Agricultural and Rural Finance Markets in Transition

Proceedings of Regional Research Committee NC-1014
St. Louis, Missouri

October 4-5, 2007
Dr. Michael A. Gunderson, Editor
January 2008
Food and Resource Economics
University of Florida
PO Box 110240
Gainesville, Illinois 32611-0240

Competing Risk Proportional Hazard Models of Farm Service Agency Direct Operating Loans

Prepared by

Bruce L. Dixon, Bruce L. Ahrendsen, Monica Foianini, Sandy Hamm, and Diana Danforth
Department of Agricultural Economics and Agribusiness
University of Arkansas-Fayetteville

ABSTRACT

The USDA Farm Service Agency (FSA) direct farm loan program is designed to provide credit to family-sized farms unable to obtain credit from conventional sources at reasonable rates and terms despite having sufficient cash flow to repay and an ability to fully securitize the loan. FSA policy encourages borrowers to exit the program as soon as possible. This study uses Cox proportional hazard models in a competing risks framework to identify predictive factors of: (1) loan success or default, and (2) length of time to loan termination. Survey data from 1925 direct loans originated in federal fiscal years 1994-96 are used for analysis. Only data available to FSA at time of origination were collected. Since these data are all the information FSA has at time of loan origination, the competing risk models provide an alternative method for measuring a priori relative riskiness indicated by borrower and loan characteristics. Results indicate that borrower financial strength, intensity of borrowers' current relationships with FSA and loan characteristics are significant measures of loan risk.

Key words: Duration, Farm Service Agency, direct loans, competing risks.

JEL Classifications: C29, G28, Q12, Q14

The Farm Service Agency (FSA) makes direct and guaranteed farm ownership (FO) and operating (OL) loans to family-size farmers and ranchers who cannot obtain commercial credit from a bank, cooperative Farm Credit System institution, or other lender. FSA loans are often provided to beginning farmers who cannot qualify for conventional loans because they have insufficient financial resources. FSA also directly lends to established farmers who have suffered financial setbacks from natural disasters in the form of emergency loans (EM). The direct farm loan programs administered by FSA are designed to provide credit to family-sized farms "unable to obtain credit from conventional sources at reasonable rates and terms" despite having sufficient cash flow to repay and an ability to provide security for the loan (Dodson and Koenig, p.1). The goal of FSA lending to farmers is to help farmers graduate to commercial sources as soon as possible. FSA programs are meant to be transitory programs and not a farmer's permanent or long term credit source. Farmers can be helped to achieve success in graduating from the FSA direct loan program in a timely manner by FSA understanding how pre-origination characteristics affect the outcome and duration of direct loans. Such information should assist FSA in helping borrowers exit as early as possible and also understanding how default numbers can be minimized.

This study uses data from a comprehensive, nationwide survey of individual, FSA direct loan borrowers to identify the factors affecting loan status and duration of FSA loans. Cox proportional hazard (PH) competing risk models using these data are estimated to model both loan status and loan duration. In this study the loan status is: paid-in-full, defaulted or still performing.[39] Competing risks models are estimated because loans can end in two different ways.[40] Competing risks PH models are very close in algebraic form to the multinomial logit model so the PH model can be used to predict the relative likelihood of one terminal status to the other. Cox PH competing risks models are estimated to identify those pre-loan origination factors that are significant predictors of which outcome is most likely and time until termination. As discussed shortly, a number of prior agricultural loan studies have used logit models or other choice models to identify indicators of successful and unsuccessful loans. A characteristic of such studies is that the length of the loan until termination is not incorporated into the analysis. Such information is essential in duration models. The ability to predict durations is valuable for both practical and estimation reasons. For example, information that a loan is likely to default early as opposed to later is valuable to a lender for obvious reasons. But knowing when a loan is likely to be paid-in-full helps both in pricing the loan and forecasting the outstanding balance of a loan portfolio more accurately. In addition, the duration information should be valuable in deriving more efficient parameter estimates than opposed to, say multinomial logit models that ignore such information.

In what follows the FSA direct loan program is described in more detail and prior research regarding loan duration studies is reviewed. The survey method and data are briefly described and relevant summary statistics are presented. Competing risks models are introduced and the estimation issues in using such models are briefly explicated. The empirical section interprets the estimates of the competing risks models. Implications of the findings are then discussed.

FSA Background Information

FSA is an agency of the United States Department of Agriculture (USDA). One of its functions is to act as a lender of last resort in compensating for credit gaps that prevent otherwise credit worthy borrowers from obtaining credit. A number of loan categories are covered by FSA.

