The financial crisis of 2008 emphasized the need of financial institutions to measure and prepare for credit risk. In addition to the Dodd-Frank Wall Street Reform and Consumer Protection Act, which was enacted in 2010, regulators established an extensive list of new regulations and mandates for financial institutions relative to capital reserves. Although some financial institutions are exempt from Dodd-Frank, such as the Farm Credit System, their regulators are still acknowledging and requiring adherence to stricter standards with respect to capital adequacy. Traditionally, agricultural lenders minimize and mitigate credit risk through portfolio diversification, risk-based pricing, liens on collateral, personal guarantees, covenants, insurance, or other special conditions specific to a borrower’s situation. In addition to these traditional measures, regulators such as the Federal Reserve, Farm Credit...
Administration, Office of the Comptroller of the Currency, and the state insurance commissioners of the National Association of Insurance Commissioners require financial institutions with exposure to credit risk to maintain a level of capital adequacy, known as economic capital, sufficient to offset portfolio losses resulting from expected and unexpected loan defaults. How a given institution arrives at an estimate of the appropriate level of economic capital may differ greatly, and to date, no single method to derive an appropriate portfolio capital reserve has been required by regulators, although the Basel Accords provide guidance as to how capital adequacy may be achieved (Basel Committee on Banking Supervision 2006).

Previous studies measuring agricultural credit risk have often concentrated only on the risk of default without extending the analysis to include expected and unexpected losses to arrive at an adequate estimate of economic capital. In some instances combinations of ad hoc methods have been employed based on simple historical averages that have proven to be inadequate. This incomplete analysis does not provide a coherent estimate of the economic capital necessary to meet expected and unexpected loan losses. This may cause the portfolio risk manager to use poor estimates of expected and unexpected losses in determining necessary economic capital, leading to an inefficient use of capital or subjecting the financial institution to unplanned losses that regulators are specifically seeking to mitigate.

The goal of this study was to evaluate various methods of measuring agricultural mortgage portfolio credit risk using an extensive data set from one relatively large financial institution to determine what methods provide the best approach for determining an adequate estimate of economic capital. Various regression and other techniques were used to estimate expected losses based upon loan delinquencies, while Value-at-Risk and expected shortfall were used to estimate unexpected losses. The portfolio’s expected and unexpected loss estimates represent the recommended level of economic capital to cover potential loan losses.

Industry regulators and practitioners are seeking more robust and efficient methods to assess portfolio credit risk and estimate economic capital. Improved measurement of credit risk and economic capital directly benefits financial institutions through more accurate allocations of capital reserves, releasing funds for lending and operations. More efficient standardized credit risk assessment methods reduce the costs of regulators as well as government guarantee programs, and can reduce systemic market credit risks that lead to aggregate economic efficiency. This study helps the agricultural industry determine more optimal levels of economic capital for financial institutions with mortgage portfolios.

**Literature Review**

Few agricultural credit studies have modeled all aspects of portfolio credit risk management to arrive at an estimate for capital reserves. Pederson and Sakaimbo (2011) appear to be the exception, and found that historical systematic risks due to correlation between probability of default and loss given default through the business cycle can significantly increase the required economic capital for an agricultural lender. Most studies have focused on only one of the many components of credit risk assessment, for example Miller and LaDue (1989), who used logit regressions to assess the probability of default component, while others have compared credit vendor models to determine probability of default (Zech and Pederson 2004; Katchova and Barry 2005). Credit risk in agriculture has also been assessed as an embedded put option (Sherrick, Barry, and Ellinger 2000; Stokes and Brinch 2001), while some have determined the driving factors of prepayment and default (Dressler and Stokes 2010; Dixon et al. 2011). Featherstone and Boessen (1994) estimated a model of loan loss severity and risk premiums for a sample of farm real estate mortgages, and Yan and Barry (2006) assessed loss given default on a sample of farm real estate loans and found farmland pledged as collateral for a loan was expected to cover the loan balance with a relatively high probability. Early versions of loss given default models assumed that default rates and recovery rates are independent, but empirical evidence has shown otherwise (Altman, Resti, and Sironi 2004). As a result, extensions to loss given default modeling have been developed to include a non-zero correlation between default and recovery rates driven by a systematic factor. Others have viewed issues of default from a copula function approach (Anderson and Sidenius 2004). More recently, models have been developed that address the
dependency between idiosyncratic and systemic risk (Brownlees and Engle 2011).

A lengthy overview of exposure at default estimation concepts may be found in Engelmann and Rauhmeier (2006), while techniques are offered by Valvonis (2008) on how exposure at default may be estimated. Yang and Tkachenko (2012) compared the performance of various types of models that may be used to empirically measure exposure at default and loss given default including logistic regression, Bayesian models, and neural networks.

Portfolio loan loss distributions may be estimated by various methods depending on how the dependencies between portfolio sector loan losses are handled (Jouravlev and Maurer 2009). Differentiating between the sectors of an agricultural loan portfolio is critical since the sources of cash flow that are generated from the underlying assets that service debt repayment vary greatly, and accordingly are accompanied by differing risks. Of interest are the dependencies between the portfolio sector loan losses in the extreme right tail of a loan loss distribution since these dependencies directly affect the estimates of Value-at-Risk and expected shortfall (Klugman, Panjer, and Willmot 2008). Typically, parametric methods are implemented to estimate a loan loss distribution that at minimum control for the dependencies between the portfolio sector loan losses (Altman 2011). The resulting loan loss distribution features a long right tail to account for the potential of a severe economic downturn that could generate large portfolio loan losses. A common method used to model the dependencies (i.e., correlations) between the portfolio sectors is a copula.

