Measuring the Credit Risk of AgriBank Loans under the New Basel Capital Accord: A Logit Regression Approach

Enrique Hennings and Bruce Sherrick
JEL code: Q21; G21

Contact Information:

Enrique Hennings PhD Student, Department of Agricultural and Consumer Economics University of Illinois 1301 W. Gregory Dr. Urbana, IL 61801 ph: (217) 333-2658 fax: (217) 333-2312 emhennin@uiuc.edu	Bruce J. Sherrick Professor, Ag and Applied Finance University of Illinois 1301 W. Gregory Dr. Urbana, IL 61801 ph: (217) 244-2637 fax: (217) 333-2312 sherrick@uiuc.edu

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Portland, OR, July 29-August 1, 2007

Introduction

In 1999, the Basel Committee on Banking Supervision proposed a new set of capital standards that have collectively referred to as the Basel II Accord. This Accord has undergone an extensive consultative period, and the final version was released in 2004. The Accord has three fundamental pillars: 1) minimum capital requirements, 2) supervisory review, and 3) market discipline. The Basel II Accord retains the basic definition of bank capital and the market risk provisions of the 1996 Amendment, but substantially changes methods for evaluating credit risk, emphasizing frequency and severity of loan defaults. It also establishes guidelines to determine the economic capital needed by financial institutions to protect themselves from market, credit, and operating risk.

The Basel II Accord has generated a series of incentives for financial institutions to develop models that assess credit risk and to allocate their economic capital to different segments of their portfolios. Understanding the credit risk of these portfolios is of critical importance for lenders and for regulators, and provisions in the Basel II Accord establish guidelines to assist financial institutions in screening borrowers, and in determining how much credit should be extended and at what price given the institution risk bearing capacity. The proposal prescribes relationships between expected risk and the real return on investments, and recognizes the linkage between capital levels and the realized rate of return on equity (Gustafson et al, 2003). At the same time, better credit risk assessment allows regulators to be assured that institutions have adequate capital to tolerate prescribed levels of insolvency.

Agricultural lenders generally have less developed credit risk models than commercial banks, but more specific industry knowledge and informal credit evaluation methods. The implementation of the Basel II Accord will require agricultural financial institutions to change their credit risk models, implement new risk rating systems while collecting and gathering more information from each borrower; and to improve their technologies and processes for collecting and managing data. Currently, there are no standard risk-rating models used by agricultural financial institutions since the quality of data differs among institution in terms of loan types and approaches. This makes these type of institutions less competitive and vulnerable to major risk. Furthermore, given inherent difficulties in data acquisition, fewer advances have been made with respect to agricultural credit portfolios in general. Zech (2003) indicates that one of the primary limitations in using credit risk models that follow the Basel II Accord is the quality of the information. For example, the Internal Rating-Based (IRB) Approach mandates five years of Probability of Default (PD) and seven years of Loss Given Default (LGD) estimates to be available.

Furthermore, the Accord does not differentiate among sectors such as agriculture and industry, making it difficult for agricultural lenders to adopt the Accord as some of the assumptions used when modeling credit risk in commercial banking may not hold for agricultural financial institutions (Zech, 2003). As a consequence, developing of robust risk-rating model is essential for agricultural lenders to reduce credit risk by accurately determine borrowers' creditworthiness, as well to determine the economic capital that they are required to hold to avoid being overcapitalized and, therefore lose profitable investment opportunities. A parallel benefit of having an accurate credit risk assessment model is that it helps management to identify and remove borrowers that present excessive credit risk, and to determine how much credit should be extended and at what price to eligible borrowers. Credit risk assessment assists institutions in aligning their expectations for risk and return.

The agricultural finance literature on credit scoring has developed alternative statistical models and data sets that provide a useful description of several alternative statistical approaches on which to score credit (Chhikara, 1989). One of the objectives of applied research on credit risk assessment models is to identify good model characteristics. This study represents an empirical investigation of the probability of default within a loan portfolio to determine the set of characteristics to incorporate in a risk-rating model, as there is no uniformly adopted risk-rating model in the applied finance literature. The hypothesis behind the model is that the likelihood of repayment (default) and borrower creditworthiness can be determined by applying statistical models to measurable characteristics of borrowers at the individual transaction level following the New Basel Accord. Institutions that do adopt the New Basel Accord would have to standardize, at least partially, their data collection methods before employing a reliable internal rating based approach. Among major agricultural lenders, AgriBank is relatively well prepared to adopt the Basel II Guidelines as it has been collecting data and developing internal rating systems for more than 5 years already.

The general objective of this study is to asses the impact of the Basel II guidelines on measures of AgriBank credit risk. The specific objectives are the following:

- Estimate the probability of default for each loan in the portfolio of AgriBank while incorporating the guidelines of the New Basel Capital Accord,
- Test three alternative definitions of default and estimate different prediction models for default,
- Assess different data categories for the estimation of default probability (origination process, loan type, association, payment frequency)
- Identify models with superior predictive capabilities to quantify credit risk and estimate capital requirements for agricultural lenders (AgriBank) under the New Basel Capital Accord
- Measure time patterns of default and incorporate them into the predictive default models

New Basel Accord (Basel II)

In 1999, the Basel Committee proposed a new capital accord referred to as the Basel II Accord. This proposal undertook an extensive consultative period ending up with the committee releasing additional proposals for consultation in 2001 and 2003. The finalized Basel II Accord was released in 2004 and has three pillars: 1) minimum capital requirements, 2) supervisory review, and 3) market discipline. The Basel II Accord retains the definition of bank capital and the market risk provisions of the 1996 Amendment. It replaces the old treatment of credit risk, and it requires capital for operating risk.[1] The basic capital requirement for banks can be calculated dividing capital by the sum of credit risk, market risk and operating risk. This ratio needs to be greater or equal to 8%. As with market risk under the 1996 Amendment, banks have options as to how they value their credit risk and market risk. For credit risk, they can choose from among the Standardized Approach, the Foundation Internal Rating-Based (IRB) Approach, and the

[1] Operational risk refers to direct or indirect loss resulting from inadequate or failed internal processes, people, and systems, or from external events

Advanced-IRB Approach. For operating risk, the choices are among the Basic Indicator Approach, the Standardized Approach, and an Internal Measurement Approach.

The Basel II Accord seeks to improve the existing rules by aligning regulatory capital requirements more closely to the underlying risks that banks and financial institutions face and to promote a modern approach to capital supervision. The Basel II Accord requires banks to identify the risks they face today and the risks they may face in the future and to improve their ability to manage them. The changes in the financial conditions due to the Basel II Accord can be summarized as follows: 1) the new accord requires financial institutions to reevaluate their credit, market and operating risk in order to ensure the safety and soundness of capital management, 2) the accord requires new classes in the internal risk-rating and the use of a dual risk-rating system (Probability of default (PD) and potential loss given default occurs (PLD))[2], and 3) it suggests eight criteria for banks to measure when implementing a risk-rating system including the evaluation of a firm's repayment capacity, solvency, earning, operating leverage, financial efficiency, liquidity, management, and industry standing.

Impact of Basel Accord in Agricultural Financial Institutions

It is important to recognize that large and internationally financial institutions are required to adopt the Basel II guidelines. Agricultural financial institutions will also have to complying with the Basel II guidelines; therefore, it is important to recognize the necessity to review the impact of these measures for banks and firms that are subject to prudential banking or securities regulation. There are several issues that have to be addressed by financial institutions with respect to the treatment of double-default effects for covered transactions, short-term maturity adjustments; the internal ratings-based approach; improvements to the current trading book regime, especially with respect to the treatment of specific risk; and the design of a specific capital treatment for failed transactions and transactions that are not settled through a delivery-versus-payment framework (non-DvP). Agricultural finance institutions need to be able to respond to three basic questions in order to assess the implication of implementing the Basel II: 1) How much risk is present? 2) What should be the institutional tolerance to risk? and 3) How much capital is needed to offset that risk? Barry (2001) points out that answering these questions is necessary to manage risk effectively.

The implementation of Basel II could adversely affect the competitive position of agricultural finance institutions that do not adopt the advanced internal risk ratings-based (A-IRB). For example, the A-IRB may reduce the minimum regulatory capital and thereby lower the marginal costs of agricultural lending for adopters. The substitution effect of a decline in marginal costs at A-IRB agricultural finance institutions relative to non-A-IRB institutions may encourage A-IRB finance institutions to reduce prices or, increase quantities of agricultural lending. This could potentially lead to other financial institutions lowering the prices that they charge and could cut their market shares.

In the U.S., the adoption of Basel II could have significant competitive effects in the market for any assets or off-balance sheet activities in which the A-IRB agricultural finance institutions have substantially lower or higher capital requirements than the Basel I requirements to which the non-A-IRB agricultural finance institutions are subject. Thus, the adoption of the Basel II

---

[2] The probability of default indicates how frequently a loss may occur, while loss given default indicates the severity of the default

Accord can benefit agricultural financial institutions by allowing them to increase the segmentation of their loan portfolios by risk rating and result the better management of risk and capital provision. Under the Basel II proposal, banks with sufficiently sophisticated risk measurement and management systems can use their own internal systems to estimate key risk parameters that determine regulatory capital minimums.

The implementation of the Basel II will require agricultural financial institution to change their credit risk models implementing new risk rating systems while collecting and gathering more information for each borrower and improving their technologies and processes to collect data. Zech (2003) indicates that one of the biggest limitations in using a credit risk model that follows the Basel II accord is the quality of the information since for example for the IRB Approach, five years of PD and seven years of LGD are required. Furthermore, the accord does not differentiate between sectors such as agriculture, industry and so forth making it more difficult for agricultural lenders to adopt the Accord since some of the assumptions used when modeling credit risk in commercial banking may not be the same for agricultural financial institutions (Zech, 2003). Barry (2001) indicates, "Agricultural lending is characterized by the cyclical performance of farm business, length, seasonal production pattern, high capital intensity (real state), extensive leasing of farmland, participation in government programs, and annual versus monthly payments on intermediate and long-term loans". Barry (2001) also argued that even thought loan losses are infrequently in agricultural lending, the losses can be significant and be highly correlated across production units, geographic areas and time due to the lack of diversity among farms. He also pointed out that risk and capital adequacy problems might arrive through higher lending cost for small-business and reduced earning from lending rather than loan losses.

The successful implementation of the Basel II accord depends on both lenders and regulators understanding the special circumstances of each type of business. This is particularly true in the case of credit risk management in agricultural lending institutions.

Credit risk assessment models in Agriculture

The purpose of a credit risk assessment is to evaluate the risk exposure of lenders and shareholders at the transaction and portfolio levels given the fact that high credit risk might affect negatively the rate of returns on a loan portfolio. Gustafson, et al. (2003) argue that credit risk model seek to asses the borrower's creditworthiness by applying econometric methods to measurable characteristics at the individual transaction level. Credit score models are used to predict the likelihood of loan repayments or loan defaults incorporating specific measurable factors that cluster borrowers in different risk groups that reflect their creditworthiness. These models are also useful to assist in loan approval decisions, loan pricing and to meet regulatory requirements (Ellinger et al., 1992). Gustafson, et al. (2003) indicates that one of the limitations of the credit risk models is that they usually do not account for misclassification costs or, the degree of defaults in a loan portfolio.[3] Traditionally, credit risk assessment is based on quantitative assessment tools such as credit scoring models as well as lenders' own qualitative criteria concerning the management capacity and repayment capacity of the borrowers. However, in recent years, quantitative risk evaluation models have become increasingly important and have

[3] Misclassification cost refers to Type I and Type II errors, where Type I occurs when a bad borrower is consider acceptable while Type II error occurs when a bad borrower is classified as acceptable.

evolved into complex models that are more accurately able to predict borrower behaviors and better assess management capacity and repayment capacity.