---

[39] Defaulted includes loans that have experienced foreclosure, bankruptcy or debt write-off but are still not terminated.

[40] See Kalbfleisch and Prentice, Chapter 8, for an extensive introduction to the competing risks model.

These include OL loans which can be one year or seven years and are used to pay for operating and/or farmer living expenses. These loans typically cover variable input expenses and can also cover intermediate capital items such as farm machinery. Farm ownership loans can be used to purchase farm land, construct or repair buildings and to pay for soil and water conservation activities. Emergency loans can be used for a variety of expenses induced by a natural disaster. For the loans in the sample the maximum OL and FO indebtedness is $200,000 per category and EM loans could range up to $500,000. More detail on FSA loans is available in Nwoha et al.

In addition to serving the general category of family-sized farms, federal legislation compels FSA's FO and OL lending programs to target specific subgroups falling under the family farm umbrella. These groups are socially disadvantaged (SDA) and beginning farmer (BF) applicants. FSA defines an SDA farmer or rancher as "one of a group whose members have been subjected to racial, ethnic, or gender prejudice because of their identity as members of the group without regard to their individual qualities. For purposes of FSA programs, socially disadvantaged groups are women, African Americans, American Indians, Alaskan Natives, Hispanics, and Asian Americans and Pacific Islanders." (USDA/FSA, 2004a)

For definitional purposes, "A beginning farmer or rancher is an individual or entity that (1) has not operated a farm or ranch for more than 10 years; (2) meets the loan eligibility requirements of the program to which he/she is applying; (3) substantially participates in the operation; and, (4) for FO loan purposes, does not own a farm greater than 30 percent of the average size farm in the county. (Note: all applicants for direct FO loans must have participated in business operation of a farm for at least three years.) If the applicant is an entity, all members must be related by blood or marriage and all stockholders in a corporation must be eligible beginning farmers" (USDA/FSA, 2004b). General eligibility requirements include U.S citizenship, not being able to obtain credit elsewhere, an acceptable credit history and other requirements.[41]

To hasten borrower graduation to commercial credit sources, lifetime limits for FSA borrowers to receive direct credit were established for OL loans. Borrowers became ineligible for direct OL loans after receiving OL loans for 10 years, and ineligible for guaranteed loans after 15 years (USDA/ERS). However, as financial hardship increased in the late nineties and the ensuing decade, the Farm Security and Rural Investment Act of 2002 (2002 Farm Act, P.L. 107-171) legislated changes in FSA direct farm loan programs to loosen borrower eligibility rules. The eligibility time limits for direct OL loans were eliminated to allow longer access to FSA funded farm programs (USDA/FSA, 2004b). Nonetheless, current policy encourages direct loans to be paid back when the borrower is financially able to obtain loans from commercial lenders (USDA/FSA, 2007).

Past Loan Duration Studies

Duration models have been used in a variety of loan studies in the business literature, mainly focusing on home mortgages. Hakim and Haddad model conventional mortgage defaults as a function of borrower and loan characteristics. This study adopts an accelerated failure-time (AFT) methodology, which allows variations in the curvature of the baseline hazard (default) rate. The AFT model incorporates a baseline default hazard function and independent variables alter the hazard rate. The hazard rate is defined as total mortgages defaulting in month t divided by total mortgages still performing at the start of month t.

Disaggregated and regionally diversified observations which include a large set of borrower characteristics are used. The data consist of a nationwide random sample of 9,000 mortgages

[41] For more detail go to: http://www.fsa.usda.gov/dafl/directloans.htm#Eligibility, accessed June 1, 2005.

originated between 1973 and 1980 from Freddie Mac conventional mortgages. The study groups the explanatory variables into two categories: loan characteristics and borrower attributes. The principal loan characteristic variable is the loan-to-value ratio (LTV) at purchase. This variable measures the lender's risk to changes in collateral value. Several variables represent the impact of borrower characteristics on default. Monthly household income less fixed expenses measures repayment capacity. The authors use loan amount as a proxy for borrower wealth. Borrower characteristics include age, borrower's number of years on most recent job, number of household dependents and borrower location in one of six regions of the U.S.

Six models are estimated with one model per region. The authors interpret the results as indicating that high-risk borrowers are associated with small loans and high expense-to-income differentials. The loan principal coefficient is negative and significant, implying that higher loan amounts are related to lower likelihoods of default. The income-expense differential variable is generally significant and negatively signed, as expected. The significance of these coefficients indicates a sensitivity to income fluctuations. Loan-to-value is uniformly significant across regions and positively signed. Number of dependents increases the default probability. The borrower's age is insignificant in five of six regions. Borrower job tenure is insignificant in five of six regions.