Sklar (1959) showed that a multivariate joint probability distribution function may be constructed from a set of univariate marginal distribution functions and a selected copula. The dependence structure between the marginal distributions is captured by the copula and is independent of the form of the marginal distributions. With respect to loan portfolios, a copula may be estimated using a time series of historical portfolio sector loan loss rates (i.e., the marginal distributions) together with the estimated sector one-year ahead expected losses to generate a portfolio loan loss distribution from which Value-at-Risk and expected shortfall may be determined (McNeil, Frey, and Embrechts 2005). Although Value-at-Risk is typically the metric selected to determine capital reserves to cover unexpected losses, it has recently been criticized as a credit risk measure and one replacement metric that has been recommended is expected shortfall (Artzner et al. 1997). By construction, expected shortfall is always greater than or equal to Value-at-Risk as it represents the weighted average over the remaining right tail of the loan loss distribution above Value-at-Risk.

Estimation of Expected and Unexpected Losses in an Agricultural Mortgage Portfolio

The Basel Committee on Banking Supervision (BCBS) defines expected losses for a portfolio of loans as the composite formula

\[
POD_i \times LGD_i \times EAD_i = EL_i
\]

for each loan \(i\), where \(POD\) is the probability of default measured as a percentage, \(LGD\) is the loss given default measured as a percentage, \(EAD\) is the exposure at default measured in currency units, and \(EL\) is the expected loss for a loan measured in currency units. A portfolio loan loss distribution is then constructed from a probability density function estimated by coupling the one-year ahead estimate of expected loan losses with historical portfolio loan losses. From this portfolio loan loss distribution, estimates of unexpected loss represented as Value-at-Risk and expected shortfall are determined; together expected loss and unexpected loss determine economic capital.

Consider first the probability of default (\(POD\)) component of equation (1). Since the goal of a credit risk model is to determine an estimate of the appropriate economic capital reserve, the resulting estimate of probability of default should reflect the likelihood of a defaulted loan ending in foreclosure. One way to accomplish this is by multiplying the probability of delinquency by the conditional probability of a delinquent loan incurring a loss in the next period or one year. Therefore, the product of the probability of delinquency and the probability of loss is one way to represent the
probability of default as defined in the Basel framework.

Survival analysis was chosen to estimate the probability of delinquency because both the event of interest (i.e., delinquency) and the timing of the event are critical, as was identified in Dressler and Stokes (2010). Censoring occurs when the event of interest is not observed within the specified time period that the data span. Loan delinquency involves random right censoring, indicating that delinquency may occur after the end of the data time period. Standard methods of survival analysis require that random right censoring be non-informative, such that the censoring of a subject provides no additional information about the survival time of a subject since that information is not available (Allison 1995). In addition, the data may also be left truncated as some loans may have existed prior to the start of the observation period.

Associated with survival analysis is both a survivor and a hazard function. The survivor function reveals the probability of a loan surviving to time \( t \), while the hazard function, or conditional density, specifies the instantaneous rate of failure at \( T=t \), given that the loan has survived up to time \( t \). The hazard rate is more useful when interpreting survival data because it represents a measure of risk throughout time, with greater hazard rates coinciding with a greater risk of the event occurring.

In the context of loan delinquency, the basic survivor function \( S(t) \) represents the probability that a loan will survive to at least time \( t \) is

\[
S(t) = P(T \geq t) = 1 - F(t) = \int \limits_{t}^{\infty} f(x) \, dx
\]

where \( T \geq 0 \), and \( F(t) \) denotes a cumulative distribution function with the corresponding probability density function denoted by \( f(x) \). A summary of survival estimation in economics is found in Keifer (1998). In this study more detailed survivor methods incorporating recurrent delinquencies are estimated, using the intensity model of Anderson and Gill (1982) and proportional rate and mean model of Lin et al. (2000). A brief introduction to these models is provided below and the interested reader may find further details with respect to these models in a supplementary appendix online.

The Andersen and Gill (1982) model is considered a counting process with the Cox (1972) type of intensity function for recurrent events and may be defined as

\[
\lambda_k(t) = \exp(x_i(t_i)^* \beta) \lambda_0(t)
\]

where \( \lambda_0(t) \) is the baseline hazard function assumed equal for every failure type. The model consists of two components: (1) all the influence of the prior events on future recurrences, if any exist, is mediated through the time-varying covariates, and (2) the covariates have multiplicative effects on the instantaneous rate of the counting process. In the Lin et al. (2000) model, the baseline hazard function is allowed to differ by failure type.