The attention to credit score models by agricultural lenders has increased during the 1980's and 1990's due to the large number of loan defaults by failure farms (Turvey, 1991). Prior to the 1970's, credit risk evaluation relied primarily on subjective assessments such as how well the lender knew the farmer and the size of loan (Gustafson et al., 2003). Since then, agricultural finance institutions and researchers have developed quantitative and less subjective decision criteria to evaluate borrower risk. Allcott (1985), Kohl and Forbes (1982), Kohl (1987) and Tongate (1984) developed methods that include comprehensive financial measures such as liquidity, solvency, profitability, efficiency, repayment capacity, and management ability.

Fisher and Moore (1987) proposed a logistic function with fewer explanatory variables to address the problem of multivariate normality of explanatory variables. Miller and LaDue (1989) used the logistic method with borrower repayment based on ratio theory data instead of subjective lender information. They also accounted for covariance between regions and farm types and considered costs of misclassification. Gustafson et al. (2003) reported that Lufborrow, Barry and Dixon (1984) were the first ones to link credit assessment with loan pricing using a probit model and claimed that their results enabled lenders to advise borrowers to improve their credit score. Barry and Ellinger (1989) built up a multi-period model that endogenized credit, investment and loan pricing decisions. LaDue (1990) developed sixteen financial ratios in cooperation with agricultural bankers in an effort to standardize the information requirement by lenders resulting in the creation of the Farm Financial Standards Taskforce. Chhikara (1989) and Gustafson (1988) supported to implementation of portfolio analysis type of models to assess credit risk since most previous studies limited themselves to only default rates.

Turvey (1991) compared and contrasted the performance of four alternative credit evaluation models such as discrimant analysis (DA), probit, logit, and linear probability (LPM), and found that the four methods have similar classification accuracies. However, the LPM and DA techniques have some estimation problems when correcting for heteroscedasticity (LPM) and the assumption of normally distributed random variables (DA). On the other hand, the probit and logit are less restrictive in terms of the underlying distributional assumption. Gustafson et al. (2003) argue in favor of the change in the direction of credit risk research around 1980's, recognizing the asymmetric information problem. Borrowers know more about their credit risk than lenders because they are more familiar with their business, financial position and repayment intentions. Gustafson et al. (2003) argued that asymmetric information is the source of adverse selection and moral hazard problems[4] and point out that lenders have responded to these problems by focusing on relationship information including borrower motivation, commitment, and intentions. Ellinger et al. (1992a) found that overall model consistency is better when predicting low performance loans than when predicting high performance loans; however the differences when using subjective measures such as management affect negatively the consistency among models. Model consistency decreased as more detail analysis such as type of loan, purpose of loan and type of borrowers were included (Gustafson et al., 2003). Data

[4] Adverse selection occurs when the lender is unable to distinguish between high- and low-risk borrowers. Moral hazard is the ability of a borrower to use loan funds to engage in activities that are riskier than the lender anticipated. Only the borrower can know their true intentions for the loaned funds and their future ability and willingness to repay the loan

limitations in terms of number of year available and consistency between institutions make it hard to validate the credit score models since there is no uniform model used by lenders. Gustafson et al. (2003) conclude that "it is not surprising that a significant part of the recent agricultural finance literature has focused on the potential for improving the consistency (or robustness) of the models".

This review reveals that there is a lack of concern regarding which models apply to different loan types or business, which variables are most important to use to predict loan performance, or which models are most suitable for specific type of data. Miller and LaDue (1989) found that no specific factor has consistently been used in the credit-scoring framework and that there is a huge variety among models in the literature. Variables that have been considered by authors to measure credit risks include: borrower liquidity, leverage, collateral, repayment ability, repayment history, profitability, efficiency measures, farm type, geographic region, and financial and non-financial factors.

Gallagher (2001) indicates that a prediction model without non-financial variables would suffer from model misspecification. He combined non-financial characteristics of loans such as manager, lender experience, and the use of a financial adviser to see which agribusiness loans perform better. Zech and Pederson (2003), using a linear and logistic regression models, identified the debt-to-asset ratio as a major predictor of repayment ability. They also found that factors such as family living expenses and farm financial efficiency are excellent predictors of overall financial performance, however the authors indicate that the lack of robustness of risk-rating models is evident confirming that models developed for specific periods may not be used for subsequent periods. Katchova and Barry (2005) developed models for quantifying credit risk in agricultural lending. They calculated probabilities of default, loss given default, portfolio risk, and correlations using data from farm businesses. They showed that the expected and unexpected losses under the Basel II Accord critically depend on the credit quality of the loan portfolio and the correlations among farm performances.

Another important change in the direction of credit risk research is the recognition that assuming a specific distribution to model credit risk restricts the applicability of the model and may undermine its results. Thus, a proliferation of nonparametric approaches such as recursive partitioning algorithms and mathematical programming techniques have been implemented and compared to those obtained with parametric statistical approaches (Gustafson et al., 2003). For example, Ziara et al. (1995) found that either mathematical programming techniques or statistical models performed equally well, and that mixed integer-programming models perform better than parametric models. One clear advantage of non-parametric model is that they can fit several distribution functions and mathematical function, and at the same time, they are flexible enough for sensibility analysis. Furthermore, when the data sample is small or the data is contaminated, non-parametric model may behave better (Gustafson et al., 2003).

Credit risk migration analysis has been also used to model credit risk in agriculture. Phillips and Katchova (2004) used annual migration rates of credit scores to test path dependence conditional on the business cycle. They found that upgrades to credit scores are followed by downgrades. In addition, upgrades are more likely to occur in an economic expansion phase while downgrades are more likely in economic recession. Escalante et al. (2005) used a probit regression for determining path dependence by accounting for demographic, financial, and macroeconomic variables such as farmland value, aggregate money supply, the S&P 500 Index, and long-term agricultural interest rates. Using data from agricultural banks, Gloy et al. (2006) conducted a credit risk migration study. He used a logistic regression model to assess credit downgrades and found that the probability of a downgrade is different among lending institutions and when farms are facing a downgrade economic cycle, they are more likely to downgrade.

Engelman, Hayden and Tashe (2003) provide an empirical comparison of the logit model and the Altman's score Z-core (discriminate analysis on a large sample of SMEs. They show that the logit model significantly outperforms Altman's score Z-core in terms of rank ordering (the ROC coefficient). Featherstone et al. (2006) used data on loans in the Seventh Farm Credit District Portfolio to model loan defaults and map them into S&P probability of bond default for publicly rated companies. The authors employed a binary logit model to calculate the PD for each loan before mapping it to the S&P publicly rated firms. They discovered that repayment capacity, owner equity, and working capital origination loans are determinants of PD.

Data employed

The data employed for this study was obtained from AgriBank, a lending institution in the Farm Credit System (FCS).5 AgriBank is a member-owned cooperative that provides credit for agriculture to the shareholders in Arkansas, Illinois, Indiana, Iowa, Kentucky, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Tennessee, Wisconsin, and Wyoming. AgriBank was established in 1992 when the Farm Credit Banks of St. Louis merged with the Farm Credit Banks of St. Paul. Subsequent mergers took place in 1994 and 2003 with the Farm Credit Bank of Louisville and with AgAmerica Farm Credit Bank, respectively. AgriBank has more than $47 billion in assets and around $2 billion in equity.

The database obtained from AgriBank contains information from two sources: 1) customer financial information (371,260.00 observations) which includes data such as income, asset, collateral, financial ratios and balance sheet information, and 2) the default database (460,274.00 observations) which contains loan level information such as loan amount, repayment capacity, working capital, payments information and dates, etc. The study utilizes AgriBank's loan level data rather than customer level data. The data based was cleaned by removing master notes with association and participations, negative values or non-commitment (unless scored loans) and leases. Records with blank or missing underwriting fields (liquidity, solvency and repayment capacity) were also removed from the database.

According to the guidelines of the New Basel Capital Accord, four conditions need to be met for a loan to be considered in default: 1) the borrower could not pay in full her obligations; 2) the loan is past due for 90 days or more; 3) a credit loss event happened, for example a write off, specific provision, debt restructuring, interest or fees, etc; 4) the borrower has filed for bankruptcy. The database does not contain a specific variable for default; therefore, in this study, default was defined in three different ways. First, a loan that has been 90 days past due or more at least one time since origination was considered to be in default. Some loans that appear under the category of 120 days past-due were not recorded in the 92 days past due category, thus

[5] The Farm Credit System (System) is a network of borrower-owned lending institutions and related service organizations serving all 50 states and the Commonwealth of Puerto Rico. These institutions specialize in providing credit and related services to farmers, ranchers, and producers or harvesters of aquatic products.

to avoid excluding them from the default category, all loans that were more than 90 days past due were considered delinquent. This measure of default is consistent with the literature and is used by the industry as well. The second default measure is similar to the first one with key difference being that default loans have to be non-accrual loans as well. The accrual code points out those loans that were classified as non-accrual without first being classified in any other past-due category. The third definition of default loan extends this second definition to include those loans in which the borrowers has filed for bankruptcy.

All three variables (definitions) take a value of one if the default has occurred, or a zero if one has not. The third default definition follows all the New Accord's conditions since the definition used by AgriBank for non-accrual loans status is consistent with the concept of unlikeness to pay obligations (i.e. loans are placed in non-accrual status when the principal or interest is delinquent for 90 days or more). Thus, definition three satisfies conditions one and two of the Accord. The third condition is fulfilled by default and fourth condition is captured by the Litigation_code variable that indicates whether a loan is in litigation (one category within this variable is bankruptcy). Given these three conditions of default, 217 observations were deleted from the database since they did not include information that indicated if they had, at any point in time, been in default (i.e. they included missing observations). Furthermore, data entries that did not contain information about the origination process were assessed for abnormalities and a further 3354 observations were deleted. The resulting dataset contained 456,703 observations.

For practical purposes, it was assumed that the probability of default related to the remaining lifetime of the loan. Furthermore, in order to be consistent, all the variables employed in this paper were based on origination data since some type of loans such as mortgage loans do not contain updated financial variables, or should not be updated at any time (i.e. the variable owner equity percentage).

Econometric Model

To calculate the probability of default, a binomial logit model is employed. This approach is consistent with Barry, Sherrick and Featherstone (2005) and Featherstone, Roessler and Barry (2006). The logistic regression method is used when the dependent variable is a dichotomous and the independent variables are of continuous and/or categorical. It is also used to determine the percent of variance in the dependent variable explained by the independent variables. Logit regression uses maximum likelihood estimation to estimate the probability of a certain event occurring by calculating the changes in the log odds of the dependent variable and not the changes in the dependent variable itself as is the case of OLS regression.[6] The maximum likelihood method is consistent and asymptotically efficient (large sample produces normally distributed estimates) which allows researchers to use typical hypothesis testing techniques (Cramer, 1986 and Eliason, 1993). Logistic regression does not assume a linear relationship between the independent and dependent variables, does not require normally distributed variables and do not assume homoscedasticity. However, it does require independent observations and that the independent variables are linearly related to the logit of the dependent variable.

[6] The logs odds refer to the natural log of the odds of the dependent variable occurring or not.

The logit regression has the advantage that it always returns a probability between zero and one overcoming the inherently unbounded problem of linear functions. Westgaard and Van der Wijst (2001) indicate that logit and probit models are especially useful to overcome the problem of the bounded dependent variable since they transform the probabilities in such a way that they are no longer bounded.[7] In the logit model, the upper bound is removed by transforming the probability *p* to the odds ratio of *p/(1-p)* while the lower bound is removed by taking the logarithm of the odd ratio: ln (*p/(1-p)*). These transformations make the logit model linear in log-odds, making the coefficients easier to interpret compared to the coefficients found in probit models. Thus, to determine the probability of default of any loan within the AgriBank portfolio of loans conditional on the realization of its variables (financial ratios), the following empirical model was used:

Eq. 1 $$\ln\left(\frac{\text{prob. of default}}{\text{1-prob. of default}}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_I x_I + \varepsilon_\iota$$

Where β_0 is a constant, β_i *(i=1…I)* are re the parameter estimates, x_n *(i=1…I)* are the explanatory variables and ε is an error term.

