems are more costly, but the benefits of using these systems may outweigh the higher costs with larger loans due to economies of size and spreading the cost over a larger loan volume base.

As financial institutions evaluate and improve their risk-rating systems, portfolio managers are moving from judgmental assessment toward greater dependence on quantitative assessment (Czuszak). Quantitative assessment speeds up the risk-rating process and can lead to more standardized results (Yu, Garside, and Stoker). To improve the efficiency of the quantitative models, financial institutions need to ensure they are implementing sufficient granularity in the design of the risk classes.

The New Basel Accord suggests a minimum of six classes for performing loans and two classes for nonperforming loans. The benefits of increased granularity include more complete distinctions of risk, and improved portfolio and profitability analysis as it relates to loan assets (Barry 2003). If the level of granularity is too small, the system may not generate a true distribution of loans across the risk classes and will be less useful in management's decision making. Conversely, too many grades may provide too much detail or lead to a false sense of security (Yu, Garside, and Stoker). Increasing the number of risk-rating classes allows financial institutions to more closely calibrate the risk ratings to the individual characteristics of each loan and the borrower's business.

Credit Scoring Literature

There has been a long tradition of studying credit scoring models in the agricultural finance literature (Betubiza and Leatham). Miller and LaDue examined the status of credit scoring and developed a model for 203 New York borrowers. Turvey and Brown examined 9,403 Canadian loans of which 63% were current and 37% were in arrears. They found that estimating the model using 1981 data and using 1982 and 1983 data as a hold out sample resulted in a 71.7% type I error and a 8.64% type II error. Khoju and Barry compared the use of simulation models to credit scoring models to help with the issue of data generation. Novak and LaDue compared the use of multiperiod to single-period measures and found that the multiperiod measures provided more stable parameter estimates.

While the credit scoring literature provides an important backdrop for our study, it does not examine the accept/reject decision as is studied in previous analyses listed above. Using a sample of only accepted loans would make the results susceptible to Joy and Tollefson's criticism of a nonrandom sample. In addition, the cutoff level for classifying a loan as acceptable or not is a difficult issue. Changing that level can dramatically affect the level of the type I and type II error. Our article uses "credit scoring" techniques to rate a portfolio of loans on an "objective" basis. The ultimate goal of Basel is to make the rating of loan portfolios more objective and to identify loans that have a low PD versus loans that have a higher PD and appropriately price those assets.

Conceptual Framework

Three fundamental issues link credit risk and capital adequacy: how much risk is present, what is the institution's tolerance for risk, and how much capital should be held to offset that risk. EL rates can also be broken down into three elements and analyzed separately for each transaction to gauge that transaction's risk. These elements are PD, loss given default (LGD), and exposure at default (EAD). The PD is an indication of the likelihood and frequency that a loan will enter default status. LGD measures the impact on the institution from default. LGD is the net of any recovery the institution has received, either through liquidation of collateral or deficiency judgments rendered from foreclosure or bankruptcy proceedings and may be associated with quality of collateral, seniority of claims, and any guarantees associated with the loan. EAD is what the institution has at risk when the loan does enter default status. EAD is usually expressed in dollar terms and is comprised of principal outstanding, unutilized commitment, and any fees or other expenses the institution incurs in collecting the default. Both PD and LGD are usually expressed in percentage terms.

The relationship between PD, LGD, EAD, and EL is expressed as follows:

$$EL = PD * LDG * EAD. \tag{1}$$

Financial institutions attempt to minimize loss rates "at different levels of aggregation (individual loans, portfolios of loans) and for baskets of loans representing different types of borrowers and loan characteristics" (Barry 2001). Thus, PD can be predicted by the type of borrower, underwriting variables, loan size, maturity, payment frequency, and many other variables dependent on either the borrower characteristics or external economic factors. While the model is multiplicative, it is common in the finance literature to examine these aspects separately mainly due to the inability of data information systems to easily track loans through the default process. Katchova and Barry modeled the EL in this manner. In addition, Featherstone and Boessen studied loan loss severity using LGD $*$ EAD separately from the PD due to data availability.

The Basel Accord suggests eight criteria for banks to measure when implementing a risk-rating system including evaluation of a firm's repayment capacity, solvency, earnings, operating leverage, financial efficiency, liquidity, management, and industry standing. The Seventh Farm Credit District also evaluates the collateral position specific to an individual loan.

Methods

This research utilizes historical financial origination ratios based on underwriting standards currently in place within the Seventh Farm Credit District to predict the PD of a loan portfolio based on loan type (i.e., operating, term, real estate, commercial). The loans in the sample are categorized into ten risk classes.[2] The models determine the PD for each loan and map them to a similar PD grid of S&P publicly rated firms. The PD for each loan in the sample is calculated from an equation derived from binary logit regression models estimated from loan origination data. Once the mapping has been completed, the characteristics of the loans in each risk-rating class are known and lenders (AgriBank) can use these characteristics in providing benchmarks for each risk-rating class.