The method used to obtain the survival analysis model predicted probabilities is from the empirical cumulative hazard function of the survivor function. For the multiplicative hazards model the cumulative hazard function may be defined as

\[
\hat{\Lambda}(t, x) = \sum_{i=1}^{n} \int_{0}^{t_i} \frac{dN_i(s)}{\sum_{j=1}^{n} Y_j(s) \exp((x_j - x)^* \beta)}
\]

where \( N_i \) is a counting process for the \( i \)th subject, \( Y_i \) is the censoring indicator for the \( i \)th subject, \( x_i \) is a vector of covariates with respect to the \( j \)th subject, and \( x \) is a vector of the reference covariate value categories that represents an arbitrary value of the \( j \)th covariate. Therefore, the empirical cumulative hazard function estimate of the survivor function for \( x = x \) may be determined as \( \hat{S}(t, x) = \exp(-\hat{\Lambda}(t, x)) \), with first derivative \( \hat{S}'(t, x) = -\exp(-\hat{\Lambda}(t, x)) \hat{\Lambda}(t, x) \). By definition, \( h(t_i) = -\frac{\hat{S}'(t_i, x)}{\hat{S}(t_i, x)} = \hat{\Lambda}(t_i, x) \), and \( H(t_i) = -\frac{\hat{S}'(t_i, x)dt}{\hat{S}(t_i, x)} = \hat{\Lambda}(t_i, x) \cdot t_i \). For this study the one year ahead probability of delinquency may be determined by taking the difference in the empirical cumulative hazard function defined as

\[
t_i \cdot p(t_i) = 1 - \exp(-(H(u) - H(t_i)))
\]

where \( u = t_i + 366 \) represents the interval of time over which the probability of delinquency is desired, noting that 2012 was a leap year.
Delinquency data rather than defaults directly were used to model the event of default primarily because delinquencies provided many more observations, as well as the interest in modeling delinquencies by the financial institution that provided the primary data. As a result the resulting probability of delinquency is augmented by a probability of loss in order to appropriately represent the likelihood of a loan incurring a loss. A loan loss is defined as the time when a loss is realized on a defaulted loan. The probability of loss is estimated from the time when a loss is realized on a defaulted loan incurring a loss. A loan loss is defined as the default weighted loss given default, a linear regression predicted value. The fixed estimate using both a fixed estimate and a linear regression predicted value. The fixed estimate to a loan default and in this study is estimated to a loan default and is calculated as

\[
PL_{i|PD_{del,i}} = \frac{\text{count of loans that experienced a loss within one year of delinquency}}{\text{count of delinquent loans}}.
\]

The probability of delinquency is multiplied by equation (6) to reflect the likelihood of a delinquent loan ending in default within one year. Regression analysis was not used to identify the determinants of the probability of loss due to the limited number of observations in the primary data.

Next consider the loss given default (LGD) component of equation (1). Loss given default represents the percentage of the principal balance outstanding that is lost to a loan default and in this study is estimated using both a fixed estimate and a linear regression predicted value. The fixed estimate is defined as the default weighted loss given default and is calculated as

\[
\bar{\mu} = \frac{\sum_{i=1}^{N} \mu_i}{\sum_{i=1}^{N} PBO_{loss,i}}
\]

where \(\mu_i\) represents the loss given default for loan \(i\), and \(PBO_{loss,i}\) represents the principal balance outstanding at the time of loss, with

\[
\bar{\sigma}^2 = \frac{\bar{\mu}}{N} \sum_{i=1}^{N} (\mu_i - \bar{\mu})^2.
\]

As an alternative to fixed estimates of loss given default, a linear regression model incorporating a set of loan-specific and economic factors was estimated. The third component of equation (1) is exposure at default (EAD), which in this study represents the principal balance outstanding at the time expected losses were estimated. This estimate ignores any costs associated with the foreclosure and liquidation process, and also ignores the forgone future discounted cash flows from scheduled interest payments. Also, remaining balances on revolving lines of credit that are secured by real estate may be overstated or understated since the principal balances may fluctuate during different periods throughout a given year. The limitations in the measurement of exposure at default in this study are considered minimal and would not otherwise influence the results presented.

The estimates of expected losses for each portfolio sector, obtained by equation (1), were used to determine predicted sector expected loss rates over the next year. The sector-specific expected loss rate for each sector \(i\) is defined as

\[
\text{expected loss rate}_i = \frac{\text{expected losses}_i}{\text{principal balance outstanding}_i}
\]

where the index includes time \(t+1\) to indicate the estimates are forward looking one period or year. These expected loss rates were combined with the historical sector loss rates for each sector \(i\) and is defined as

\[
\text{historical loss rate}_i = \frac{\text{actual losses}_i}{\text{principal balance outstanding}_i}
\]

where the index includes time \(t-n\), \(n = 0, 1, \ldots, k-1\) depending on how many \(k-1\) historical periods plus the current period of data were available, which in this study is nine years (i.e., 2003–2011). The expected and historical loss rates from equations (8) and (9) were combined to derive a loss distribution. The loss distribution that results from the combination of the expected and historical loss rates for the different portfolio sectors was used to determine an estimate of unexpected losses. Given the existence of multiple portfolio sectors, the cross-sectional correlations between these sectors will influence the resulting loss distribution. At minimum, accounting for the cross-sectional dependence between portfolio sector loss rates is necessary. One method that may account for cross-sectional dependence between the
sector loss rates of the portfolio is the copula-based loss distribution.

A copula is a multivariate function that describes the dependence between marginal distribution functions, and is useful in risk management for describing the dependence of extreme outcomes such as loan losses. In addition, copulas may easily be simulated, which is useful in the mortgage portfolio setting as it is often the case that the marginal behavior of individual risk factors is better understood than the dependence structure between portfolio sectors (McNeil, Frey, and Embrechts 2005). Several different types of copulas exist depending on the dependence structure that is desired that links the marginal distributions represented by each portfolio sector.