In this logit estimation, "default" and "non-default" variable is the dependent variable, and this variable is regressed against a set of dependent variables. The New Basel Accord suggests eight criteria for implementing a risk-rating system: 1) repayment capacity, 2) solvency, 3) earnings, 4) operating leverage, 5) financial efficiency, 6) liquidity, 7) management and 8) industry standing. Some of these criteria were employed as independent variables. In particular, liquidity was approximate by the working capital to average gross income ratio, solvency measured as owner equity as a percentage of the loan, and repayment capacity measured as the CDRC percentage at the loan level. Due to lack of data, the remaining criteria were not approximated. Later in the estimation process, other independent variables were included to explore different model specifications and additional information data up to five prior sets of customer-level financial statement were used to adjust for seasonality effects on the probability of default.

Predicted Probabilities

For a given set of values for the independent variables, the predicted probability in a binary regression model is defined as follow:

Eq. 2 $$\Pr(y = 1 \mid x) = \Phi(x\hat{\beta})$$

Where Φ is the cumulative distribution function for the logistic distribution with a variance of $\pi^2/3$.

Marginal Change

The marginal change is defined as follow:

[7] The difference between the logit and probit is that the probit uses a normal distribution approach instead of the logit distribution which has thicker tails than the normal distribution. In practice, there is no difference unless there are too many extreme observations in the sample (Green, 1993)

Eq. 3 $$\text{marginal change} = \frac{\partial \Pr(y=1 \mid x) = \Phi(x\hat{\beta})}{\partial x_k}$$

The value of the marginal change depends on the level of all variables in the model. It is typically computed with all variables held at their means; however, it may also be calculated by obtaining the marginal change for each observation in the dataset and then taking the average across all values.

Empirical Results for Traditional Loans

The logit estimation employed default and non-default as dependent variable and was regressed against several set of variables as indicated in the previous section. For loans evaluated under the traditional approach, the following independent variables were included: a) Capital debt repayment capacity as a percentage of the loan (CDRC). This variable indicates the borrowers' ability to generate sustainable earnings adequate to service debt on a continuous basis, b) original owner equity percentage on which the loan decision was made, c) working capital to average gross income ratio, which is considered a relative measure that indicates the adequacy of working capital compared to the size of business.

The econometric analysis was conducted in STATA using the logit command. Observations that did not have values for all variables (liquidity, solvency and repayment capacity) were removed as indicated above. In addition, observations that do not have values for at least two variables were dropped when estimating the regression leaving a sample of 438,312 observations. The regression results for the probability of default for traditionally evaluated loans are shown in Table 1. Using the first definition of default, the likelihood ratio has a chi-square of 743.46 with a p-value of 0.000 indicating that the model as a whole fits significantly better than the reduced model. The coefficients for the owner equity percentage and for the working capital as a percentage of the gross income are both significantly different from zero and negative at 99% level; for one unit increase in owner equity percentage, the log odds of being delinquent (vs. not being delinquent) reduces by 0.022. By taking the exponent of the coefficient (exp(-0.022)=0.97), it can be said that for a one unit increase in the owner equity percentage, the odds of default the loan change by a factor of 0.97. Additionally, an increase of one unit in working capital as a percentage of the gross income is associated with a 0.002 decrease in the log odds of default, or for one unit increase in the working capital as a percentage of gross income, the odds of default change by a factor of 0.99 (exp(-0.002)=0.99). To express these results in probability, the original probability of default (2.18/97.82=0.0222) multiplied by the odds ratio for the owner equity percentage (0.97) result in a new odd of the dependent variable of 0.0222*0.97=0.0217. Let x by the new probability such as $x/(1-x)=0.0271$ since the odds are defined as the probability of default (x) divided by the probability of non-default $(1-x)$. Solving for x results in the new probability of default. Thus for an original probability of default (2.18%), a logistic *b* coefficient of 0.97 means that a unit increase in owner equity percentage decreases the probability of default to 2.13% (a reduction of 0.05% from the original probability). Furthermore, an increase of one unit in working capital as a percentage of the gross income results in a decrease of the probability of default of 0.003% to 2.177. The coefficient for the repayment capacity is not significantly different from zero.

When using the second and third definition of default, the results are very similar as shown in Table 1. The likelihood ratio indicates that in both cases, the model fits significantly better than the reduced model. An increase of one unit in owner equity percentage decreases the log odds of default by 0.024 higher compared to the coefficient of 0.022 when using the first definition. In other words, an increase in one unit in owner equity percentage decreases the probability of default by 0.06% (from 2.60% to 2.54%). On the other hand, for one unit increase in working capital as a percentage of the gross income, the log odds of being delinquent (vs. not being delinquent) decreases by 0.001. In this case, the coefficient is lower than the coefficient 0.002 when using the first definition. In probability terms, a unit increase in working capital as a percentage of the gross income decrease the probability of default by 0.01%. The repayment capacity coefficient is again not significantly different from zero. When comparing the McFadden's Pseudo-R-Square of all three cases, it can be seen that the Pseudo-R-Square increased from 0.022 to 0.027 when using the second and third definition of default as dependent variables. Similar results are obtained when looking at the Cragg and Uhler's R-Square. However, this trend does not indicate anything. The Pseudo- R-Square cannot be compared directly because they came from three different frequency distributions. This measure will be helpful later.

Logit Regression Results after Correcting for Outliers and Influential Points

To asses which observations are outliers, standardize residuals were generated after the logit regression on the entire set of observations. Observations were sorted from the lowest to highest values of independent variable and index numbers were created for each one. Hosmer and Lemeshow (2000) indicate that there are no fast rules to indicate which residuals are "large". However, by sorting the observations according to the independent variables helps to search for problematic residuals. For example, for the working capital is clear that there are several observations that can be responsible for the lack of fit especially among the high-working capital observations. Moreover, the CDRC variable shows a disproportionate number of cases with large residuals relative to the others either for low, medium and high values while the owner's equity variable shows large residuals relative to the others mostly among high-value observations. Thus, it is necessary to identify these observations for further inspection. This was done by means of the index numbers, observations that have a Standard Pearson Residual greater than 10 were marketed as outliers.

To further analyze the cases with "large" residuals and see if these points have a strong influence on the estimated parameters, the Pregibon's measure (dbeta) was employed. There are several points that are not well fit by the model; however, there are influential points that cannot be clearly identified. Thus, a cutoff point equal to 4/[n - p] was used to isolate these points resulting in 3578 observations. The initial inspection of the data did not reveal any inconsistency in the data with the exception of few data entry errors. The major problem encountered when applying these methodology was that all the default observations were included within these observations. Thus, it was not possible to run the model eliminating all these points as the methodology suggested. These results are not surprising since the data set is highly unbalanced (97.82% non-default and 2.18 % defaults). It was expected that given the nature of this data set, the few default observations (2.18%) would have a high leverage (h_i) and therefore a high dbeta value.

A second way to identify outliers was also employed to overcome the problem mentioned above. Outliers are cases with extreme values with respect to a single variable. Thus, values of included variables greater than the mean plus three standard deviations or lower than the mean minus three standard deviations were considered outliers, and consequently were adjusted back to three times the standard deviation above and below the mean. The observations adjusted using this method were compared to the observations identified using the dbeta measure. It turned out that the majority of the observations identify by the dbeta were outside the mean plus/minus three standard deviation range. As a result, less than 2% (3158 observations) of all loans were adjusted for any of the variables employed in the regressions.

Table 1 shows the regression results after adjusting for outliers in the independent variables. The likelihood ratios for all three definitions of default indicate that the corrected models are significantly better than the one with a constant only. When using the first definition of default, the coefficient for repayment capacity is still not significantly different from zero. The coefficient of the owner equity percentage is significantly different from zero and negative, and when compared to the case without the outliers' adjustments, the effect is slightly stronger. Furthermore, the probability of default for a one-unit increase in the owner equity percentage decreases from 2.18 % to 2.13%. The working capital as a percentage of gross income is also negative and significant; however, the effect is one point stronger than the case without outlier corrections. In other words, for one unit increase in working capital as a percentage of the gross income, the probability of default decreases by 0.01%. For the second and third definitions of default, the results are similar. The repayment capacity is not significantly different from zero. The other two variables are statistically significant at 99% percent level and negative as was expected.[8] Compared to the case without outlier adjustments, the effect of the owner equity percentage in both cases is stronger by 0.001 points; the effect of the working capital as a percentage of the gross income is also stronger by 0.001 points. There is no difference between independent variables coefficients when using the second and third definitions. In all cases, none of the coefficients is a large number indicating that it will take a major change to have a large impact on the probability of default. When comparing the McFadden's Pseudo-R-Square, it can be seen that the R-Square increased from 0.022 to 0.024 using the first definition and from 0.027 to 0.028 when using the second and third definition of default. These results indicate a little improvement in the strength of the association. Now, the R-Square measure can be used to compare the model with outlier corrections against the model without correction since both models have the same frequency distributions of dependent variable.

Square term inclusion

Loan defaults were further analyzed by including square terms in the model. A Box-Tidwell transformation (test) was conducted to check for the assumption of linearity relationship between the independent variables and the log odds of the dependent variable. The test results indicate that the coefficient for the interaction terms, in the case of the owner equity percentage was not significant, while for the repayment capacity and for the working capital were significant indicating the presence of nonlinearity in the relationship. Table 1 shows the logit regression results when square terms were included as new variables while attempting to improve the fitting of previous models. For the first definition of default, the coefficient of the repayment capacity is negative and significantly different from zero, an important improvement. The coefficient of

[8] The expected sign are shown in **Error! Reference source not found.**

the owner equity percentage and working capital are also negative and significantly different from zero. Translating these results in probability terms, it can be said that for a one-unit increase either in the CDRC as percentage of the loan or in the working capital ratio, the probability of default decreases from 2.18 % to 2.17%. Furthermore, for one unit increase in owner equity percentage, the probability of default decreases by 0.03%. For the second and third definitions of default, the results are similar. The repayment capacity is again negative and significantly different from zero at 99% percent level. The other two variables are also statistically significant at 99% level and negative as was expected. Compared to the case with outliers' adjustments, the effect is slightly stronger. In terms of probabilities, for a one-unit increase in either the CDRC as percentage of the loan or in the working capital ratio, the probability of default decreases from 2.18 % to 2.59%. For one unit increase in owner equity percentage, the probability of default decreases from 2.18 % to 2.55% and for the working capital ratio to 2.58%.

For all definitions of default, the square terms for the repayment capacity and for the working capital turned to be positive and significantly different from zero at a 99% significance level, while the square term coefficient for owner equity was not significant. These results corroborated the outcome from the Box-Tidwell test showing a strong non-linear relationship between the independent variables and the log odds of the dependent variable. The coefficients for the square term of debt repayment capacity and working capital indicate an increasing marginal effect of CDRC and working capital on the log odds of default. Even though, these coefficients are very small, they are highly significant indicating the presence of a non-linear relationship between dependent and independent variables.

Furthermore, a joint test for all square terms was conducted to confirm the non-linear relationship between the dependant and independent variables, to test this relationship, the following hypothesis was used, the Ho: $\beta_4= 0$, $\beta_5 =0$ and $\beta_6=0$. The hypothesis that the effects of the square terms are simultaneously equal to zero can be rejected at the 0.01 level ($X^2= 153.78$, $df =3$, $p= 0.000$). This result confirms the non-linear relationship between the dependant and independent variables. Additionally, to test the equality of coefficient for the square terms, the following hypothesis was tested, Ho: $\beta_4= \beta_5 =\beta_6$. The Chi-Square test conclude that the null hypothesis that the effect of the square terms are equal is significant at the 0.01 level ($X^2=10.49$, $df =2$, $p= 0.0053$). This result suggests that there is strong evidence that the effects are equal, however, the regression results indicate that the square term coefficient for owner equity was not significant. Thus, three additional hypotheses were tested: a) Ho: $\beta_4= \beta_5$. The Chi-Square test for this hypothesis indicates that the null hypothesis, that square term for the square repayment capacity is equal to the square term for the owner's equity percentage, can be rejected ($X^2=0.02$, $df =1$, $p= 089$); b) the second additional hypothesis is Ho: $\beta_5 =\beta_6$. The Chi-Square test for the hypothesis that the estimate for the square term of owner's equity as a percentage is equal to the estimate for the square term of working capital can also be rejected ($X^2=0.00$, $df =1$, $p= 0.9997$). c) The Chi-Square test for the hypothesis Ho: $\beta_4=\beta_6$ indicates that effect of square terms for the repayment capacity and the working capital are equally significant at 0.01 level ($X^2=10.40$, $df =1$, $p= 0.0013$).