The financial standards include repayment capacity, solvency, liquidity, and collateral measures, which currently are the most complete data available for the objective criteria discussed above. Table 1 shows existing minimum underwriting standards (see Roessler for data definitions). AgriBank and the individual associations have set minimum guidelines for each of the above measures. At the time of loan application and analysis, loan officers evaluate the borrower's financial position (utilizing recent balance sheet and historical earnings trends as well as projections for the next year) and determine if they meet all or most of the minimum underwriting standards. The standards are generally applied to projected performance, although the projection must reasonably reflect the borrower's recent performance. These factors are considered along with the firm's management and previous repayment histories, either with the Farm Credit institution or other

Table 1. Seventh Farm Credit District minimum underwriting standards (as of 12/31/02)

Standard	Minimum
Repayment capacity	115%
Owner equity	50%
Working capital	15%
Loan to value	65% for real estate loans; 75% for term and operating loans

Source: Roessler.

creditors. If the borrower meets all of the underwriting standards, the loan will generally be approved and the financial data along with other variables are the basis for a risk rating being assigned to that particular loan.

If the borrower does not meet all of the underwriting standards, there must be significant mitigating or other adverse external factors for the loan to be approved. Such factors include very low collateral penetration (loan amount is low relative to the amount of collateral being offered), stress in the industry or local geographic area, a government guarantee, or additional controls placed on the borrower. These mitigating factors are often not measurable, and thus create a weakness in statistical models that may be utilizing historical financial data as a predictor of default. The financial position of the borrower, along with the mitigating strengths of the loan structure, is taken into consideration when determining a risk rating.

Data

Data were obtained from the Seventh Farm Credit District's loan accounting data base. During the time period studied (1995–2002), the Seventh District was made up of eleven states: Arkansas, Illinois, Indiana, Kentucky, Michigan, Minnesota, Missouri, North Dakota, Ohio, Tennessee, and Wisconsin. The District began using this data system in 1995. The data collected were quite extensive and were categorized both by customer level and loan level. The customer level data contained customer number, total assets, total liabilities, industry, gross farm income, and net income. The loan level data contained customer number, loan number, volume, commitment amount, date of origination, maturity date, loan product code, repayment capacity percentage, owner equity percentage, working capital percentage, loan to value percentage, payment frequency and type, delinquency history, collateral code and description, estimated market value of collateral, and credit bureau score(s). Not all elements were available for all customers and/or loans, thus standards were established for keeping and discarding loans.

When a loan is approved and "booked" into the data system, certain information on the customer and the loan is captured. Some of that information is updated over time, such as collateral values and earnings and balance sheet information; however, loan origination data are not changed. Loan and customer information are updated in the database at random intervals after origination and not all loans are updated. Because there was no way to determine the timing of data for the loans in the database (i.e., did total assets represent assets at origination or at some point after), loan origination data were considered the most consistent source of information. Thus, origination ratios (repayment capacity percentage, owner equity percentage, and working capital percentage) were used as the main variables in our model. Commitment amount was also considered. Commitment amount represents the total principal approved for that loan, whether or not it was disbursed. One of the secondary objectives was to determine if loan size is related to the PD.

Default is defined as a payment being ninety days or more past due at least once since origination. Ninety days or more is a typical definition of default in the market. This variable takes the value of (1) if default has occurred and (0) otherwise.

Table 2. Adjustments to origination ratio data

Ratio	Mean	Standard Deviation (SD)	Mean + 3SD	Mean − 3SD
Repayment	154.98%	74.98%	379.86%	−70.03%
Owner equity	64.13%	17.56%	116.82%	11.44%
Working capital	41.38%	88.0%	305.38%	−222.62%

Data Screening

Loan level is the unit of analysis. First, loans approved using scorecard systems were separated from those that were traditionally approved. Because scorecard-approved loans do not have traditional origination ratios, they cannot be analyzed in the same model as traditionally approved loans. We eliminated loans without origination ratios (repayment capacity, owner equity, and working capital) or that had only one of the three ratios.

As with all database systems, data integrity is an issue. Not all of the loan data had been input correctly at origination and some of the origination ratio values were obvious outliers. Outliers can occur when the denominator of a ratio gets close to zero. The outlying values for repayment capacity, owner equity, and working capital were adjusted back to three times the standard deviation above and below the mean of all loans for each ratio thus retaining these loans in the data base. For example, if owner equity percentage was negative 800 in the original data, the mean was 50 and the standard deviation was 10, the negative 800 was adjusted to 20. This method will capture 99% of all loans if the data are normally distributed and indeed, less than 1% of all of the loans were adjusted using the means and standard deviations in table 2. A financial institution must rate all loans in the portfolio and as such, outliers must be rated. To examine the effect of this adjustment, the econometric models were reestimated with these observations deleted and compared to those estimated with the adjusted observations. The results were robust to this adjustment.

Table 3 presents descriptive statistics for these variables that are presented in the All-Loans section. There were 157,853 traditionally approved loans of which 2,892 had defaulted, yielding a total default percentage of 1.83%. Table 3 also reports the descriptive statistics for loans categorized as nondefault and default. The origination ratio means for all defaulted and nondefaulted loans are above the minimum underwriting standards; however, for each ratio, the means for those loans that did not default are higher. The minimums and maximums are explained by the adjustments made to the loans as discussed above except for the maximum owner equity reported for the loans that did default. Owner equity has the smallest variation in values. Working capital has the largest range in values. The mean commitment amount for the loans that defaulted is higher than for the loans that did not default.