The Student’s $t$-copula (i.e., $t$-copula) was used to model the dependency between portfolio loss rates between four portfolio sectors. The $t$-copula was selected because it allows for greater probability mass in the tails of the multivariate distributions. This is appealing to the risk manager as the $t$-copula captures the likelihood of extreme events, or events that have a very low likelihood of occurring. Given these features, the $t$-copula complements the process in determining estimates of Value-at-Risk and expected shortfall. Kole, Koedijk, and Verbeek (2007) compared Gaussian and Gumbel copulas to the $t$-copula but considered those inferior for use with credit risk data. These authors show that the Gaussian copula tends to underestimate the likelihood of unexpected events while the Gumbel copula overstates the likelihood of unexpected events.

The first part of this section was used to develop the components of equation (1), which is the measure of expected loss. The remainder of this section focuses on unexpected loss, as measured by Value-at-Risk and expected shortfall. These two unexpected loss measures are obtained from the loss distribution, as described above. Value-at-Risk ($VaR_a$) has been one of the most widely used risk measures in financial risk management. Let $F_L(l) = \Pr(L \leq l)$ represent the distribution function of a corresponding loss distribution. Given some confidence level $z$, the $VaR_a$ of a portfolio is given by the smallest number $l$ such that the probability that the loss $L$ exceeds $l$ is no larger than $1 - z$. However, a criticism of $VaR_a$ is that it is not subadditive. Subadditivity requires that the aggregate risk for two combined risks can be no more than the risks individually summed together (Artzner et al. 1999).

A risk measure closely related to $VaR_a$ that is subadditive is expected shortfall ($ES_a$), and it is considered to be a coherent risk measure (Acerbi and Tasche 2002). Specifically, for a loss $L$ with $E([L]) < \infty$ and distribution function $F_L$, the expected shortfall at confidence level $z \in (0, 1)$ is defined as

$$ES_z = \frac{1}{1 - z} \int_z^1 q_u(F_L)du$$

where $q_u(F_L) = F_L^1(u)$ is the quantile function of $F_L$; $ES_z$ is related to $VaR_a$ by

$$ES_z = \frac{1}{1 - z} \int_z^1 VaR_a(L)du.$$

By construction, $ES_z \geq VaR_a$, and $ES_z$ may be interpreted as the expected loss that is incurred in the event that the $VaR_a$ is exceeded (McNeil, Frey, and Embrechts 2005). Therefore, $ES_z$ is an average (i.e., weighted average) over all the levels $u \geq z$ of the loan loss distribution.

**Data**

The primary data consists of loan information from a financial institution’s agricultural mortgage portfolio that includes four agricultural sectors: annual crops, livestock production, permanent crops, and agribusiness. As of December 31, 2012 the annual crops sector represented 45% of the portfolio principal balance outstanding, the livestock production sector represented an estimated 15%, the permanent crops sector represented an estimated 22%, and the agribusiness sector represented an estimated 18%. To preserve the identity of the data source, the data are not summarized, although the construction and definitions of the variables are discussed.

The majority of the sample loan-level data originated during the time period beginning on January 1, 2003 and ending on December 31, 2011, although some loan originations are dated prior to January 1, 2003 (i.e., left truncated observations). Primary data include the dependent variables and several quantitative and qualitative independent variables that
represent characteristics of the loan or borrower. Secondary data include quantitative independent variables that represent macroeconomic characteristics.

Loans in the portfolio are classified into four types based upon payment structure: interest only (ION), level installment payments (LIP), fixed principal amount with the same payment frequency for principal and interest (FSN), and fixed principal amount with different payment frequency for principal and interest (FDN). The annual crops, livestock production, and permanent crops sectors are dominated by the LIP and FDN types. The agribusiness sector is dominated by the LIP type. Loan interest rates are classified as fixed (FIX), adjustable (ARM), or variable (VAR), with the annual crops, livestock production, and permanent crops sectors dominated by the ARM interest rate type. The agribusiness sector is dominated by the FIX and VAR interest rate types.

Loan delinquency is defined as ≥90 days past due on the interest portion of the loan payment. The interest portion of the payment is the determining factor of delinquency, as typically principal payments may be deferred for a period of time to allow for a recovery considering the situation restricting scheduled loan repayment. The loan data includes measures of loan delinquency that include a binary indicator representing whether a loan experienced delinquency in a given period of time, and the age of the loan measured as the number of days from origination at the time of delinquency. Recurrent delinquency loans had multiple binary indicators across time.

The probability of loss was estimated from the loans that experienced a loss during the time period covering 2003 to 2011. For a few loans that experienced a loss, only the year of loss was available. For those loans, an estimate of the time to loss was obtained by randomly selecting simulated normally distributed truncated time to loss data calibrated on the average time to loss for the loans for which complete time data were available.

The sample of loss given default loans covers the time period of 1999 to 2011 for the fixed estimate of loss given default. Missing data at the loan level precluded the 1999 to 2002 data from being included in the regression analysis; therefore, the linear regression of loss given default utilizes loss given default data that covers the time period of 2003 to 2011. Exposure at default was represented by the principal balance outstanding at the time that expected losses were estimated.

The independent loan level and macroeconomic variables used in estimating the recurrent survival analysis regressions and the loss given default linear regression for the annual crops sector of the portfolio sample are summarized in tables 1 and 2. The independent variables represent values either at loan origination or across the loan life.

Results and Discussion

The focus of this study was to specify and estimate the probability of delinquency, probability of loss, loss given default, and exposure at default, to use those estimates to arrive at expected losses, and then arrive at unexpected losses using Value-at-Risk and expected shortfall techniques. The annual crops sector from the sample of loans is used to detail the various estimates, although the other sectors were also estimated and incorporated into the final aggregate portfolio estimates presented in table 8. The interested reader can find the other three sector estimates in a supplementary appendix online.