Hakim and Haddad note that the default model could be paired with a prepayment model, thereby providing a framework for mortgage valuation. Such a proposal implies a competing risks model that would give a comprehensive analysis of when a mortgage is likely to terminate.

Jackson and Kaserman test two alternative default theories as a basis for interpreting and explaining empirical models of default risk on home mortgage loans. The first theory, referred to as the equity theory of default, posits that borrowers make default decisions by comparing the current value of equity with the discounted stream of debt service. The second view, labeled the ability to pay default theory, assumes mortgagors do not default on a loan as long as their income stream remains high enough to meet the debt service without undue financial burden.

The equity model relies on the market value of the mortgaged property at time t, the outstanding mortgage loan balance at time t, the loan-to-value ratio on the mortgage loan at origination, the contract interest rate, and the contract life of the mortgage obligation from origination to maturity. The ability to pay model requires that the default probability be positively related to loan-to-value ratio and interest rate as in the equity model but negatively related to maturity—the opposite from the equity hypothesis.

Using Federal Housing Administration data from 1736 individual mortgages as the observational unit, the relationships between default probability and financing terms (contract length, mortgage interest rate and loan-to-value ratio) were estimated by both ordinary least squares and the probit. The results support the equity theory of default behavior. All three financing variables are positive but only loan-to-value is significant. The findings emphasize the importance of borrowers having a positive equity position. The authors further note the low predictive power of their models and attribute this to the nature of the observations being individual behavior.

While including the loan-to-value ratio as an independent variable has undoubted theoretical justification, it is not always found to be significant. Lambrecht, Perraudin and Satchell estimate a Weibull duration model of United Kingdom mortgage defaults. The sample is limited to 5272 observed defaults and therefore does not have to deal with censoring issues of still performing loans or loans paid-in-full. As with our study the data are pre-origination and limited to a degree greater than ours. They utilize loan-to-value at origination, interest rate, salary and marital status

at time of loan origination. Results indicate loan-to-value is not important and ability to pay is important.

Ambrose and Capone examine the differences in the relative hazards of prepayment and first and second defaults. As part of their analysis they utilize a competing risks framework to estimate time to default or time to prepayment and use the exponential form of the hazard function. The observations used for this analysis are drawn from 724,666 Federal Housing Administration single-family residential mortgages originated in 1989 to give a sample of 42,764 for estimation after adjusting for location and missing data. The duration variable is the number of months between mortgage origination and default or prepayment. The independent variables are divided into financial characteristics of the borrower and loan. Financial variables include one variable indicating if prepayment is optimal as a function of interest rates, loan-to-value and payment-to-income ratios, and a variable indicating a negative equity position in the loan. Borrower non financial characteristics are income and race as well as the borrower's state unemployment rate.

Results indicate that falling interest rates after origination lead to more frequent prepayment. In the default model, an increasing probability of negative equity results in a higher default hazard. Loan-to-value at origination seems not to affect prepayment or default hazards. The negative race coefficient indicates minority borrowers have a lower probability of a first default than do non minority borrowers. Borrowers in the lower and middle income categories are more likely to prepay while upper income borrowers have lower prepayment and default probabilities.

Agricultural Loan Studies

Although the use of duration methods is not new to agricultural economics (Key and Roberts and Huffman and Feridhanusetyawan) we found no published studies where the durations of agricultural loans were analyzed in a competing risks framework or, for that matter, in a duration framework of any sort. However, a rich literature exists on agricultural lending in the form of credit scoring studies (Lufburrow, Barry and Dixon, Splett, Barry, Dixon and Ellinger, Turvey, and Limsombunchai and Lee). These studies identify financial and borrower related characteristics that are important in loan approval. These studies focus on variables observable at loan origination. Since such variables are considered important for loan approval, it follows that such variables should also be important in predicting time to loan termination. As discussed below, many variables in the agricultural credit scoring models are the same as those used in the duration studies of non-agricultural loans discussed above.

Turvey and Weersink estimate a logistic credit scoring model. The independent variables for the logistic credit scoring function are: liquidity ratio, debt-asset ratio, anticipated change in contribution margin defined as revenue minus variable costs, return on assets, repayment ratio, loan-to-security ratio, loan contract amount, interest rate charged, and a binary variable to indicate a loan used for refinancing. Regional and farm type variables are also included.