Models comparison based on information criterions

To be able to compare competing models, the Akaike's Information Criterion (AIC*n) and Bayesian Information Criteria (BIC) were employed. These measures of fit provide an index to asses whether a model is adequate. The results of theses information criteria are shown in Table 1. The lowest AIC*n criterions for the first, second and third definition of default were found in the results from model that includes square terms. Comparing the AIC*n criterion within the square terms model, the lowest value belongs to the first definition of default. Thus, based on the AIC*n criterion the first definition of default in the square terms model is the most adequate model. Conversely, the lowest BIC criterions for all the definitions of default are found in the square terms model. The differences among BIC values from the three models are greater than 10 for all definitions of default, thus following the strength classification developed by Raftery (1996), there is strong evidence favoring the square terms model over the other two models. Within the square terms model results, the lowest value for the BIC criterion belongs to the third definition of default. The difference between the first and second definitions of default favors the second definition, while the difference between the second and third definition is 2.5 indicating that there is positive evidence favoring the third definition. Based on the lower values for the information criterion, the model that included the square terms was chosen. This model will be used as the basic model for the following analysis.

Marginal Change in the Predicted Probabilities

Marginal changes were calculated with all the variables held at their mean and using the *prchange* command in STATA as shown in Table 2. The first column show the change in the predicted probability as *x* changes from its minimum to its maximum. The strongest negative marginal change is for the working capital ratio; thus, moving from the lowest value of working capital to the highest, the probability of default would reduce approximately by 0.15 holding the other variables at the mean. Recall, that the marginal change is the instantaneous rate of change. Consequently, it does not equal to the actual change for a finite change in the working capital unless you are in a region of the probability curve that is approximately linear. The smallest marginal change belongs to the square term for the owner equity as a percentage of the loan. The second column in

Table 2 shows the change in the predicted probability as *x* change from 0 to 1. In this case, the strongest marginal change belongs to owner equity percentage. The marginal change for the other two independent variables is weaker and almost zero for the square terms. The third column shows the change in predicted probability as *x* change from ½ units below the base to ½ units above. For example for a loan that averaged in all categories, an additional one percent in owner equity as percentage of the loan decreases the probability of default by 0.0006. The fourth column indicates the change in predicted probability as *x* changes from ½ standard deviations below base to ½ standard deviations above. A standard deviation change in working capital ratio centered on the mean will decrease the probability of default by 0.01 holding other variables at their means. The fifth column shows the partial derivative of the predicted probability/rate with respect to the given set of independent variables, while the sixth and seventh columns show the 95% confidence interval for a discrete change. In this sense, it is possible to say with 95% confidence that the true change in the default probability associated with an increase on the owner equity percentage is between -0.0007 and -0.0004.

Using a cutoff of 3% for classifying default, the model correctly predicted 71.72% of the loans that would default, 50.02% of the loans that actually defaulted and 72.30% of the loans that did not default as indicated by the third and fourth columns in Table 3. The cutoff of 3 % was chosen because it is the closest value to the actual default rate of 2.6 %. Table 3 shows different cutoff points and the correctly percentage of prediction. It is clear that the cutoff point for loans that would default, defaulted, or did not default affects the percentage correctly predicted. As the cutoff percentage increases, the sensitivity decreases while the specificity increases.

Regression Result Analysis by Loan Type

There are five types of loans: operating loans, intermediate term loans, real state loans, rural residence loans and commercial loans. Operating loans are those to purchase inputs and to pay expenses. These are short-term loans used to finance daily farm operating expenses such as labor, input costs, equipment repairs, feeder livestock, etc. Intermediate loans are those to finance equipment, livestock, irrigation systems, vehicles or other capital items for up to 7 years. Intermediate term loans can also be used for forestry harvest products, providers of farm marketing and processing, service contract growers and farm related businesses. Real state loans are those for the acquisitions of any type of real state. They are long-term loans, and are used to finance or refinance farmland purchases, improvements to farmland, and agricultural facilities and buildings, and land contract. Loan terms range from 5-30 years. Rural residence loans are those to finance or refinance homes. They are designed for homes on acreage, homes in rural subdivisions and for homes in rural towns. Sometimes, they involved lot and construction financing. Moreover, commercial loans are usually a full range of loan products and services to agri-businesses, agricultural coops, implement dealers, and large farming operations. However, for this study, the commercial loan will refer to those loans with more than 5 million dollars in commitment.

The loans were analyzed by loan type to asses if there are any improvements in fitting the models and their predictions capabilities. Using the three different default definitions, separate models were run for each of the loan categories. Previous results as well as the present analysis by loan type show that some coefficient improves significantly when using second and third definition of default, however in general there are not many differences among the second and third definitions. Therefore, to simplify the presentation of results, only the results for the third definition, that follows closely the Basel II criteria, will be explained. The regression results when using the third definition of default by loan type are presented in Table 4. In general, these results are consistent with those in Table 1. Notice that only four out of five loan categories are presented in the results, commercial loans could not be analyzed because the maximum likelihood estimation is not possible when the dependant variable does not vary within one of the categories of the independent variables. In the case of commercial loans, the working capital as a percentage of gross income does not have a corresponding defaulted loan, thus the model cannot be fitted because the coefficient for the working capital variable is effectively negative infinity. Stata’s solution is to drop this variable along with all observations. These findings indicate that a much larger number of default cases within the commercial loans are needed to fit the model for commercial loans.

For the data set without corrections, the owner’s equity coefficient is always significant at 99% level except in the rural resident loan model. For the operating loan model, no other coefficients are significant. The likelihood ratio has a chi-square of 207 with a p-value of 0.000. This indicates that the model as a whole fits significantly better than the reduced model. Likewise, for the intermediate terms and rural resident loans, the likelihood ratio is also significantly better than the reduced one. The coefficient for working capital is also significant at 99% level, while for the rural residence loans none of the coefficients are significant. These results are confirmed with the non-significance of the likelihood ratio. Similarly, for the data set with outliers' corrections, the results are analogous with the only difference that the working capital coefficient is significant at 99% level for all loan types. The likelihood ratio test is significant for all type of loans except for the rural residence loans.

For the model with square term, all the origination ratios are statistically significant at 99% level, and the sign for all of the coefficients are negative for the operating and intermediate term loan models as shown in Table 4. Similarly, the sign for the all square term except for the owner's equity percentage are significant at 99% level and positive. For the real state loan model, only the coefficient for working capital is significant at 99% level. Furthermore, only the coefficient for repayment capacity square term is not significant while the other square terms are significant at 99% level. Real estate loans are typically long-term, thus lenders put more attention to borrower's long-term assets. This is why borrower's repayment capacity and owner's equity as percentage are not significant, suggesting that alternative variables should be incorporated to better fit a logit model for real state loans. In the rural residence model, none of the variables are significant. These results can be explained due to the lack of degrees of freedom when estimating the model since only four of the 883 loans defaulted. This is evident when observing the low log likelihood (Table 4) and noting that is the only ratio that is not significant among loan types. These results are expected, as rural residence loans are usually evaluated using a scorecard system (Featherstone et al, 2006).

The results for information criteria are shown at the bottom of Table 4. The lowest AIC*n criterions for the operating and real state loans belong to the model that includes square terms; for the intermediate loans the lowest criterion corresponds to the data set without corrections. Since the model for rural residence is not significant it does not matter which criterion is the smallest. Thus, based on the AIC*n criterion the square term model is a better fit to model loan defaults. Comparing among definitions, the first definition of default would be the most adequate model. Conversely, the lowest BIC criterions among definitions of default are found for the third definition. When comparing the BIC criterions among loan types, the lowest values are found in the square terms model for the operating and intermediate loans while for the real state model, the lowest BIC belongs to the model with outliers' corrections. The differences in BIC between the model with outliers' correction and without indicate strong/positive evidence that the model with correction is better since the BIC differences are greater than six for operating/intermediate loans. For real state loans, the evidence is weak according strength classification scale developed by Raftery (1996). Comparing the corrected model with the model that includes square term, the evidence is strong favoring the square terms model over the other two models for the first two types of loans, while for real state loan, the evidence suggests strongly that the model with outlier's correction is better. Based on the lower values for the information criterion, the model that included the square terms should be more appropriate to model operating and intermediate loans while to model real state loans, the model with outliers' correction should be employed.

Table 5 shows the marginal change in the predicted probabilities for operating loans. The strongest negative marginal change belongs to the working capital ratio; thus, moving from the lowest to the highest value of working capital, the probability of default would reduce approximately by 0.18 holding the other variables at their mean. The smallest marginal change belongs to the square term for the owner equity as a percentage of the loan. When changing positive values alone (from 0 to 1), the strongest marginal change belongs to owner equity percentage. The marginal change for the other two independent variables is weaker and almost zero for the square terms. As moving from ½ units below the base to ½ units above, again the owner's equity percentage has the greatest marginal change (0.0008). A bigger change such as ½ standard deviation below base to ½ standard deviation above in the owner's equity will decrease the probability of default by 0.016 holding other variables at their means while the same change in repayment capacity will result in a change of 0.0113 in the probability of default. The lowest marginal effect (0.0095) corresponds to a change in the working capital ratio. The fifth column shows the partial derivative of the predicted probability/rate with respect to the given set of independent variables. It is possible to say with 95% confidence that the true change in the default probability associated with an increase of the owner equity percentage is between -0.0011 and -0.0004 for operating loans.

Table 6 and Table 7 show the marginal effects on the probability of default for intermediate and real state loans. The biggest marginal change in default probabilities corresponds to the working capital ratio for both intermediate and real estate loans, while the lowest marginal change in default probabilities corresponds to repayment capacity for intermediate loans and for owner's equity for real state loans. The implications of these results are straightforward for lenders. For operating loans, lenders should put special attention on the working capital ratio. For intermediate term loans and real estate loans, lenders should put more attention to the owner's equity percentage and repayment capacity respectively. The magnitude of marginal change in predicted probabilities when moving up from the minimum to maximum values of independent variables is an indicator of the slope of the curves.

The closest value to the actual default rate for each loan type is approximately 3 %, thus using a cutoff of 3% for classifying default, the model correctly predicted 49.14 % of the operating loans that would default, 68.06 % of the operating loans that actually defaulted and 49.14% of the operating loans that did not default as shown in Table 8. In the case of intermediate term loans, the model correctly predicted 62.43% of the loans that would default, 60.73% of the ones that actually defaulted and 62.48% of the intermediate terms loans that did not default. Furthermore, the results for real state loans and rural residence loans are misleading since they indicate that the model predicted correctly 92.94% and 99.43% respectively. However, the regression results indicate that the model does not fit well the data.

Regression Result Analysis by Association

AgriBank is one of the districts within the Farm Credit System, and is composed by seventeen Agricultural Credit Associations (ACAs) that make short, intermediate, and long-term loans and one Agricultural Credit Bank that provide with funds and services to the local ACAs. The analysis of the portfolio credit risk could be further enhanced if segmented by type of association since each association has its own management style and its own method to evaluate the creditworthiness of borrowers. The analysis by association could help to asses if there is any difference within a group of lenders that provide roughly the same products, share a common source of funds, and have operated as a unified entity. With the purpose of attracting borrowers, the associations offer distinct type of credits relationship that are more appealing to the credit needs of specific types of farm borrowers. Thus, the associations could adopt different models to quantify credit risk and to calibrate their minimum capital requirements. The loan segmentation by association also helps to asses any improvements in fitting the models and their predictions capabilities.