Econometric Methods

A binomial logit model was used to estimate the PD. This technique uses maximum likelihood estimation to "choose coefficient measurements that maximize

Table 3. Distribution of variables used in analysis of probability of default for all traditionally approved loans, nondefaulted loans, and defaulted loans

Variable	Mean	Standard Deviation	Minimum	Maximum
All loans				
Repayment capacity	153.04%	68.08%	−70.03%	379.86%
Owner equity	64.16%	17.48%	11.44%	116.82%
Working capital	36.75%	67.09%	−222.62%	305.38%
Commitment amount	$149,104	$512,683	$5	$45,523,844
Observations	157,853			
Nondefaulted loans				
Repayment capacity	153.26%	68.0%	−70.0%	379.9%
Owner equity	64.3%	17.4%	−11.4%	116.8%
Working capital	37.1%	67.2%	−222.6%	305.4%
Commitment amount	$148,806	$501,791	$5	$45,523,844
Observations	154,961			
Defaulted loans				
Repayment capacity	141.0%	71.9%	−70.0%	379.9%
Owner equity	56.1%	19.3%	11.4%	100.0%
Working capital	19.5%	58.0%	−222.6%	305.4%
Commitment amount	$165,086	$924,635	$900	$45,000,000
Observations	2,892			

the likelihood of the sample dataset being observed" (Studenmund). The maximum likelihood method is consistent and asymptotically efficient, and with very large samples produces normally distributed coefficient estimates which allow the user to employ typical hypothesis testing techniques (Studenmund).

To determine whether or not specific loan origination data gathered by the Seventh Farm Credit District are statistically significant in predicting default, the following empirical model was used:

$$\text{(2)} \quad \text{Ln(probability of default/[1} - \text{probability of default])} \\ = \beta_0 + \beta_1\text{(repayment capacity percentage)} + \beta_2\text{(owner equity percentage)} \\ + \beta_3\text{(working capital percentage)} + \beta_4\text{(commitment amount)} + \varepsilon_I.$$

PD Results

The regression results indicate that all of the variables are statistically significant at the 99.99% level (table 4). The coefficients for the origination ratios are negative, as expected.[3] The relatively small size of the coefficients indicates that sizeable changes in variable levels would be needed to have a large impact on the PD. This relationship is illustrated graphically later. The chi-square value for the likelihood ratio (732.8) indicates that this model is statistically significant in

Table 4. Logistic regression results of probability of default for all loans in the portfolio

Variable	Estimate	Standard Error	Chi-Square	Pr > Chi-Square
Intercept	−2.3643	0.700	1,139.3093	<.0001
CDRC	−0.00135	0.000329	16.7658	<.0001
OE	−0.0217	0.00108	408.8323	<.0001
WC	−0.00399	0.000420	90.0451	<.0001
Likelihood ratio (H_o: $B_i = 0$)			732.7583	<.0001
−2 Log L	28,132.288			
Number of observations	157,853			
Number defaulted	2,892			
Percentage defaulted	1.83%			
Volume defaulted	$477,427,989.00			
Percentage volume defaulted	2.03%			

Goodness of fit statistics

R^2	Max Rescaled R^2	Correct (%)	Sensitivity (%)	Specificity (%)
0.0046	0.0277	65.4	55.6	65.6

predicting whether or not a loan will default. The null hypothesis that all of the variables are zero is rejected.

The percentage defaulted refers to the actual number of loans that defaulted as a percentage of all of the loans in this model (table 4). The volume was calculated by multiplying the number of defaulted loans by the mean commitment amount of defaulted loans.[4] Percentage volume defaulted was calculated by dividing the volume defaulted by the sum of total volume of loans made obtained by multiplying the average loan size by the number of loans.

Using a cutoff of 2% for classifying default, the model correctly predicted 65.4% of the loans that would default. The model predicted correctly 55.6% of the loans that did default, and 65.6% of the loans that did not default, as indicated by the sensitivity and specificity outcomes, respectively.

The 2% cutoff was chosen because it was close to the mean default rate of 1.83%; however, a number of different cutoffs could have been chosen. Table 5 reports the percentage correctly predicted to default based on cutoffs ranging from 1% to 10% for the base model (table 4). The percentage correctly predicted to default or not is sensitive to cutoff changes in the 1–6% range. Once the cutoff reaches 7%, the percentage correct and sensitivity and specificity percentages do not change significantly.

To test this model further, a fourth variable, commitment amount is added to determine whether larger loans have higher probabilities of default. However, commitment amount is not statistically significant in predicting the PD for a loan

Table 5. Percentage correct predictions based on differing cutoffs

Cutoff Percentage	Correct (%)	Sensitivity (%)	Specificity (%)
1	18.0	90.1	16.7
2	65.4	55.6	65.6
3	89.1	23.7	90.3
4	95.3	10.3	96.9
5	97.1	4.4	98.8
6	97.9	1.7	99.7
7	98.1	0.2	99.9
8	98.1	0.0	100.0
9	98.2	0.0	100.0
10	98.2	0.0	100.0

or portfolio of loans. Thus, loan size does not significantly influence whether or not a loan will enter default status despite the higher average size of default versus nondefaulted loans table 3.