The variables included in the survival analysis regressions are described in tables 1 and 2. A summary of the delinquency data is listed in table 3. The importance of considering recurrent delinquencies becomes evident because on an aggregate portfolio basis, approximately 52% of delinquencies in the sample were recurrent from 2003 to 2011.

As noted in the previous section, survival analysis in a recurrent delinquency framework was used to estimate the probability of delinquency. Two types of models that handle within group independence (i.e., Anderson and Gill 1982) and within group dependence (i.e., Lin et al. 2000) were estimated. A static variables only model was estimated and compared to the combination of static and time-varying variables model because credit risk assessment is only performed at time of loan origination at some financial institutions that do not regularly update borrowers’ financial information.

The results of the static variable only Anderson and Gill (1982) survival analysis regression were estimated using equation (3) and are presented in table 4. The dependent variable in equation (3) represents the counting process (i.e., cumulative count of delinquencies)
of recurrent delinquencies, and the independent variables represent loan level and macroeconomic factors. The parameter estimates were obtained using the Breslow likelihood method. All numeric variable parameter estimates have the anticipated signs validating that the model estimated the effect of each variable as expected. The interested reader may find the results of the Lin et al. (2000) survival analysis regression model in a supplementary appendix online.
Measures of liquidity, solvency, profitability, and controls for unobserved heterogeneity (i.e., categorical variables) are effective at explaining the variability in probability of delinquency. For example, the hazard ratio for loan to value at origination is 2.004, indicating the estimated percentage change in the hazard rate for a one unit increase in this variable is $((2.004 - 1) \times 100 = 100.4\%$). Thus, for a 0.1 unit increase (10%) in loan to value at origination there is a 10.04% increase in the hazard of delinquency. For categorical variables the hazard ratio is the ratio of hazards between the listed category and the reference category. For example, relative to reference category loan type LIP, loan type FDN is 1.648 times more likely to experience delinquency. Greater aggregate debt to asset ratio at origination and greater interest rate at origination are associated with greater delinquency risk. The cash flow coverage ratio at origination and the working capital at origination indicate that greater levels of profitability and liquidity at origination are associated with lower levels of delinquency risk.

Using the static variable regression model estimated in table 4, predicted probabilities of delinquency were estimated for each loan that was outstanding as of December 31, 2012. Missing values for any numeric variables in the sample were replaced with the median of the sample variable so a predicted probability of delinquency was made for each loan. The regression results are presented in table 4. The static variable regression model is based on the Anderson and Gill (1982) model for 134 delinquencies with reference categories loan type LIP, US census region West.

### Table 4. Annual Crops Sector Survival Regression Results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Static Variables Only</th>
<th>Static and Time-varying Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loan to value ratio at origination</td>
<td>0.695*** (0.134)</td>
<td>0.665*** (0.157)</td>
</tr>
<tr>
<td>Loan to value ratio</td>
<td></td>
<td>1.944</td>
</tr>
<tr>
<td>Debt to asset ratio at origination</td>
<td>2.125*** (0.511)</td>
<td>1.902*** (0.441)</td>
</tr>
<tr>
<td>Loan interest rate at origination</td>
<td>0.461*** (0.106)</td>
<td>1.586</td>
</tr>
<tr>
<td>Loan interest rate</td>
<td></td>
<td>6.702</td>
</tr>
<tr>
<td>Cash flow coverage ratio at origination</td>
<td>-0.002* (0.001)</td>
<td>0.998</td>
</tr>
<tr>
<td>Working capital at origination</td>
<td>-0.014 (0.006)</td>
<td>0.986</td>
</tr>
<tr>
<td>Real loan origination amount</td>
<td>0.002* (0.002)</td>
<td>1.002</td>
</tr>
<tr>
<td>Real State cropland Pct. Change (1-1)</td>
<td></td>
<td>-0.007* (0.004)</td>
</tr>
<tr>
<td>Real State exports (1-1)</td>
<td></td>
<td>0.993</td>
</tr>
<tr>
<td>Loan type FDN</td>
<td>0.499* (0.207)</td>
<td>0.344* (0.199)</td>
</tr>
<tr>
<td>Loan type FSN</td>
<td>0.337 (0.298)</td>
<td>0.424 (0.273)</td>
</tr>
<tr>
<td>Loan type ION</td>
<td>0.659 (1.057)</td>
<td>-0.494* (0.762)</td>
</tr>
<tr>
<td>U.S. Census Region Midwest</td>
<td>-0.668 (0.287)</td>
<td>0.036 (0.279)</td>
</tr>
<tr>
<td>U.S. Census Region South</td>
<td>0.423 (0.286)</td>
<td>0.443 (0.279)</td>
</tr>
<tr>
<td>−2 Log Likelihood</td>
<td>1,315.95 (1,431.86)</td>
<td>1,414.61 (1,586.99)</td>
</tr>
</tbody>
</table>

Note: Asterisks *, **, and *** represent statistical significance at the 0.10, 0.05, and 0.01 levels, respectively. Results are estimated using Breslow likelihood with ties in failure time method on the Anderson and Gill (1982) model for 134 delinquencies with reference categories loan type LIP, US census region West.