The regression results by association are presented in Table 9. These results found that overall model consistency is similar among associations. The log-likelihood ratio test shows that the model a whole is significant at a 99% level for all associations. Only three out of the seventeen associations (Association 8, 11 and 16) have all coefficients of the origination ratios statistically significant. In the case of the Association 8, the repayment capacity percentage and the owner's equity ratio are negative and significant at 99% level, while the working capital ratio is negative and significant at only 90% level. In contrast, for the Association 11, the results show that both the repayment capacity percentage and the working capital ratio are negative and significant at 99% level, while the owner's equity ratio is negative and significant at the only 90% level. The regression coefficients for the repayment capacity percentage and the owner's equity ratio of the Association 16 are negative and significant at 95%, whereas the working capital coefficient is negative and significant at the 90% level. Alongside these results, six out of the seventeen associations have two out of the three coefficients negative and significant, while six out of the seventeen associations have only one coefficient negative and significant. The Association 6 is the only one that has no significant coefficients and the Association 9 was dropped from the analysis because none of its loans was defaulted. These findings are a clear indicator of the consistency of the model across associations. Working capital was the variable that better behaved (twelve associations has significant coefficients) follow by the repayment capacity variable (nine associations), and last, the owners' equity variable (seven associations). Overall results show that associations should put more emphasis on liquidity indicators, then on the repayments capacity of the borrowers, and finally on the borrowers' solvency.

Similarly, the sign of the square CDRC term is positive and significant for ten out of seventeen associations. Contrary, the square term of the owner's equity ratio is significant for only three of the seventeen associations. However, the coefficients for the Association 12 and 14 are negative indicating an increasing marginal effect of owner's equity on the log odds of default, while the coefficients of the Association 18 is positive indicating a decreasing marginal effect of the owners' equity on the log odds of default. This is sign of the variability among associations and an indicator of the non-linearity of relationship between dependent and independent variables.

The analysis by association was conducted comparing three alternative models, the same way as the analysis by loan type. Based on the AIC*n criterion the square term model combined with the third definition of default is the best mode to predict default. The results for the information criterions analysis are shown at the bottom of Table 9. The lowest AIC*n criteria for almost all association belong to the model that includes square terms. Within the square terms model, the lowest AIC*n criterion belongs to the Association 15. Similarly, the lowest BIC criterions among definitions of default are found for the third definition of default, and within the square terms model. These results provided strong evidence that the model using the third definition of default, correcting for outlier and including the square terms, is the best model to predict default at the associations' level. The lowest BIC criterion belongs to the Association 14.

Table 10 shows the marginal change in the predicted probabilities for Association 8. The strongest negative marginal change belongs to the repayment capacity percentage. Thus, when moving from the lowest to the highest value the repayment capacity percentage, the probability of default would reduce approximately by 0.73 holding the other variables at their mean, the smallest marginal change among origination ratio corresponds to the working capital percentage. It is interesting to see, that also the square terms displayed a considerably marginal changes in predicted probabilities. The smallest marginal change of all regressors belongs to the square term of the owner equity percentage. Contrasting, when changing positive values only (i.e. from 0 to 1), the strongest marginal change belongs to owner equity percentage. The marginal change for the other two independent variables is weaker and almost zero for the square terms. As moving from ½ units below the base to ½ units above, again the owner's equity percentage has the greatest marginal change (0.001). The marginal change for the other two independent variables is weaker and almost zero for the square terms as before. For bigger changes, such as ½ standard deviations below base to ½ standard deviations above the base, the biggest marginal change belongs now to the repayment capacity percentage; it decreased the probability of default by 0.02 holding other variables at their means. The same change in owners' equity resulted in a change of 0.01 in the probability of default. The lowest marginal effect (0.009) corresponds to a change in the working capital ratio. With 95% confidence, the true change in the default probability associated with an increase of the repayment capacity will be between -0.0004 and -0.002; the change in the default probability associated with a change in owners' equity is between -0.0016 and -0.0003, and for working capital is between -0.0003 and 0.000.

Table 11 and Table 12 show the marginal effects on the probability of default for Association 11 and Association 16 The biggest marginal change in default probabilities corresponds to the working capital ratio for the Association 11 and to the owners' equity for the Association 16 when moving from the lowest to the highest value of original ratios. On the contrary, for the same range, the lowest marginal corresponds to the to the owners' equity for the Association 11 and to the working capital ratio for the Association 16. Furthermore, when changing values (i.e. from 0 to 1), the strongest marginal change belongs to owner equity percentage for both the Association 11 and Association 16. The marginal change for the other two independent variables is weak, and almost zero for the square terms in both cases. When moving from ½ units below the base to ½ units above, again the owner's equity percentage has the greatest marginal change by -0.006 and by -0.0018 for Association 11 and Association 16, respectively. The marginal change for the other two independent variables is weaker and almost zero for the square terms as before. When changing ½ standard deviations below base to ½ standard deviations above the base, the biggest marginal change is for the repayment capacity percentage in the case of the Association 11, and for the owners' equity in the case of the Association 16. With 95% confidence level, it can be said that the true change in the default probability associated with an increase of the repayment capacity will be between -0.0003 and -0.0001 for the Association 11 and between -0.0005 and -0.0001 for the Association 16. The change in the default probability associated with a change in owners' equity is between -0.0011 and -0.000 for the Association 8 and between -0.0031 and -0.0005 for the Association 16; and for working capital is between -0.0003 and 0.0001, and between -0.0004 and -0.000 for the Association 11 and the Association 16, respectively.

Regression Result Analysis by Payment Frequency

There are six types of payment frequencies: monthly payment, quarterly payment, semi-annually payment, annually payment, variable payment and final maturity payment. Monthly payments are those occurring once a month. Quarterly payments are those occurring four times a year, usually March, June, September, and December. Semi-annually payments are those occurring twice a year. Annually payments are those occurring once a year. Variable payments are those whit specific installment options or special conditions regarding the frequency of payment such as deferments. Finally, final maturity payments are those that have a single payment at the end of the loan term. The loans were analyzed payment frequency to asses if there are any improvements in fitting the models and their predictions capabilities. Analyzing the frequencies of payment helps to understand which of the frequency options give the borrower the best chance of repayment. Consequently, it helps to assess potential repayment problems, and its influence on the default probabilities.

The same way as the analysis by loan type and associations, three different default definitions were employed. Consistent with previous results, the regression results improves significantly when using third definition of default. Therefore, to simplify the exposition of results, only the results for the third definition and with the outlier corrections are explained. The regression results by payment frequency are presented in Table 13. The likelihood ratio indicates that the model as a whole fits significantly better than the reduced model for all payment frequencies except for the quarterly payments (p-value of 0.223).All the coefficients of all the origination ratios are statistically significant at 99% level, and the sign for all of the coefficients are negative for the annually and final maturity payment type. For monthly payment type, the sign for the all origination ratio are negative and statistically significant at 99% level except for the repayment capacity percentage. In the case of the seminally payment type, the owner's equity percentage is not significant, while the other two origination ratios are negative and significant at 99% level. The quarterly and variable payments types do not show any significant coefficients. These results are in a certain way expected since the frequency of payments differs among loans; typically, intermediate- and long-term loans are structured with monthly, quarterly, semiannual or annual payments. The majority of the observations employed in this study correspond to intermediate and long-term loans thus are expected to have significant coefficients for monthly, semiannual or annual payment types. However, the results for the quarterly payment type are surprisingly. One possible explanation could be the fact that the majority of the quarterly payment loans are within the real state loans where only working capital is significant. In the case of the variable payment type, it is not estrange that none of the variables are significant, since it is difficult to fit a model for loans that have especial conditions such as deferments. . It is preferable to have payments, which correspond with high unmodified cash inflows.

Analyzing the McFadden's Pseudo-R-Square, it can be seen that the R-Square ranges from 0.0029 to 0.04. According to this statistic the model, the annually payment type is the model that shows the higher strength of association. Furthermore, comparing the McFadden's Pseudo-R-Square among default definitions, results indicate that there is an improvement in the strength of the association favoring again the third definition of default.

Table 14 shows the marginal change in the predicted probabilities for annually payments. The strongest negative marginal change, among origination ratios, belongs to the working capital percentage; thus, moving from the lowest to the highest value of working capital, the probability of default would reduce approximately by 0.17 holding the other variables at their mean. The smallest marginal change belongs to the owner equity as a percentage of the loan. Among square term, the negative marginal effect of the square term of repayment capacity displays a0.21 change in predicted probability. When changing positive values i.e. from 0 to 1, the strongest marginal change belongs to owners' equity percentage. The marginal change for the other two independent variables is weaker and almost zero for the square terms. As moving from ½ units below the base to ½ units above, again the owner's equity percentage has the greatest marginal change (0.0005). Changing from ½ standard deviations below base to ½ standard deviations above in the repayment capacity (the strongest marginal change) will decrease the probability of default by 0.0094 holding other variables at their means while the same change in working capital ratio will result in a change of 0.0086 in the probability of default. The lowest marginal effect (0.008) corresponds to a change in the owners' equity ratio. The fifth column shows the partial derivative of the predicted probability/rate with respect to the given set of independent variables. It is possible to say with 95% confidence that the true change in the default probability associated with an increase of the repayment capacity will be between -0.0002 and -0.0001; between -0.0007 and -0.0002 for an increase in the owners' equity percentage; between -0.0001 and -0.0000 for working capital ratio. All these marginal effect correspond to Annually Payments.

Table 15 shows the marginal change in predicted probability of default for final maturity payments. The biggest marginal change in default probabilities corresponds to the working capital ratio, while the lowest marginal change in default probabilities corresponds to the owner's equity when moving from the lowest to the highest values of each original ratio. Table 16 displays the marginal change in predicted probability of default for monthly payments and Table 17 displays the marginal change in predicted probability of default for semi-annually payments. The strongest marginal change in default probabilities corresponds to the owners' equity for monthly payments and to the repayment capacity percentage for the semi-annually payments when moving from the lowest to the highest value of these ratios. Furthermore, the lowest marginal change corresponds to the repayment capacity percentage for monthly payments and to the owners' equity percentage for the semi-annually payments. Additionally, when changing values i.e. from 0 to 1, the strongest marginal change belongs to the owners' equity for both monthly and semi-annually payments. The marginal change for the other two independent variables is weak, and almost zero for the square terms in both cases. When moving from ½ units below the base to ½ units above, again the owner's equity percentage has the strongest marginal change for both type of payment frequencies. The marginal change for the other two independent variables is weaker and almost zero for the square terms as before. When changing ½ standard deviations below base to ½ standard deviations above the base, the strongest marginal change is for the owners' equity for monthly payments and for the repayment capacity percentage in the case of semi-annually payments. The 95% confidence interval shows that the true change in the default probability associated with an increase of one unit in the repayment owner's equity will be between -0.0014 and -0.0007 for monthly payments, and between -0.0011 and -0.0002 for semi-annually payments. The lenders can interpret these results straightforward; for annually and final maturity payments, lenders should put special attention on the working capital ratio; for monthly payments, lenders should put more attention to the owner's equity percentage and for the semi-annually payments, lenders should put more attention to the repayment capacity and working capital ratio.

Summary and Conclusions

Using AgriBank's loan level data, a binomial logit model was estimated to determine the probability of default for any loan within the loan portfolio of AgriBank conditional on the realization of financial ratios. The New Basel Accord suggested eight criteria for implementing risk rating system; however, due to lack of data, only three criteria were employed. From the empirical analysis using the logit model, financial measures of liquidity, solvency and repayment capacity were found to be important determinants of the probability of default. Liquidity was approximate by the working capital to average gross income ratio; solvency was measured as owner equity as a percentage of the loan, and repayment capacity was measured as the CDRC percentage at the loan level.