Results show that loan default decreases with increased liquidity (liquidity ratio), profitability (return on assets), repayment ability (repayment ratio), and security (loan-to- security ratio). Loan default increased with financial leverage (debt-asset ratio), change in contribution margin, loan amount, interest rate, and if the loan was for refinancing. The overall prediction of the credit scoring model is 61.3%. Non-current loans are predicted with less accuracy. Turvey observes that unanticipated factors such as death, divorce, drought or disability are idiosyncratic and therefore difficult to predict. Nonetheless, such factors can be very influential in loan payback performance.

An interesting aspect of the Turvey and Weersink analysis is that no non financial borrower characteristics other than region or farm type are included among the explanatory variables. As noted above with the non-agricultural loan duration studies, borrower characteristics have been included in loan duration models. Such variables also appear in some prior agricultural credit scoring models. Gallagher uses borrower and lending officer years of experience in addition to various financial and loan characteristics. Featherstone et al. use character, productive capacity and financial record keeping abilities as independent variables in loan approval but do not have demographic characteristics such as age, race or gender.

It appears race and gender are seldom used characteristics in agricultural loan studies. Escalante et al. focus directly on impacts of gender and race on the loan application approval process for FSA direct and guaranteed loans. In the logistic equations they estimate, applicant gender and race are independent variables along with many customary borrower financial measures. Dixon et al. estimate a multinomial logit model to identify the important variables known at loan origination that can predict if and how FSA direct loan borrowers exit FSA programs. They use the same data set used in this study. The exit status of borrowers initiating FSA direct loans in federal fiscal years 1994-1996 as of November 30, 2004 is the dependent variable in the logit model. Four "exit" possibilities exist for the borrower: (1) if borrower has active direct OL, FO or EM loans as of November 30, 2004, (2) if borrower has no active direct OL, FO or EM loans and is still farming using conventional sources of credit or no credit at all, (3) if borrower has no active direct OL, FO or EM loans and left farming voluntarily or retired, and (4) if borrower has no active direct OL, FO or EM loans and left farming involuntarily (other than death).

Age and race are significant but gender is not. Loan type is important as expected and numbers of existing farm ownership, operating and emergency loans at origination are also significant as are whether the loan is beginning farmer or socially disadvantaged. Debt-to-asset ratio and net worth are significant. Variables measuring repayment capacity, annual household income and a measure of income diversity from farm and non farm sources are not. Financial difficulty with federal loans prior to direct loan origination is a significant indicator for those exiting involuntarily.

Methods

Estimation of duration models focuses attention on the hazard function (Kalbfleish and Prentice) and is estimated in this application as well. The hazard function is based on the survival function that is derived from the underlying cumulative distribution function for the distribution of time, *F(t).* The distribution function, *F(t),* describes the probability that a variable T takes on a value less than or equal to a number t and it is defined as:

$$F(t) = \int_0^t f(s)ds = \Pr(T \le t). \qquad (1)$$

Where *f(s)* is the probability density function corresponding to *F(t).* Since durations can only be positive, the underlying cumulative distribution function can only be defined over the positive portion of the real line.

The survival function *S(t)* describes the probability that a variable *T* takes on a value greater than a number t and it is given by:

$$S(t) = 1 - F(t) = \Pr(T \ge t). \qquad (2)$$

The hazard function *h*(t) is the ratio of the probability density function *f(t)* to the survival function *S(t).* It is the rate at which durations are completed after duration t, given that they last at least until t:

$$h(t) = \frac{f(t)}{S(t)}.$$

The estimation problem is to specify a parametric function for *f(t)* and then estimate the implied parameters via the hazard function. Two popular forms for estimating *h(t)* are the Cox PH or the Weibull although only the latter is a fully parametric approach.

Competing risks models for the present application essentially estimate separate hazard functions for each of two possible outcomes. One hazard function is estimated for time to default and a companion hazard function for time to paid-in-full. All observations not experiencing a default are treated as censored in the default hazard function. Defaults are treated as censored in the paid-in-full model. The ratio of the respective hazard functions then gives the relative likelihood of one type of outcome to the other.[42]

A competing risks model consisting of individually estimated hazard functions with some observations censored requires that the censoring be uninformative (Kalbfleisch and Prentice). For the sample in this study, the censoring is partially random. This occurs because the loan origination dates of the loans were not fully controlled in the sampling scheme. Within a given state, loan type, borrower race and gender, loans in the sampling frame were ordered by origination date and then a sample selected by systematic sampling. This resulted in loans being distributed roughly uniformly over the three year time span. The censoring date was fixed at November 30, 2004 for all loans. This set the censoring time and could therefore be considered as Type I censoring but for the randomness of origination date.