### Table 3. Summary of First and Recurrent Delinquencies 2003–2011 (i.e., ≥ 90 days past due interest)

<table>
<thead>
<tr>
<th>Sector</th>
<th>First Delinquency</th>
<th>Recurrent Delinquency</th>
<th>Mean (St. Dev.) of Loan Age at First Delinquency (Years)</th>
<th>Mean (St. Dev.) Loan Age at Recurrent Delinquency (Years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual crops</td>
<td>201</td>
<td>189</td>
<td>6.26 (5.10)</td>
<td>7.19 (4.67)</td>
</tr>
<tr>
<td>Livestock production</td>
<td>50</td>
<td>82</td>
<td>4.66 (3.28)</td>
<td>5.38 (3.46)</td>
</tr>
<tr>
<td>Permanent crops</td>
<td>80</td>
<td>78</td>
<td>6.48 (4.67)</td>
<td>6.77 (3.76)</td>
</tr>
<tr>
<td>Agribusiness</td>
<td>35</td>
<td>43</td>
<td>1.68 (1.85)</td>
<td>4.40 (3.28)</td>
</tr>
<tr>
<td>Portfolio</td>
<td>368</td>
<td>406</td>
<td>5.78 (4.87)</td>
<td>6.40 (4.16)</td>
</tr>
</tbody>
</table>

Note: First delinquencies are the number of first delinquencies observations in each of the lending sectors. Recurrent delinquencies reflect that some loans recover from the first delinquency but then become delinquent again.
probability of delinquency could be obtained for each loan in the sample. The median was chosen compared to the mean as the numeric variable values for some loans were large enough to skew the mean. The probability density of the predicted probabilities of delinquency is shown in figure 1.

The limiting feature of these models is that the time-varying measures of data are excluded from the analysis. Financial institutions often only have access to information when the loan originates, with little or no updating of conditions over the life of a loan. However, the risk of delinquency changes over the life of a loan and is influenced by the current economic environment. Therefore, the results of incorporating loan level and macroeconomic time-varying variables within the Anderson and Gill (1982) multiplicative hazards regression model with loan type LIP and U.S. Census Region West as reference categories are also shown in table 4.

Compared to the static model, the static and time-varying model in table 4 includes two macroeconomic variables: real state cropland percentage change (t-1) and real state exports (t-1). In addition, the model excludes the measures of cash flow coverage ratio at origination and real loan amount at origination. Instead, the current values of loan to value ratio and loan interest rate were used. Real state cropland percentage change (t-1) and real state exports (t-1) replace the profitability measure of cash flow coverage ratio at origination.

Inspection of the parameter estimates and the test statistics of the static and time-varying variables model in table 4 indicate that measures of liquidity, solvency, profitability, and controls for unobserved heterogeneity (i.e., categorical variables) are effective at explaining the variability in probability of delinquency within the annual crops sector from 2003 to 2011. The debt to asset ratio at origination variable remains the most influential variable, as a 10% increase is associated with a 57.02% increase in the hazard of delinquency.

The increase in significance of the loan interest rate given the other variables in the model is expected as the current interest rate directly affects scheduled debt repayment. The significance in the real state cropland percentage change (t-1), and real state exports

Figure 1. Probability density of the survival analysis predicted probabilities of delinquency for the static variables only model in table 4

Note: For scaling reasons, 4.74% of the data included in the < 0.02% category were excluded from the probability density.
variables indicates the importance that current profitability has on the risk of delinquency. Working capital at origination is not as significant as in the static variable-only model.

Using the static and time-varying variable model estimated in table 4, predicted probabilities of delinquency were estimated for each loan that was determined to be outstanding as of December 31, 2012. The probability density of the predicted probabilities of delinquency is plotted in figure 2. The predicted probabilities are more left-skewed than the predicted probabilities of delinquency in figure 1 derived from the static survival analysis regression model.

The probability of loss was estimated as defined in equation (6). A time period of one year was selected to match the time period that expected losses and unexpected losses are estimated. In addition, separate probability of loss estimates were estimated for first and last defaults since some loans have recurrent defaults. Table 5 summarizes the probability of loss in one year for last defaults in the loan sample for each of the four loan sectors.

The probability of loss in one year is greatest for the agribusiness sector. The low incidence of probability of loss in one year after delinquency for the annual crops and livestock production sectors may be attributed to the length of time the legal process of foreclosure takes in the state where the loan is located. The low incidence of probability of loss in one year for the annual crops and permanent crops sector may also be attributed to

![Figure 2. Probability density of the survival analysis predicted probabilities of delinquency for the static and time-varying variables model in table 4](https://academic.oup.com/ajae/article-abstract/98/5/1470/241554)

Note: For scaling reasons, 4.57% of the data included in the < 0.02% category were excluded from the probability density.

<table>
<thead>
<tr>
<th>Primary Property Type</th>
<th>Average Days to Loss from Last Default</th>
<th>Probability of Loss in One Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual crops</td>
<td>622</td>
<td>1.99%</td>
</tr>
<tr>
<td>Livestock production</td>
<td>577</td>
<td>6.00%</td>
</tr>
<tr>
<td>Permanent crops</td>
<td>357</td>
<td>8.33%</td>
</tr>
<tr>
<td>Agribusiness</td>
<td>454</td>
<td>22.86%</td>
</tr>
<tr>
<td>Portfolio</td>
<td>510</td>
<td>5.95%</td>
</tr>
</tbody>
</table>

Note: Average Days to Loss from Last Default is estimated as the average of the number of days between the last occurrence of default as measured by ≥ 90 Days Past Due Interest and the day a loss was realized. Probability of Loss in One Year is estimated as defined in equation (6).
the likelihood of a delinquent loan recovering to a current status. In addition, based on the observations from this sample these two sectors, on average, are less leveraged than the other sectors, creating a greater likelihood of recovery in less than one year as these borrowers are able to extract liquidity from their unleveraged equity to become current on scheduled loan payments.