Three alternative definition of default were employed to incorporate the four conditions that need to be met for a loan to be considered in default. None mayor differences were observed among the three default definitions. However, the regression coefficients obtained out of the econometric model favor the third definition of default. The likelihood ratios indicated that the model as a whole fits significantly better than the reduced model. The McFadden's Pseudo-R-Square improved when using the second and third definition of default as dependent variables, confirming that the third definition of default is the most appropriate one.

Loans were analyzed to look for outliers and influential point using standard Pearson residual and Pregibon's measure (dbeta), respectively. The results with the outlier's corrections are better when compared against the model without corrections. The McFadden's Pseudo-R-Square increased when using the second and third definition of default. These results indicate a little improvement in the strength of the association. The Box-Tidwell test showed a strong evidence for a nonlinear relationship between the log odds of the dependent variable and the repayment capacity, and the working capital variables. The coefficients for the square term of debt repayment capacity and working capital indicate an increasing marginal effect of CDRC and working capital on the log odds of default. Even though, these coefficients are very small, they are highly significant indicating the presence of a non-linear relationship between dependent and independent variables. The overall results show that in all cases, none of the coefficients is a large number; thus, it will take a major change to have a large impact on the probability of default. The result of the AIC*n and BIC information criterions provided positive evidence favoring the third definition of default and the inclusion of square terms in the econometric model.

The marginal changes for each variable were calculated holding the other variables at their mean. The strongest negative marginal change is for the working capital ratio; thus, moving from the lowest value of working capital to the highest, the probability of default would reduce approximately by 0.15 holding the other variables at the mean. The smallest marginal change belongs to the square term for the owner equity as a percentage of the loan. Using a cutoff of 3%

for classifying default, the model correctly predicted 71.72% of the loans that would default, 50.02% of the loans that actually defaulted and 72.30% of the loans that did not default.

Only four out of five loan types were analyzed; commercial loans could not be analyzed because the maximum likelihood estimation is not possible when the dependant variable does not vary within one of the categories of the independent variable. This indicates that a larger number of observations are needed to fit the model adequately for the commercial loans. The results for the other loan types shows that all the origination ratios are statistically significant negative for the operating and intermediate term loans. For the real state loan model, only the coefficient for working capital is significant. Real estate loans are typically long-term, thus lenders put more attention to borrower's long-term assets. This is why borrower's repayment capacity and owner's equity percentage are not significant, suggesting that alternative variables should be incorporated to better fit the logit model in the case of real state loans. Rural residence loans did no have enough default loans to estimate the model. For operating loans, the strongest marginal change belongs to the working capital ratio. In the case of intermediate loans, the strongest marginal change corresponds to the owner's equity percentage, while for real state loan corresponds to the repayment capacity. The implications of these results are straightforward for lenders; for operating loans, lenders should put special attention on the working capital ratio; for intermediate term loans and real estate loans, lenders should put more attention to the owner's equity percentage and repayment capacity respectively. In general, i.e. using a cutoff of 3% for classifying default, the model correctly predicted 49.14 % of the operating loans that would default, 68.06 % of the operating loans that actually defaulted and 49.14% of the operating loans that did not default.

The analysis by association helped to assess difference among groups of lenders that provide roughly the same products. The regression results by association displays an overall consistency among associations. Only three out of the seventeen associations (Association 8, 11 and 16) have all coefficients of the origination ratios statistically significant. Six out of the seventeen associations have two out of the three coefficients negative and significant, while six out of the seventeen associations have only one coefficient negative and significant. These findings are a clear indicator of the consistency of the model across associations. Overall results show that associations should put more emphasis on liquidity indicators, then on the repayments capacity of the borrowers, and finally on the borrowers' solvency. In the case of Association 8, the strongest negative marginal change belongs to the repayment capacity percentage while the strongest marginal change in default probabilities for the Association 11 corresponds to the working capital ratio and to the owners' equity percentage for the Association 16.

The regression results by payment frequency show that the model as a whole fits significantly better than the reduced model for all payment frequencies except for the quarterly payments. The quarterly and variable payments types do not show any significant coefficients. These results are in a certain way expected since the frequency of payments differs among loans; typically, intermediate- and long-term loans are structured with monthly, quarterly, semiannual or annual payments. The majority of the observations employed in this study correspond to intermediate and long-term loans thus are expected to have significant coefficients for monthly, semiannual or annual payment types. In the case of the variable payment type, it is not estrange that none of the variables are significant, since it is difficult to fit a model for loans that have especial conditions such as deferments. Based on these results, lenders should put special attention on the working capital ratio for annually and final maturity payments, and to the owner's equity percentage, for monthly payments and to the repayment capacity and working capital ratio for the semi-annually payments.

Table 1: Logit Regression Result of the Probability of Default for All Loans in the Portfolio

	No correction			Outlier correction			Square terms or model modification		
Variables	**1st Definition**	**2nd Definition**	**3rd Definition**	**1st Definition**	**2nd Definition**	**3rd Definition**	**1st Definition**	**2nd Definition**	**3rd Definition**
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Repayment Capacity as % of the loan	0.000	-0.000	0.000	0.000	0.000	0.000	-0.005***	-0.005***	-0.005***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	-(0.001)	-(0.001)	-(0.001)
Owner equity % of the loan	-0.022***	-0.024***	-0.024***	-0.023***	-0.025***	-0.025***	-0.023***	-0.026***	-0.026***
	(0.001)	(0.001)	(0.001)	-(0.001)	-(0.001)	-(0.001)	-(0.004)	-(0.004)	-(0.004)
Working/Av. Income capital	-0.002***	-0.001***	-0.001***	-0.003***	-0.002***	-0.002***	-0.005***	-0.005***	-0.005***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Square CDRC as % of the loan	-	-	-	-	-	-	0.000***	0.000***	0.000***
	-	-	-	-	-	-	(0.000)	(0.000)	(0.000)
Square Owner equity % of the loan	-	-	-	-	-	-	0.000	0.000	0.000
	-	-	-	-	-	-	(0.000)	(0.000)	(0.000)
Square Working/Av. Income capital	-	-	-	-	-	-	0.000***	0.000***	0.000***
	-	-	-	-	-	-	(0.000)	(0.000)	(0.000)
Constant	-2.417***	-2.084***	-2.082***	-2.364***	-2.024***	-2.021***	-1.923***	-1.581***	-1.577***
	(0.057)	(0.052)	(0.052)	-(0.061)	-(0.056)	-(0.056)	-(0.123)	-(0.111)	-(0.111)
Num. of observations	158426	158426	158426	158426	158426	158426	158426	158426	158426
Log likelihood	-16255	-18594	-18604	-16234	-18578	-18588	-16169	-18503	-18512
LR chi2(3)	743.46	1021.52	1023.81	784	1054	1056	915	1205	1207
Prob >chi2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
McFadden's R- Square	0.022	0.027	0.027	0.024	0.028	0.028	0.028	0.032	0.032
Cragg and Uhler's R-Square	0.025	0.030	0.030	0.026	0.031	0.031	0.030	0.035	0.035
AIC*n:	32517.385	37196.976	37216.43	32476.636	37164.917	37148.271	32351.387	37019.664	37038.795
BIC':	-707.544	-985.606	-987.886	-748.293	-1017.664	-1020.045	-884.623	-1132.998	-1135.603

Standard errors in parentheses
* Significant at 10%; ** significant at 5%; *** significant at 1%

Table 2: Marginal Change of Predicted Probabilities

	min>max	0>1	-+1/2	-+sd/2	Mg. Effect	[95%	C.I.]	Mean(x)	Sd(x)
Rep. Capacity as % of the loan	-0.1040	-0.0003	-0.0001	-0.0095	-0.0001	-0.0001	-0.0001	153.383	77.279
Owner equity % of the loan	-0.0812	-0.0024	-0.0006	-0.0100	-0.0006	-0.0007	-0.0004	63.232	17.317
Working/Av. Income capital	-0.1577	-0.0001	-0.0001	-0.0102	-0.0001	-0.0001	-0.0001	43.754	97.163
Sq. CDRC as % of the loan	0.1469	0.0000	0.0000	0.0097	0.0000	0.0000	0.0000	29498.400	37757.600
Sq. Owner Equity % of the loan	0.0057	0.0000	0.0000	0.0008	0.0000	0.0000	0.0000	4298.190	2135.980
Sq. Working/Av. Income Cap.	0.0930	0.0000	0.0000	0.0069	0.0000	0.0000	0.0000	11355.100	45881.700

Table 3: Percentage of Correct Predictions for Different Cutoff Points

Cutoff Percentage	Correct (%)	Sensitivity (%)	Specificity (%)
1	7.37	98.33	4.94
2	41.22	78.23	40.23
3	71.72	50.02	72.30
4	86.31	26.67	87.90
5	92.18	15.23	94.23
6	94.70	9.58	96.98
7	95.87	5.67	98.87
8	96.53	4.00	99.00
9	96.93	2.55	99.46
10	97.18	1.45	99.74

Table 4: Logit Regression Result of the Probability of Default by Loan Type

Variables/Loan Type	No correction				Outlier correction				Square terms or model modification			
	Operating	Int. Term	Real State	Rural Res.	Operating	Int. Term	Real State	Rural Res.	Operating	Int. Term	Real State	Rural Res.
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Repayment Capacity as % of the loan	-0.000	-0.000	-0.001	-0.001	-0.000	-0.001	-0.001	-0.001	-0.004***	-0.005***	-0.002	0.136
	(0.000)	(0.000)	(0.001)	(0.007)	(0.000)	(0.000)	(0.001)	(0.008)	(0.001)	(0.001)	(0.002)	(0.173)
Owner equity % of the loan	-0.019***	-0.028***	-0.029***	-0.011	-0.019***	-0.028***	-0.029***	-0.011	-0.025***	-0.030***	0.003	0.232
	(0.001)	(0.001)	(0.002)	(0.028)	(0.002)	(0.001)	(0.002)	(0.030)	(0.006)	(0.005)	(0.011)	(0.326)
Working/Av. Income capital	-0.001	-0.001***	-0.001**	0.000	-0.002**	-0.002***	-0.001**	0.000	-0.005***	-0.004***	-0.003***	-0.002
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.003)	(0.001)	(0.001)	(0.001)	(0.006)
Square CDRC as % of the loan	-	-	-	-	-	-	-	-	0.000***	0.000***	0.000	-0.000
	-	-	-	-	-	-	-	-	(0.000)	(0.000)	(0.000)	(0.001)
Square Owner equity % of the loan	-	-	-	-	-	-	-	-	0.000	0.000	-0.000**	-0.002
	-	-	-	-	-	-	-	-	(0.000)	(0.000)	(0.000)	(0.003)
Square Working/Av. Income capital	-	-	-	-	-	-	-	-	0.000***	0.000***	0.000**	0.000
	-	-	-	-	-	-	-	-	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-2.185***	-1.742***	-2.325***	-4.676*	-2.123***	-1.690***	-2.265***	-4.623*	-1.639***	-1.275***	-2.895***	-21.926
	(0.084)	(0.077)	(0.137)	(1.962)	(0.092)	(0.082)	(0.146)	(2.036)	(0.188)	(0.157)	(0.321)	(16.757)
Num. of observations	44258	64246	48994	883	44258	64246	48994	883	44258	64246	48994	883
Log likelihood	-6394	-8414	-3540	-25.5	-6390	-8411	-3539	-25.5	-6360	-8375	-3531	-24.1
LR chi2(3)	207	559	233	.174	215.003	565.935	234.764	0.162	275	637	251	2.87
Prob >chi2	0.000	0.000	0.000	0.982	0.000	0.000	0.000	0.983	0.000	0.000	0.000	0.825
Pseudo R-squared	0.016	0.032	0.032	0.003	0.017	0.033	0.032	0.003	0.021	0.037	0.034	0.056
AIC*n:	12795.712	16386.106	7087.477	58.984	12787.908	16829.128	7085.492	58.996	12733.847	16764.03	7075.421	62.287
BIC':	175.105	-525.745	-200.381	20.176	-182.909	-532.723	-202.365	58.996	-210.877	-570.61	-186.038	37.829