Loan Type Results

The loans can further be analyzed based on the type: operating, term, real estate, rural residence, and commercial. Commercial loans are those with a commitment amount greater than $5 million. The other loan types are based on the collateral used and are less than $5 million. For each type of loan,[5] the independent variables (capital debt repayment capacity [CDRC], owner equity, and working capital) were regressed on the default category. All of the origination ratios are statistically significant at the 99% level and the signs for all of the coefficients were negative for the operating loan model (table 6). Although all of the signs on the variable coefficients for term loans are negative, only owner equity and working capital are statistically significant at the 95% level.

While the parameter estimates are negative for each of the variables predicting default on real estate loans, only the owner equity origination ratio is significant at the 99% level (table 6). The working capital ratio is significant at the 95% level, and the repayment capacity ratio is not statistically significant. Because real estate loans are typically long-term, most lenders place greater emphasis on the borrower's long-term asset base when evaluating these loans. Therefore, a borrower's financial position as it relates to repayment capacity and working capital is not as statistically significant as the owner equity ratio in predicting the default of real estate loans.

The very low value of the -2 Log L measure (table 6) indicates this model is not statistically accurate in predicting the PD of rural residence loans. Only four loans of the 934 observed defaulted, thus there are not enough degrees of freedom for an accurate model to be estimated. This is not unexpected, as rural residence loans do not utilize traditional farm underwriting standards but more often use a scorecard analysis. The commercial loan model also is not statistically

Table 6. Logistic regression estimates of probability of default for traditionally approved operating, term, real estate, rural residence, and commercial loans

Variable	Operating	Term	Real Estate	Rural Residence	Commercial
Intercept	−2.3055***	−2.1594***	−2.4611***	−3.1546	−1.6958
CDRC	−0.00155***	−0.00082*	−0.00160	−0.00600	-4.7×10^{-6}
OE	−0.0186***	−0.0249***	−0.0316***	−0.0150	−0.0405
WC	−0.00288***	−0.00387***	−0.00167**	−0.00949	−0.0123
−2 Log L	10,888.09***	11,971.70***	4,874.22***	47.62	50.85
Number of observations	51,742	60,093	44,876	934	201
Number defaulted	1,159	1,271	452	4	6
Percentage defaulted	2.24%	2.12%	1.01%	0.43%	2.99%
Volume defaulted	$170,243,354	$125,405,210	$101,116,911	$712,876	$79,949,628
Percentage volume defaulted	2.63%	2.44%	1.04%	0.53%	3.89%

Goodness of fit statistics

Model	R^2	Max Rescaled R^2	Correct (%)	Sensitivity (%)	Specificity (%)
Operating	0.0040	0.0209	46.2	70.2	45.6
Term	0.0057	0.0308	55.8	63.7	55.6
Real estate	0.0040	0.0379	91.5	21.9	92.3
Rural residence	0.0043	0.0792	99.1	0.0	99.6
Commercial	0.0153	0.0651	37.8	33.3	37.9

*, **, *** Represent statistical significance at the 10%, 5%, and 1% level, respectively.

significant and cannot be relied upon to accurately predict the PD of commercial loans. The majority of commercial loans are not underwritten with traditional standards. Instead, they are underwritten with more typical commercial ratios such as fixed charge coverage and interest coverage ratios utilizing earnings before interest, taxes, and depreciation allowances (EBITDA) cash flow measures, which are utilized by other lenders in the industry.

To test whether the parameter estimates are statistically different by loan class, the log likelihood statistics are summed across the classes and subtracted from the likelihood statistic for the original model (table 4). The difference is distributed as a chi-square statistic with $(n - 1) * k$ degrees of freedom where n is the number of classes and k is the number of parameters estimated in each equation. The results of the chi-square distribution test was 160.43, which results in a statistical probability of 6.75 E-24, indicating statistically significant differences between the model that separated the loans into different types and one that estimated default for all loans.

Changes in PD

Figures 1–3 portray the marginal effects on the PD for the respective underwriting variables for each of the loan types where the logit regression was statistically significant. Each graph shows the pattern of the PD change as one of the underwriting variables ranges between one and two standard deviations above and below its mean, with the other two held at their mean levels.

As CDRC increases, indicating an increasing ability to repay debt, the PD for each type decreases (figure 1). The relationship between default and CDRC is

Figure 1. Change in probability of default as CDRC increases for traditionally underwritten loans