Two different estimates of loss given default were estimated: fixed estimate based on the empirical default weighted mean loss as specified by equation (7) and linear regression. The fixed estimates of loss given default in table 6 represent the default weighted mean loss given default based on sample loan loss data covering 1999–2011. Over the thirteen-year period, loss given default was lowest for the annual crops sector and highest for the agribusiness sector. The variance of the loss given default in the loss loan sample was greatest for the agribusiness sector and lowest for the annual crops sector.

Linear regression methods were used to analyze the relationship between loss given default as defined in equation (7) and a set of independent variables at the loan and macroeconomic level. The results of the static variable only linear regression presented in table 7 includes 26 observations due to missing data, and has loan type FDN, interest rate type FIX, and timber as reference categories. The timber sector was previously excluded from the probability of delinquency regressions because there was an insufficient number of timber defaults to estimate survival analysis models. One notable difference between the probability of delinquency models and the loss given default models is that while the probability of delinquency models are loan sector-specific, the loss given default models use all agricultural portfolio loans and therefore are portfolio specific. Thus, timber was included in the loss given default models because the data from the timber sector, although limited, was deemed useful.

Table 6. Summary of the Default Weighted Mean and Variance of Loss Given Default 1999–2011

<table>
<thead>
<tr>
<th>Sector</th>
<th>Default Weighted mean LGD</th>
<th>Variance LGD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual crops</td>
<td>9.27%</td>
<td>1.41%</td>
</tr>
<tr>
<td>Livestock</td>
<td>25.13%</td>
<td>5.46%</td>
</tr>
<tr>
<td>production</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Permanent crops</td>
<td>28.71%</td>
<td>2.69%</td>
</tr>
<tr>
<td>Agribusiness</td>
<td>34.96%</td>
<td>5.79%</td>
</tr>
<tr>
<td>Portfolio</td>
<td>29.34%</td>
<td>3.91%</td>
</tr>
</tbody>
</table>

Table 7. Loss Given Default Linear Regressions

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Static Variables</th>
<th>Static and Time-varying Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Std. Error</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.049</td>
<td>0.696</td>
</tr>
<tr>
<td>Debt to asset ratio at origination</td>
<td>0.371**</td>
<td>0.152</td>
</tr>
<tr>
<td>Real GDP all industries (t-1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annual crops</td>
<td>−0.168</td>
<td>0.620</td>
</tr>
<tr>
<td>Livestock production</td>
<td>0.072</td>
<td>0.626</td>
</tr>
<tr>
<td>Permanent crops</td>
<td>0.177</td>
<td>0.622</td>
</tr>
<tr>
<td>Agribusiness</td>
<td>0.074</td>
<td>0.622</td>
</tr>
<tr>
<td>Interest rate type VAR</td>
<td>0.302***</td>
<td>0.063</td>
</tr>
<tr>
<td>Interest rate type ARM</td>
<td>0.018</td>
<td>0.043</td>
</tr>
<tr>
<td>Loan type ION</td>
<td>−0.330</td>
<td>0.320</td>
</tr>
<tr>
<td>Loan type FSN</td>
<td>0.090</td>
<td>0.320</td>
</tr>
<tr>
<td>Loan type LIP</td>
<td>−0.078</td>
<td>0.320</td>
</tr>
</tbody>
</table>

Note: Asterisks *, **, and *** represent statistical significance at the 0.10, 0.05, and 0.01 levels, respectively. As a reviewer noted, with the addition of the GDP variable in the static and time-varying variables model, which varies over time, this is a multilevel model and thus the degrees of freedom are greater than with the static model. As a result we identify statistical significance at only the 0.01 level. Loss given default represents the dependent variable as defined in equation (7). Results are based on linear regression using 26 delinquencies with reference categories Timber, Interest rate type FIX, and loan type FDN. Heteroscedasticity-consistent standard errors are presented for both regression models.
not a significant addition to the model as security valuations are typically not adjusted on a regular basis for the data set this study analyzes. This is often true in agriculture as it may take several months to obtain a third party appraisal, and the appraisals are typically costly due to the unique land features and improvements. One other notable feature of the model is that the annual crops sector is associated with lower levels of loss given default relative to the timber sector as the empirical data suggested.

The results of a static and time-varying variables linear regression is also presented in table 7. For comparison purposes with the static variables-only model, the linear regression includes the real GDP all industries (t-1) macroeconomic variable, and although it is not significant, the sign on the parameter estimate is as expected. The probability density of the predicted loss given default estimates for the static and time-varying model for the annual crops sector are presented in table 7 and are shown in figure 3.

Economic capital set aside for possible loan losses must be sufficient to meet both expected and unexpected loan losses, although in any given year expected losses is what should happen. Expected loss estimates were obtained from estimating the probability of delinquency, probability of loss, loss given default, and exposure at default as previously discussed. Given the combination of models selected, different estimates of expected loss results are summarized in table 8 for the annual crops sector (i.e., Annual Crops EL column) and then the total loan portfolio (i.e., Portfolio EL, Portfolio VaR (99%), Portfolio ES (99%), Portfolio UL with Base VaR (99%), and Portfolio UL with Base ES (99%) columns) that also includes permanent crop, livestock production, and agribusiness sectors. On a portfolio basis predicted expected losses estimated from the various models ranged from $421 per $1MM of portfolio principal balance outstanding to $2,733 per $1MM of portfolio principal balance outstanding, which although relatively low on a $1MM basis, is a range from low to high of over six multiples.