Standard errors in parentheses

* Significant at 10%; ** significant at 5%; *** significant at 1%

Table 5: Change in Predicting Probabilities for Operating Loans

	min->max	0->1	-+1/2	-+sd/2	Mg.Effect	[95%	C.I.]	Mean(x)	Sd(x)
Repayment Capacity as % of the loan	-0.0957	-0.0002	-0.0001	-0.0113	-0.0001	-0.0002	-0.0001	156.0420	87.7879
Owner equity % of the loan	-0.1036	-0.0029	-0.0008	-0.0136	-0.0008	-0.0011	-0.0004	63.2202	17.9359
Working/Av. Income capital	-0.1877	-0.0002	-0.0001	-0.0095	-0.0001	-0.0002	-0.0001	30.1663	69.2144
Square CDRC as % of the loan	0.1133	0.0000	0.0000	0.0116	0.0000	0.0000	0.0000	32055.8000	43692.2000
Square Owner equity % of the loan	0.0282	0.0000	0.0000	0.0039	0.0000	0.0000	0.0000	4318.4800	2256.6700
Square Working/Av. Income capital	0.2468	0.0000	0.0000	0.0084	0.0000	0.0000	0.0000	5700.5200	29295.3000

	Non-Default	Default
Pr(y|x)	0.9690	0.0310

Table 6: Change in Predicting Probabilities for Intermediate Term Loans

	min->max	0->1	-+1/2	-+sd/2	Mg.Effect	[95%	C.I.]	Mean(x)	Sd(x)
Repayment Capacity as % of the loan	-0.1062	-0.0003	-0.0001	-0.0102	-0.0001	-0.0002	-0.0001	152.9260	77.6074
Owner equity % of the loan	-0.1157	-0.0038	-0.0008	-0.0134	-0.0008	-0.0010	-0.0005	63.3411	17.4457
Working/Av. Income capital	-0.1505	-0.0001	-0.0001	-0.0084	-0.0001	-0.0001	-0.0001	30.8164	75.6337
Square CDRC as % of the loan	0.1455	0.0000	0.0000	0.0103	0.0000	0.0000	0.0000	29409.2000	37313.7000
Square Owner equity % of the loan	0.0087	0.0000	0.0000	0.0013	0.0000	0.0000	0.0000	4316.4400	2153.6700
Square Working/Av. Income capital	0.1410	0.0000	0.0000	0.0067	0.0000	0.0000	0.0000	6670.0300	34058.6000

	Non-Default	Default
Pr(y|x)	0.9738	0.0262

Table 7: Change in Predicting Probabilities for Real State Loans

	min->max	0->1	-+1/2	-+sd/2	Mg.Effect	[95%	C.I.]	Mean(x)	Sd(x)
Repayment Capacity as % of the loan	-0.0186	0.0000	0.0000	-0.0019	0.0000	-0.0001	0.0000	151.6690	65.9364
Owner equity % of the loan	0.0033	0.0000	0.0000	0.0005	0.0000	-0.0002	0.0003	63.1264	16.5753
Working/Av. Income capital	-0.0477	0.0000	0.0000	-0.0047	0.0000	-0.0001	0.0000	71.0270	129.0890
Square CDRC as % of the loan	0.0106	0.0000	0.0000	0.0012	0.0000	0.0000	0.0000	27351.1000	32061.9000
Square Owner equity % of the loan	-0.0399	0.0000	0.0000	-0.0071	0.0000	0.0000	0.0000	4259.6700	1998.3900
Square Working/Av. Income capital	0.0231	0.0000	0.0000	0.0033	0.0000	0.0000	0.0000	21708.4000	64115.5000

	Non-Default	Default
Pr(y|x)	0.9882	0.0118

Table 8: Percentage of Correct Predictions by Loan Type

	Loan type			
	Operating	Int. Term	Real State	Rural Res.
Cutoff (%)	3.0%	3.0%	3.0%	3.0%
Correctly (%)	49.78%	62.43%	92.94%	99.43%
Sensitivity (%)	68.06%	60.73%	13.34%	0.00%
Specificity (%)	49.14%	62.48%	94.09%	99.89%

Table 9: Logit Regression Result of the Probability of Default by Association

Variables/Associations	Association Number							
	2	3	4	5	6	7	8	10
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Repayment Capacity as % of the loan	-0.009**	-0.010**	-0.006***	-0.009	-0.001	-0.011	-0.016***	-0.001
	(0.003)	(0.003)	(0.002)	(0.005)	(0.003)	(0.006)	(0.003)	(0.001)
Owner equity % of the loan	0.022	-0.044	-0.009	-0.055	-0.020	-0.052	-0.054***	-0.027**
	(0.034)	(0.023)	(0.012)	(0.045)	(0.016)	(0.032)	(0.016)	(0.009)
Working/Av. Income capital	-0.003	-0.016***	-0.004**	-0.013**	-0.003	-0.010*	-0.009*	-0.007***
	(0.003)	(0.004)	(0.001)	(0.004)	(0.002)	(0.004)	(0.004)	(0.001)
Square CDRC as % of the loan	0.000***	0.000**	0.000***	0.000	0.000	0.000	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Square Owner equity % of the loan	-0.001	0.000	-0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Square Working/Av. Income capital	0.000	0.000***	0.000	0.000**	0.000	-0.000	0.000**	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-2.693**	-0.417	-2.296***	-1.714	-2.709***	-1.177	1.019*	-1.792***
	(1.035)	(0.663)	(0.323)	(1.440)	(0.461)	(0.962)	(0.475)	(0.263)
Num. of observations	8577	7750	16349	5704	9626	2880	5899	21755
Log likelihood	-605	-702	-1663	-322	-1110	-259	-631	-3215
LR chi2(3)	52.9	108	103	21	19.5	18.4	151	188
Prob >chi2	0.0000	0.0000	0.0000	.00181	.00342	.00536	0.0000	0.0000
Pseudo R-squared	.0419	.0713	.03	.0316	.0087	.0343	.107	.0284
AIC*n:	1224.0659	1417.9517	3340.1241	658.88371	2233.1071	532.07517	1275.1097	6444.3504
BIC':	1.4502886	-54.131616	-44.673395	30.861201	35.551139	29.41869	-99.076344	-127.89342

Standard errors in parentheses
* Significant at 10%; ** significant at 5%; *** significant at 1%

Table 9: Continued

Variables/Associations	Association Number							
	11	12	13	14	15	16	17	18
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Repayment Capacity as % of the loan	-0.008***	-0.005	-0.011***	-0.005***	0.012	-0.008**	-0.006	-0.009**
	(0.002)	(0.003)	(0.003)	(0.002)	(0.011)	(0.003)	(0.003)	(0.003)
Owner equity % of the loan	-0.026*	0.005	0.032	0.002	-0.061	-0.050**	-0.067**	-0.097***
	(0.012)	(0.017)	(0.036)	(0.011)	(0.043)	(0.017)	(0.023)	(0.016)
Working/Av. Income capital	-0.010***	-0.006**	-0.008*	-0.003***	-0.043***	-0.005*	-0.002	-0.004
	(0.002)	(0.002)	(0.004)	(0.001)	(0.012)	(0.003)	(0.002)	(0.002)
Square CDRC as % of the loan	0.000*	0.000	0.000*	0.000***	-0.000	0.000**	0.000*	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Square Owner equity % of the loan	-0.000	-0.001**	-0.001	-0.000*	0.000	0.000	0.000	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Square Working/Av. Income capital	0.000***	0.000	0.000	0.000	-0.000	0.000*	0.000	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-1.091***	-2.043***	-2.003*	-1.803***	-3.285*	0.024	-0.531	0.079
	(0.305)	(0.433)	(1.003)	(0.370)	(1.508)	(0.459)	(0.662)	(0.463)
Num. of observations	13336	9564	3957	33124	2115	3879	3510	10393
Log likelihood	-1653	-855	-528	-4461	-130	-685	-390	-882
LR chi2(3)	210	156	67.4	309	29	99.1	42	86.1
Prob >chi2	0.0000	0.0000	0.0000	0.0000	.000061	0.0000	0.0000	0.0000
Pseudo R-squared	.0598	.0835	.06	.0335	.1	.0675	.0511	.0465
AIC*n:	3320.0938	1723.6308	1069.958	8936.7835	274.14781	1383.7686	794.47151	1777.1402
BIC':	-153.41043	-100.69494	-17.741997	-247.00538	16.942271	-49.56999	6.9346053	-30.583881

Standard errors in parentheses
* Significant at 10%; ** significant at 5%; *** significant at 1%

Table 10: Change in Predicting Probabilities for Association 8

	min->max	0->1	-+1/2	-+sd/2	Mg.Effect	[95%	C.I.]	Mean(x)	Sd(x)
Repayment Capacity as % of the loan	-0.7325	-0.0026	-0.0003	-0.0229	-0.0003	-0.0004	-0.0002	172.6670	78.2474
Owner equity % of the loan	-0.3424	-0.0135	-0.0010	-0.0145	-0.0010	-0.0016	-0.0003	70.3388	14.7038
Working/Av. Income capital	-0.0561	-0.0002	-0.0002	-0.0092	-0.0002	-0.0003	0.0000	36.4917	55.5224
Square CDRC as % of the loan	0.7372	0.0000	0.0000	0.0210	0.0000	0.0000	0.0000	35935.6000	39947.5000
Square Owner equity % of the loan	0.0262	0.0000	0.0000	0.0049	0.0000	0.0000	0.0000	5163.7100	1940.0200
Square Working/Av. Income capital	0.8885	0.0000	0.0000	0.0085	0.0000	0.0000	0.0000	4413.8500	19933.4000

	Non-Default	Default
Pr(y|x)	0.9819	0.0181

Table 11: Change in Predicting Probabilities for Association 11

	min->max	0->1	-+1/2	-+sd/2	Mg.Effect	[95%	C.I.]	Mean(x)	Sd(x)
Repayment Capacity as % of the loan	-0.1929	-0.0005	-0.0002	-0.0121	-0.0002	-0.0003	-0.0001	141.0130	66.2899
Owner equity % of the loan	-0.0730	-0.0023	-0.0006	-0.0102	-0.0006	-0.0011	0.0000	57.3770	17.2016
Working/Av. Income capital	-0.6176	-0.0003	-0.0002	-0.0107	-0.0002	-0.0003	-0.0001	24.7745	48.3100
Square CDRC as % of the loan	0.1548	0.0000	0.0000	0.0081	0.0000	0.0000	0.0000	24278.6000	31298.0000
Square Owner equity % of the loan	-0.0100	0.0000	0.0000	-0.0016	0.0000	0.0000	0.0000	3588.0000	1976.5400
Square Working/Av. Income capital	0.9471	0.0000	0.0000	0.0107	0.0000	0.0000	0.0000	2947.4600	17382.3000

	Non-Default	Default
Pr(y|x)	0.9771	0.0229

Table 12: Change in Predicting Probabilities for Association 16

	min->max	0->1	-+1/2	-+sd/2	Mg.Effect	[95%	C.I.]	Mean(x)	Sd(x)
Repayment Capacity as % of the loan	-0.1987	-0.0008	-0.0003	-0.0226	-0.0003	-0.0005	-0.0001	141.3200	77.4558
Owner equity % of the loan	-0.3497	-0.0124	-0.0018	-0.0316	-0.0018	-0.0031	-0.0005	61.0753	16.9142
Working/Av. Income capital	-0.0902	-0.0002	-0.0002	-0.0107	-0.0002	-0.0004	0.0000	25.4953	53.5646
Square CDRC as % of the loan	0.4611	0.0000	0.0000	0.0226	0.0000	0.0000	0.0000	25969.1000	36369.3000
Square Owner equity % of the loan	0.1132	0.0000	0.0000	0.0104	0.0000	0.0000	0.0000	4016.2100	2035.3000
Square Working/Av. Income capital	0.6898	0.0000	0.0000	0.0104	0.0000	0.0000	0.0000	3518.4400	17723.4000