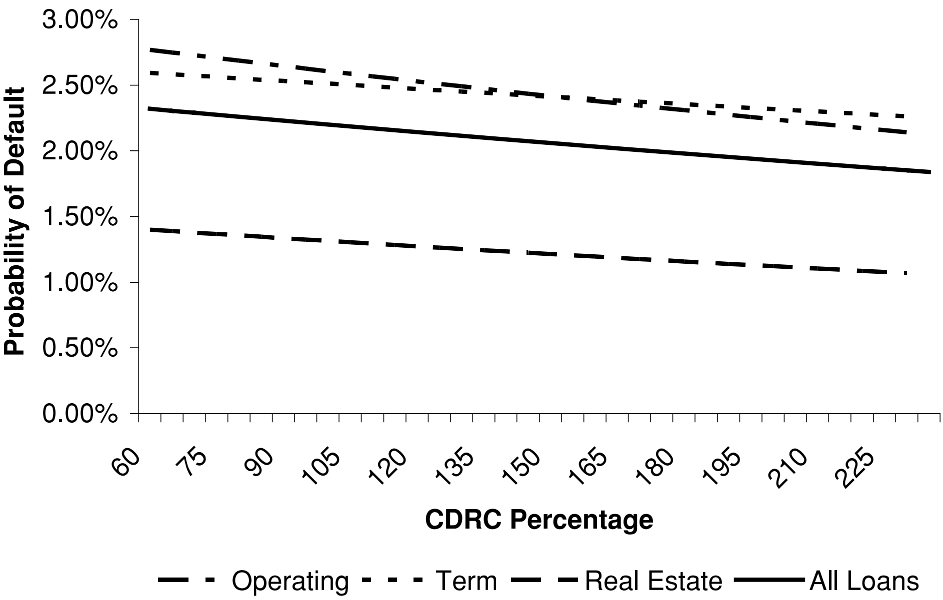
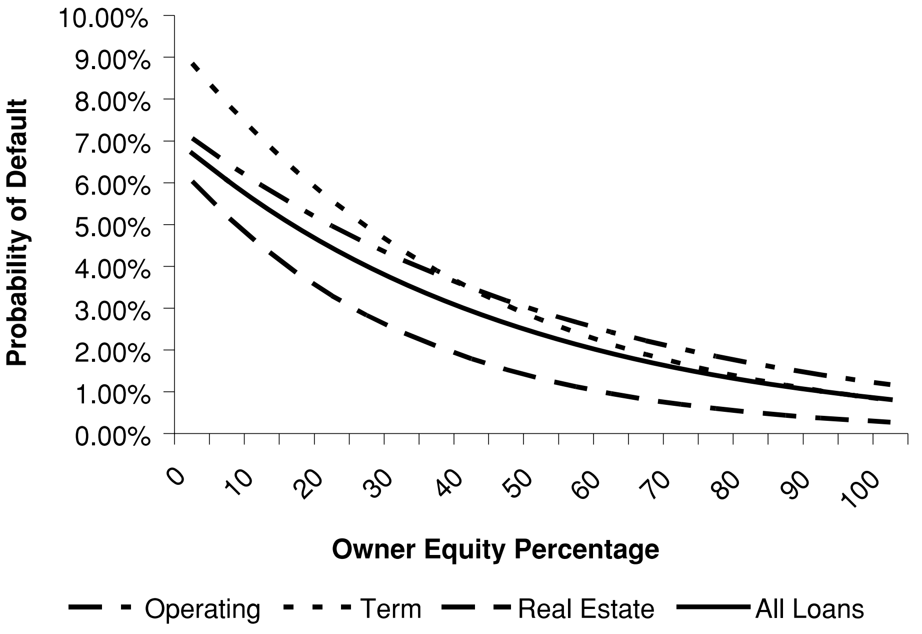


Figure 2. Change in probability of default as owner equity increases for traditionally underwritten loans



more flat for the real estate and term loans indicating that a lender should place greater emphasis on the repayment capacity ratio for operating loans than for real estate and term loans.

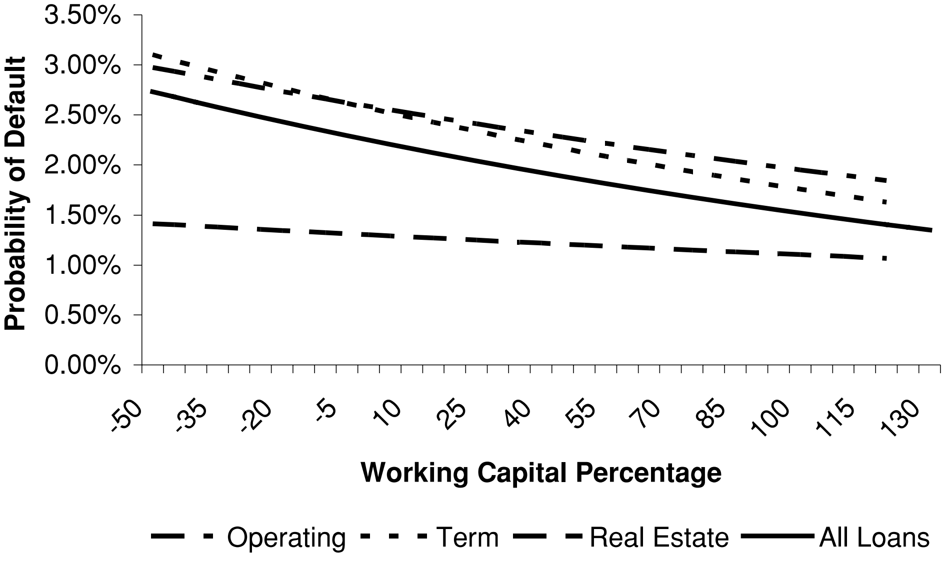
As owner equity increases from 0% to 50%, the PD decreased at a decreasing rate for all categories (figure 2). Changes in the owner equity ratio have the smallest impact on operating loans and the all loans category, a greater impact on term loans and the greatest impact on real estate loans.