Various portfolio loss distributions were determined using the t-copula with the historical loss rate data to generate unexpected losses. Estimates of Value-at-Risk and

![Figure 3. Annual crops probability density of the linear regression static and time-varying variables predicted loss given default for the model in table 7](image)

Note: For scaling reasons, 54.27% of the data at a predicted value of 0% were excluded from the probability density.
Overall, on a portfolio basis, predicted VaR ranges from $4,837 per $1MM of portfolio principal balance outstanding to $4,936 per $1MM of portfolio principal balance outstanding. Predicted portfolio ES ranges from $5,046 per $1MM of portfolio principal balance outstanding to $5,166 per $1MM of portfolio principal balance outstanding. The expected shortfall model arrives at a slightly greater value for unexpected loss and thus would require a more conservative allowance for economic capital.

Potential loan losses can result from both expected and unexpected losses. The combined average predicted portfolio loss ranged from $3,686 per $1MM of portfolio principal balance outstanding when combining the expected losses with the unexpected losses derived from the VaR model, to $3,897 per $1MM of portfolio principal balance outstanding when combining expected losses with unexpected losses from the expected shortfall model.

The economic capital estimates derived from the survival analysis probability of delinquency best match the empirical loss data from the financial institution studied. Specifically, the models with static and time-varying variables to estimate expected losses came closest to the actual portfolio loss of $2,190 per $1MM of principal balance outstanding that was observed at year-end 2012. This assumes that the observed losses were the expected value from the actual loss distribution.

**Conclusion**

The probability of delinquency, probability of loss, loss given default, and exposure at default were estimated and used to determine expected and unexpected losses from loan-level and industry-level data using various methods. The primary data consists of loan information from a financial institution over the period 2003–2011 including four agricultural sectors: *annual crops*, *livestock production*, *permanent crops*, and *agribusiness*, with secondary data from published sources.

Probability of delinquency was analyzed from recurrent delinquency data using survival analysis. Two types of models were estimated using static and a combination of static

---

**Table 8. Portfolio Expected and Unexpected Losses with Various Model Combinations (Per $ Million of Principal Balance Outstanding)**

<table>
<thead>
<tr>
<th>Probability of Delinquency</th>
<th>Linear Regression Fixed Estimate</th>
<th>Linear Regression Static and Time-Varying Variables</th>
<th>Linear Regression Static Variables Only</th>
<th>Annual Loss</th>
<th>Static and Time-Varying Variables Only</th>
<th>Static Only</th>
<th>Average</th>
<th>Portfolio EL (VaR) 99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>EL</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>$2,038</td>
<td>$2,038</td>
</tr>
<tr>
<td>Portfolio ES (99%)</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
</tr>
<tr>
<td>Portfolio UL with Base VaR</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
</tr>
<tr>
<td>Portfolio UL with Base ES</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
<td>$2,038</td>
</tr>
</tbody>
</table>

Note: EL is expected loss, VaR is value at risk, ES is expected shortfall, and UL is unexpected loss.
and time-varying loan-level and industry-level data (within group independent, and within group dependent settings) to compare results when controlling for correlation between recurrent delinquencies. Results showed that measures of liquidity, solvency, profitability, and controls for unobserved heterogeneity are important when measuring delinquency.

Loss given default was analyzed across each loan portfolio sector using a fixed estimate and linear regression. In addition, linear regression was used to model loss given default at an aggregate portfolio level given a set of static and a combination of static and time-varying variables. The results indicated that measures of leverage, state-level economic output, and controls for unobserved heterogeneity are important for explaining loss given default.

Expected loss estimates were determined for each sector and at the portfolio level using the predicted and fixed values of probability of delinquency, probability of loss, loss given default, and exposure at default as previously discussed and estimated. Sector estimates of expected losses were then coupled with the historical sector loss histories to develop a sample from which unexpected losses were estimated as Value-at-Risk or expected shortfall.

A t-copula was utilized to account for the correlation in loss rates between sectors. The copula was simulated to produce a random sample from which a loss distribution was estimated corresponding to each set of model combinations. From the estimated loss distributions Value-at-Risk and expected shortfall were determined. Slight differences in both estimates were observed. Unexpected losses were then estimated using either Value-at-Risk or expected shortfall. The results show the importance of accounting for the timing of delinquency when considering delinquency risk.

Few studies exploring agricultural credit risk have concurrently addressed all components of the economic capital estimation as was accomplished in this article. As such, the results are useful for both academics and practitioners to further understand credit risk and develop a more standardized approach to developing processes to determine estimates of economic capital that regulators would consider more coherent. Limited data were available as indicated for certain parts of the analysis, so future research could focus on finding ways to incorporate methods that accommodate limited data when estimating economic capital for a loan portfolio. Other areas of improvement could include obtaining more data to allow the inclusion of model validation procedures, expanding on the analysis presented in this study through the use of more alternative credit risk methods, or developing new methods to analyze loan portfolio credit risk and economic capital estimation.

Supplementary Material

Supplementary material is available online at the American Journal of Agricultural Economics online at http://oxfordjournals.org/our_journals/ajae/.

References