	Non-Default	Default
Pr(y|x)	0.9622	0.0378

Table 13: Logit Regression Result of the Probability of Default by Frequency Type

Variables/Payment Frequency	Annually	Final Maturity	Monthly	Quarterly	Semi-Annually	Variable
	b/se	b/se	b/se	b/se	b/se	b/se
Repayment Capacity as % of the loan	-0.007***	-0.006***	-0.000	-0.001	-0.012***	-0.005
	(0.001)	(0.001)	(0.001)	(0.005)	(0.002)	(0.003)
Owner equity % of the loan	-0.026***	-0.023**	-0.046***	0.036	-0.020	-0.027
	(0.006)	(0.008)	(0.007)	(0.025)	(0.014)	(0.030)
Working/Av. Income capital	-0.005***	-0.005***	-0.004***	-0.002	-0.006***	-0.007
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.004)
Square CDRC as % of the loan	0.000***	0.000***	0.000	0.000	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Square Owner equity % of the loan	-0.000	0.000	0.000**	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Square Working/Av. Income capital	0.000***	0.000***	0.000**	0.000	0.000**	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-1.312***	-1.712***	-1.645***	-4.154***	-1.205**	-0.686
	(0.186)	(0.230)	(0.209)	(0.754)	(0.430)	(0.843)
Num. of observations	74154	28744	40114	2940	11265	1209
Log likelihood	-7661	-3922	-4723	-439	-1357	-296
LR chi2(3)	680	169	320	8.21	103	18.3
Prob >chi2	0.000	0.000	0.000	.223	0.000	.00548
Pseudo R-squared	.0425	.021	.0327	.00927	.0366	.03
AIC*n:	15335.615	7858.6622	9460.0549	891.89208	2728.5011	605.71645

Standard errors in parentheses

* Significant at 10%; ** significant at 5%; *** significant at 1%

Table 14: Change in Predicting Probabilities for Annually Payments

	min->max	0->1	-+1/2	-+sd/2	Mg.Effect	[95%	C.I.]	Mean(x)	Sd(x)
Repayment Capacity as % of the loan	-0.1244	-0.0003	-0.0001	-0.0094	-0.0001	-0.0002	-0.0001	154.4	75.0
Owner equity % of the loan	-0.0736	-0.0022	-0.0005	-0.0080	-0.0005	-0.0007	-0.0002	66.1	16.6
Working/Av. Income capital	-0.1726	-0.0001	-0.0001	-0.0086	-0.0001	-0.0001	-0.0001	41.5	86.7
Square CDRC as % of the loan	0.2168	0.0000	0.0000	0.0102	0.0000	0.0000	0.0000	29467.3	36755.5
Square Owner equity % of the loan	-0.0099	0.0000	0.0000	-0.0016	0.0000	0.0000	0.0000	4645.6	2117.9
Square Working/Av. Income capital	0.1683	0.0000	0.0000	0.0071	0.0000	0.0000	0.0000	9233.4	39656.5

	Non-Default	Default
Pr(y|x)	0.9812	0.0188

Table 15: Change in Predicting Probabilities for Final Maturity Payments

	min->max	0->1	-+1/2	-+sd/2	Mg.Effect	[95%	C.I.]	Mean(x)	Sd(x)
Repayment Capacity as % of the loan	-0.1393	-0.0004	-0.0002	-0.0137	-0.0002	-0.0002	-0.0001	149.9	83.1
Owner equity % of the loan	-0.0823	-0.0022	-0.0006	-0.0112	-0.0006	-0.0011	-0.0002	61.3	17.2
Working/Av. Income capital	-0.1994	-0.0002	-0.0001	-0.0094	-0.0001	-0.0002	-0.0001	28.6	67.6
Square CDRC as % of the loan	0.1534	0.0000	0.0000	0.0121	0.0000	0.0000	0.0000	29384.9	40510.7
Square Owner equity % of the loan	0.0283	0.0000	0.0000	0.0036	0.0000	0.0000	0.0000	4053.2	2093.0
Square Working/Av. Income capital	0.2924	0.0000	0.0000	0.0085	0.0000	0.0000	0.0000	5391.6	28315.0

	Non-Default	Default
Pr(y|x)	0.9710	0.0290

Table 16: Change in Predicting Probabilities for Monthly Payments

	min->max	0->1	-+1/2	-+sd/2	Mg.Effect	[95%	C.I.]	Mean(x)	Sd(x)
Repayment Capacity as % of the loan	-0.0029	0.0000	0.0000	-0.0004	0.0000	-0.0001	0.0001	154.9	79.4
Owner equity % of the loan	-0.1921	-0.0088	-0.0010	-0.0184	-0.0010	-0.0014	-0.0007	59.5	17.5
Working/Av. Income capital	-0.1437	-0.0001	-0.0001	-0.0121	-0.0001	-0.0001	-0.0001	54.1	123.6
Square CDRC as % of the loan	0.0052	0.0000	0.0000	0.0009	0.0000	0.0000	0.0000	30290.4	38612.0
Square Owner equity % of the loan	0.1392	0.0000	0.0000	0.0095	0.0000	0.0000	0.0000	3852.2	2081.7
Square Working/Av. Income capital	0.0395	0.0000	0.0000	0.0056	0.0000	0.0000	0.0000	18192.5	61151.3

	Non-Default	Default
Pr(y|x)	0.9770	0.0230

Table 17: Change in Predicting Probabilities for Semi-Annually Payments

	min->max	0->1	-+1/2	-+sd/2	Mg.Effect	[95%	C.I.]	Mean(x)	Sd(x)
Repayment Capacity as % of the loan	-0.4362	-0.0012	-0.0003	-0.0195	-0.0003	-0.0004	-0.0002	153.1	71.4
Owner equity % of the loan	-0.0617	-0.0015	-0.0005	-0.0078	-0.0005	-0.0011	0.0002	65.0	16.8
Working/Av. Income capital	-0.2465	-0.0002	-0.0001	-0.0143	-0.0001	-0.0002	-0.0001	59.4	108.8
Square CDRC as % of the loan	0.5990	0.0000	0.0000	0.0189	0.0000	0.0000	0.0000	28552.1	35656.5
Square Owner equity % of the loan	0.0106	0.0000	0.0000	0.0015	0.0000	0.0000	0.0000	4501.6	2080.1
Square Working/Av. Income capital	0.1036	0.0000	0.0000	0.0085	0.0000	0.0000	0.0000	15376.3	51609.8

	Non-Default	Default
Pr(y|x)	0.9763	0.0237

References

Allcott, K.W “An Agricultural Loan Rating System” *The Journal of Commercial Bank Lending* 65(1985):29-38.

Barry, J. Peter “Modern Capital Management by Financial Institutions: Implications for Agricultural Lenders”, *Agricultural Finance Review,* 61 (Fall 2001): 103-122

Barry, J. Peter and P. N. Ellinger. "Credit Scoring, Loan Pricing, and Farm Business Performance." *Western Journal of Agricultural Economics.* 14(1989):15-55

Cramer, J. S. “Econometric Applications of Maximum Likelihood Methods” , Cambridge, Cambridge University Press 1986

Chhikara, R. “The State of The Art in Credit Evaluation” *American Journal of Agricultural Economics.* 71,5(1989):1138-44.

Eliason, S. “Maximum Likelihood Estimation” Newbury Park CA: Sage. 1993

Ellinger, P. N., N. S. Splett, and P. J. Barry. "Consistency of Credit Evaluations at Agricultural Banks." *Agribusiness: An International Journal.* 52(1992a):43-56.

Ellinger, P.N. et al, “Credit Evaluation Procedures at Agricultural Banks” *Financing Agriculture in a Changing Environment: Macro, Market, Policy and Management Issues – Proceedings of Regional Research Committee NC-161,* Dept. of Ag. Econ, Kansas State Univ., Manhattan, Feb. 1992, p.61-82.

Ellinger, P. N., N. S. Splett and P. J. Barry. "Consistency of Credit Evaluation Procedures at Agricultural Banks." *Agribusiness: An International Journal.* 8, (1992), 517-536.

Featherstone, A. M., L. M. Roessler, and P. J. Barry “Determining the Probability of Default and Risk-Rating Class for Loans in the Seventh Farm Credit District Portfolio”, *Review of Agricultural Economics,* Volume 28, Number 1 (2006): 4–23

Fisher, M.L. and K. Moore. “An Improved Credit Scoring Function for the St. Paul Bank for Cooperatives” ” *Financing Agriculture in a Changing Environment: Macro, Market, Policy and Management Issues – Proceedings of Regional Research Committee NC-161,* Dept. of Ag. Bus., S. Ill. Univ., Carbondale, 1987, p.196-216.

Gallagher, R. L. “Distinguishing Characteristics of Unsuccessful versus Successful Agribusiness Loans.” *Agricultural Finance Review.* 61(2001):19-36.

Gloy, B.A., M.A. Gunderson, and E.L. LaDue. “The Costs and Returns of Agricultural Credit Delivery.” *American Journal of Agricultural Economics*, 87:3(2005):703-716.

Greene, H.W. “*Econometric Analysis*” Prentice-Hall. Englewood Cliffs, NJ.
Gustafson, C., Pederson, G. and B. Gloy “Credit Risk Assessment” Working Paper, 2003

Hosmer D. W., Jr. and S. Lemeshow. “Applied Logistic Regression” Wiley Second Edition, New York 2000

Katchova, A.L. and P.J. Barry. “Credit Risk Models and Agricultural Lending.”, American *Journal of Agricultural Economics,* 87(2005):194-205.

Kohl, D.M. “Credit Analysis Scorecard*” Journal of Agricultural Lending* 1(1987):14-22.

Kohl, D.M. and S.O. Forbes, “Lending to Agribusiness: A Systematic Loan Analysis for Agricultural Lenders”, *Journal of Commercial Bank Lending* Feb. (1982):40-47.

LaDue, E.L. “The Farm Financial Standards Taskforce” *Financing Agriculture in a Changing Environment: Macro, Market, Policy and Management Issues – Proceedings of Regional Research Committee NC-161*, Dept. of Ag. Econ, Texas A&M Univ., College Station, May, 1990, p.84-97.

Long, S., “Regression Models for Categorical Dependent Variables Using STATA”, STATA *Press* Second Edition Texas 2006

Mooney, C.Z., and Duval, R.D. “Bootstrapping: A Nonparametric Approach to Statistical Inference”, *Sage Publications*, Newberry Park, Ca., 21. (1993).

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

Phillips, J., and A. Katchova. “Credit Score Migration Analysis of Farm Businesses: Conditioning on Business Cycles and Migration Trends” *Agricultural Finance Review* 64(2004):1-15.

Pregibon, D., “Logistic regression diagnostic”, *Annals of Statistics*, 9 (1981): 705-724.

Raftery, A. E., “Bayesian Model Selection in Social Research”, *In Sociology Methodology* Ed. P. V Mardsen, Oxford: Blackwell, vol 26 (1996):111-163

Tongate, R., “Risk Indexing: A Valuable Tool for Today’s Lender” *Agricultural Finance,* March (1984):12

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

Westgaard, Sjur and Van der Wijst, Nico, "Default Probabilities in a Corporate Bank Portfolio: A Logistic Model Approach." *European Journal of Operational Research*, 2001, 135(2), pp. 338-49.

Zech, L. and G. Pederson. “Predictors of Farm Performance and Repayment Ability as Factors for Use in Risk-Rating Models.” *Agricultural Finance Review.* 63(2003):41-54.

Ziari, H. A., D. J. Leatham and C. G. Turvey, "Application of Mathematical Programming in Credit Scoring of Agricultural Loans." *Agricultural Finance Review,* 55(1995):74-88.