As working capital increases from −50% to 130%, the PD again decreases at a decreasing rate (figure 3). Working capital has the least effect on the PD for real estate loans and has the greatest impact on term loans as measured by the slope of the curves. Changes in the owner equity percentage have a larger effect on the PD than changes in the CDRC or working capital ratios.

Risk-Rating Results

The next step in the analyses uses the logit results (table 4) to map the 157,853 loans into ten risk-rating classes based on an adaptation of S&P reported probabilities of default for their 17 risk-rating categories.[6] Roessler provided definitions for each of the rating classes, drawing upon guidelines by a Farm Credit System risk-rating committee. Included in these guidelines were directions for consolidation of several of the S&P categories and assignment of loans to the consolidated categories in this study based on comparisons between each loan's estimated PD and the PD intervals defined for each rating class.

Figure 3. Change in probability of default as working capital increases for traditionally underwritten loans



The S&P PD grid consists of diverse industries and firms. Because of this, it may be questioned as to why this particular grid is being used to analyze loans in the Seventh Farm Credit District, which only makes loans to borrowers in one industry. Applying the S&P grid to agricultural loans has several benefits. It provides consistency with the definitions outlined by the Farm Credit System. It is a model that has been established, used, and validated and it provides consistency in the marketplace. When the Farm Credit System sells its loans and other securities in the market, a buyer will better understand the risk in that portfolio of loans if it is equated to common ratings such as those used by S&P. Finally, the public ratings data provide a benchmark (synthetic rating) to follow in the absence of comparable measures of the PD by risk ratings on agricultural loans.

Table 7 outlines estimates of the probabilities of default by S&P. KMV is a company that creates and provides software to Moody's and S&P to determine the probabilities of default of their portfolios. The software determines the estimated default frequency (EDF) and categorizes it based on those company's individual risk classes. Year-end 2001 data are used to construct the grid in table 7 (Lopez).

Table 8 shows the risk-rating classes, the FCS guidelines for adapting publicly rated companies by risk-rating class to the S&P ratings, and the probabilities of default assigned to the FCS risk-rating classes.

The model in table 4 was used to predict the PD of each loan based on its individual characteristics in order to determine the characteristics of the loans in each of the risk-rating classes by PD. The loans were then assigned to one of the above risk-rating classes based on their estimated individual PD. Table 9 reports the means for each of the variables in the models, the actual default of the loans in each class, the predicted PD of each class, and the number of loans in each. In

Table 7. The mapping of S&P credit ratings to KMV EDF values

S&P Rating	KMV EDF Value (%)
AAA	(0.00, 0.02]
AA+	(0.02, 0.03]
AA	(0.03, 0.04]
AA−	(0.04, 0.05]
A+	(0.05, 0.07]
A	(0.07, 0.09]
A−	(0.09, 0.14]
BBB+	(0.14, 0.21]
BBB	(0.21, 0.31]
BBB−	(0.31, 0.52]
BB+	(0.52, 0.86]
BB	(0.86, 1.43]
BB−	(1.43, 2.03]
B+	(2.03, 2.88]
B	(2.88, 4.09]
B−	(4.09, 6.94]
CCC+	(6.94, 11.78]
CCC	(11.78, 14.00]
CCC−	(14.00, 16.70]
CC	(16.70, 17.00]
C	(17.00, 18.25]
D	(18.25, 20.00]

Source: Lopez.

addition, table 9 shows the lower and upper 95% mean confidence interval for the loan class. The average predicted PD for all traditionally approved loans is 1.61%. This would suggest the average S&P rating would fall in the BB− category based on the grid in table 7.[7]

As expected, no loans are reported in classes one, two, and three. The definitions of these classes indicate that they are comprised almost exclusively of large national credits that have the highest ratings assigned by S&P. According to the results, no loans in the portfolio have the same characteristics as the loans publicly rated AAA, AA, or A by S&P.

The means of the origination ratios decrease as risk-rating class increases (table 9). The repayment capacity ratio drops significantly from class four to class five, and then increases again in class six. Not all of the means for the origination ratios meet the minimum required for approval based on current underwriting guidelines. For example, in class nine, the repayment capacity and owner equity requirements are met while working capital is not. The binary logit model allows the strengths of one variable to offset the weaknesses in others, thus the decrease and increase in the repayment capacity mean and the discrepancies between existing minimum underwriting standards used by AgriBank and the mean values are reported in table 9.[8]

Table 8. Mapping of S&P probabilities of default to proposed risk-rating classes

Risk-Rating Class	FCS Guidelines	S&P Probability of Default (%)	Proposed PD Guidelines (%)
1	AAA to AA	0.00–0.04	0.00–0.035
2	AA to A	0.03–0.09	0.035–0.08
3	A to A−	0.07–0.14	0.08–0.14
4	BBB+ to BBB	0.14–0.31	0.14–0.26
5	BBB to BBB−	0.21–0.52	0.26–0.52
6	BB+	0.52–0.86	0.52–0.69
7	BB+ to BB	0.52–1.43	0.69–1.145
8	BB to BB−	0.86–2.03	1.145–1.73
9	BB− to B+	1.43–2.88	1.73–2.88
10	B and down	2.88 and up	2.88 and up

Source: Lopez and Farm Credit System risk-rating guideline definitions (Roessler).

The majority of the loans are in class nine. The Basel Accord suggests that no institution have more than 30% of all loans in one class. It appears, based on this model that 35.4% of all loans in the portfolio would be in class nine and 11.6% in class ten. These are a large percentage of the loans that are expected to enter a problem or adverse risk-rating class at some point during the life of the loan. The number of loans in classes nine and ten suggest that the Farm Credit System may want to further increase the granularity of their risk ratings or adjust the category definitions.

Figure 4 illustrates the ability of the model to predict the PD of a portfolio of loans based on risk-rating class by comparing the actual percentage of defaulted loans to the predicted percentage of defaulted loans. Actual default is higher than predicted default for classes four through seven and lower than predicted default for class eight. Class nine indicates that actual and predicted defaults are very similar. Class ten represents the worst loans in the portfolio combined into one class for this study, so it is expected that the actual events of default for this class would be high.

Summary and Conclusions

The New Basel Accord may allow lenders to benefit from more accurate risk rating of their loan portfolio. Many financial institutions that serve agriculture may need to revise their risk-rating system to benefit from these new provisions. The goals of the Basel Accord are to tailor risk management to the financial institution and to increase segmentation of the loan portfolio by risk rating. Risk classification systems generally include dual ratings—one for the PD and one for the potential loss given that default occurs.

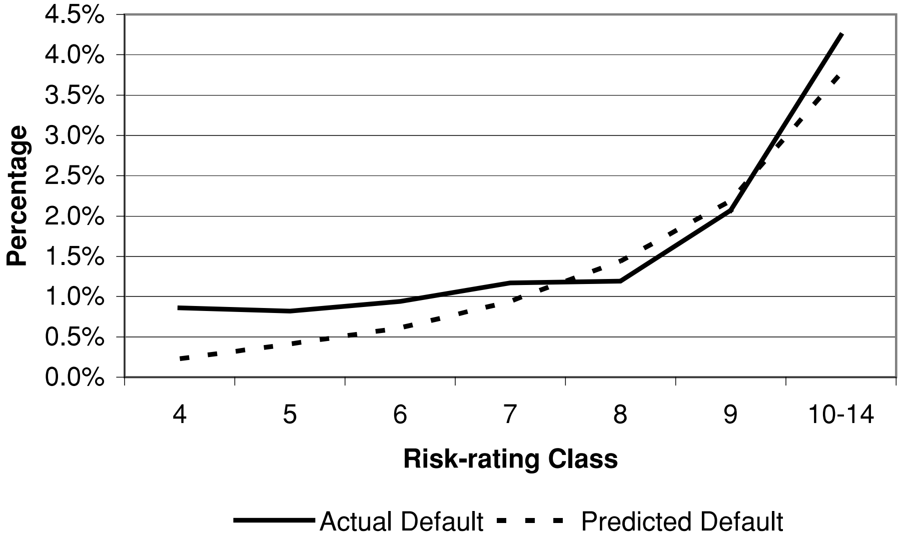
Three origination ratios: repayment capacity, owner equity, and working capital are important determinants of PD for our set of loans. Statistically significant differences exist between default models based on loan type (operating, term, real

Table 9. Characteristics of all loans in the portfolio by risk-rating class

Risk-Rating Class	CDRC Mean (%)	Owner Equity Mean (%)	Working Capital Mean (%)	Actual Default (%)	95% Lower Interval[a] (%)	Predicted Default (%)	95% Upper Interval[a] (%)	Actual Volume Default (%)	Loans	Percentage of Loans
1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0	N/A
2	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0	N/A
3	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0	N/A
4	331.5	95.1	302.2	0.86	0.18	0.23	0.30	0.13	348	0.2
5	197.6	84.2	265.2	0.82	0.33	0.41	0.50	1.03	4,619	2.9
6	220.6	82.8	157.2	0.94	0.52	0.61	0.72	0.75	4,885	3.1
7	193.7	82.5	61.2	1.17	0.85	0.94	1.04	0.81	26,788	17.0
8	151.3	71.3	28.7	1.19	1.35	1.44	1.54	1.10	47,120	29.9
9	134.6	55.7	13.2	2.07	2.09	2.19	2.30	2.46	55,852	35.4
10	121.5	34.1	−1.42	4.24	3.52	3.77	4.04	3.83	18,241	11.6

[a]Upper and lower confidence interval.

Figure 4. Actual versus predicted default for all loans by risk-rating class



estate, rural residence, and commercial), although the models for rural residence and commercial loans were not statistically significant. The default models are used to determine the PD to allow each of the 157,853 loans to be classified into a synthetic risk-rating category based on S&P credit ratings.

The analysis completed in this article is one of several ways that the data could be used to determine the PD of a portfolio of loans. Enterprise, geographic region, payment frequency, and payment type, for example, were not included in this analysis. The "seasoning" effect of loans was also not included in this analysis due to absence of information about the timing of default. As loans age, it is expected that the PD decreases.

Another area for future research includes the migration of loans from one risk-rating class to another. The current risk-rating system used by the Seventh Farm Credit District only has four acceptable risk-rating classes and five classes for problem loans, thus the range of migration is limited. However, as the District begins to compile updated data from borrowers and track the movement of loans from one risk-rating class to another, the migration patterns of the loans could be studied.

Once the new risk-rating procedures have been put into place and the Farm Credit System has gained comfort with its accuracy, the S&P guidelines as they correlate to each risk-rating class can further be used to indicate whether or not loans are being priced according to their risk. Also, the interest rate could be adjusted using the S&P ratings and the average interest rate spread on loans. A grid such as this could be used to ensure that the Farm Credit System is appropriately rewarded for the amount of risk they are taking with an individual loan or portfolio of loans. Finally, consideration of the LGD component will allow for complete implementation of the dual rating system in institutional capital management.

Acknowledgments

The authors acknowledge the helpful comments of two anonymous *Journal* reviewers. This project was supported by the National Research Initiative of the Cooperative State Research, Education, and Extension Service, USDA, Grant no. 2003-35400-12876.

Endnotes

[1]The mark-to-market approach considers the potential change of a loan's credit quality on its market value. When a loan is downgraded or enters into default status, financial institutions typically increase the interest rate associated with that loan as those institutions and financial market investors require greater risk premiums to compensate for the increased credit risk.

[2]See Roessler for more information on the definition of the risk classes.

[3]The discussion regarding the expected signs of the variables is found in the appendix with the variable definitions.

[4]This is equivalent to summing the defaulted volume on each of the loans.

[5]For more information on the means and variability of the independent variables by default category, see Roessler.

[6]The mapping could be done for each of the loan types using the regression results in table 6. These results are found in Roessler.

[7]The average S&P rating for the operating loans, term loans, and real estate loans were B+, BB−, and BB+, respectively.

[8]S&P ratings consider the overall credit worthiness of the securities issuer, including the combined effects of frequency and severity of default. In some cases, a given rating class could have a higher or lower probability of default than its neighbor, depending on differences in loss given default among these classes.

References

Barry, P.J. "Financial Management, Markets, and Models." Unpublished, University of Illinois, 2003.

——. "Modern Capital Management by Financial Institutions: Implications for Agricultural Lenders." *Agr. Rev. Finan. Rev.* 61(2001):103–22.

Basel Committee on Banking Supervision. "The New Basel Capital Accord." Basel, Switzerland: Bank for International Settlements, January 2003. Available at http://www.bis.org.

Betubiza, E., and D.J. Leatham. "A Review of Credit Assessment Research and an Annotated Bibliography." Deptartment of Agricultural Economics, Texas A&M University, College Station, June 1990.

Crouhy, M., D. Galai, and R. Mark. *Risk Management*. New York: McGraw-Hill, 2001.

Czuszak, J. "An Integrative Approach to Credit Risk Measurement and Management." *RMA J.* (June 2002):49–51.

Featherstone, A.M., and C.R. Boessen. "Loan Loss Severity of Agricultural Mortgages." *Rev. Agr. Econ.* 16(1994):249–58.

Joy, O.M., and J.O. Tollefson. "On the Financial Applications of Discriminant Analysis." *J. Finan. Quant. Anal.* 10(1975):723–40.

Katchova, A.L., and P.J. Barry. "Credit Risk Models and Agricultural Lending." *Amer. J. Agr. Econ.* 87(February 2005):194–205.

Khoju, M.R., and P.J. Barry. "Business Performance-Based Credit Scoring Models: A New Approach to Credit Evaluation." In *Regulatory Efficiency and Management Issues Affecting Rural Financial Markets*, E. LaDue and S. Allen, eds., Proceedings of North Central Region Project NC-207, Cornell Agricultural Economics Staff Paper 93–23, Department of Agricultural, Resource, and Managerial Economics, Cornell University, Ithaca, NY, December 1993.

Lopez, J.A. "The Empirical Relationship between Average Asset Correlation, Firm Probability of Default and Asset Size." Working Paper, Federal Reserve Bank of San Francisco, San Francisco, CA, 2002. Available at http://www.frbsf.org/publications/economics/papers/2002/index.html.

Miller, L.H., and E.L. LaDue. "Credit Assessment Models for Farm Borrowers: A Logit Analysis." *Agr. Finan. Rev.* 49(1989):22–36.

Novak, M.P., and E.L. LaDue. "An Analysis of Multiperiod Agricultural Credit Evaluation Models for New York Dairy Farms." *Agr. Finan. Rev.* 54(1994):55–65.

Roessler, L.M. *Determining the Probability of Default of Loans in the Seventh Farm Credit District*. Unpublished MAB thesis, Kansas State University, 2003.

Siddiqi, N., and S. Klein. "Estimating Probability of Default via External Data Sources: A Step Toward Basel II." *RMA J.* (December 2002–January 2003):68–72.

Studenmund, A.H. *Using Econometrics A Practical Guide*, 4th ed. MA: Addison, Wesley, Longman, 2001.
Turvey, C.G., and R. Brown. "Credit Scoring for Federal Lending Institutions: The Case of Canada's Farm Credit Corporations." *Agr. Finan. Rev.* 50(1990):47–57.
Yu, T., T. Garside, and J. Stoker. "Effective Credit Risk-Rating Systems." *RMA J.* (September 2001):38–43.