With random censoring, the crucial issue is obtaining unbiased parameter estimates when estimating a competing risks model. The requirement for estimating a valid model is that the random censoring be uninformative. That is, given that a loan has lasted until the censoring date, does censoring at this point indicate anything special about how that loan might ultimately terminate? For example, suppose a borrower inherits a large sum and is immediately censored. The borrower's likelihood of paying in full or not defaulting clearly increased due to the inheritance so that the censoring is informative. Given that all loans originated on a given date are censored at the same duration length, there is no a priori reason to believe the censoring is informative and, therefore customary estimation methods are valid (Cox and Oakes).
The algebraic forms of the PH and Weibull models are well known. For the PH the hazard function is written as:

$$h(t) = \lambda_0(t) \exp(\boldsymbol{\beta}'\mathbf{x})$$

where $\lambda_0(t)$ which is an arbitrary function that is not estimated, $\boldsymbol{\beta}$ is a parameter vector to be estimated and $\mathbf{x}$ is a vector of covariates for observation i. For the Weibull,

$$h(t) = \gamma(\lambda t)^{\gamma-1}\exp(\boldsymbol{\beta}'\mathbf{x})$$

where the parameters γ and λ are estimated in addition to $\boldsymbol{\beta}$. For the purposes of estimating the ratio of hazards, the PH competing risks model must be estimated with a constant term.[43] With the presence of this constant term, the likelihood of the observation terminating (J) by termination type j out of m possible termination types is given as:

$$P(J = j \mid \mathbf{x}) = \exp(\alpha_j + \boldsymbol{\beta}_j'\mathbf{x})/[\exp(\alpha_1 + \boldsymbol{\beta}_1'\mathbf{x}) + \ldots + \exp(\alpha_m + \boldsymbol{\beta}_m'\mathbf{x})]$$

[42] In the empirical application the independent variables are fixed over time since they are only observed at origination. The ratio of the PH hazards are therefore constant over time but not for the ratio of Weibull hazards.

[43] For the purposes of parameter identification one constant term is restricted to zero across the J possible outcomes. Since J = 2 in the application, only one constant term is estimated.

In the empirical models the $\boldsymbol{\beta_j}$ are estimated by partial maximum likelihood is SAS 9.1.

Data

FSA farm loan managers (FLM) were surveyed in November and December of 2004 to collect data on a systematic sample of loans originated in FY 1994–1996 in the lower forty-eight states. These three years typified the 1990s with respect to net farm income. Starting sampling before FY 1994 would have resulted in a small sample of beginning farmers. Sampling later than FY 1996 would not have given sufficient time to observe a representative set of observations on loan durations for the longer maturing loans.[44]

Systematic sampling was used to guarantee proportionate loan type, gender, racial, geographical, and loan origination date representation. Because white males were proportionately so abundant, white males were sampled at a rate of one in eighteen for OL and EM loans and all other borrowers sampled at one in nine. Data on borrower demographics and financial variables at loan origination and loan status as of November 2004 were recorded including time to termination. Observations were taken from the application and farm and home plan (forms FSA-410-1 and 431-2).

There were 34,026 OL, 3,083 FO, and 8,359 EM loans originated in fiscal years 1994-1996. The survey generated 2,715 usable responses from a sample of 3,004 for a 90% response ratio. See Nwoha et al. for the complete survey instrument.

Descriptive Statistics

The mean borrower age at loan origination is 42 years which is considerably younger than the 1997 Census of Agriculture (USDA/NASS, 1999) mean farm operator age of 54. Borrowers have a mean of 17 years of farming experience although this is substantially lower for borrowers in beginning farmer programs. Mean number of family members is 3.2 with the two beginning, non-SDA farmer programs having fewer members. The total annual sales mean is $157 thousand and individual observations vary widely. The farm operators are overwhelmingly white and male with only seven percent of the loans being SDA although minorities did obtain a relatively few non SDA loans. About 77 percent of the loans went to repeat FSA borrowers.

Borrower financial variables show that the surveyed farms with mean assets of $282 thousand and mean net worth of $118 thousand are below the U.S. farm means ($441,000 in assets and $376,000 for net worth) for 1994-1996.[45] The mean farm debt-to-asset ratio (0.71) for the surveyed farmers exceeds that for U.S. farmers (0.15).

Table 1 contains descriptions of the loan status as of November 30, 2004 and the mean durations by loan type. Several prominent characteristics are clear. First, defaults for the one-year loans are long in coming. All four one-year loan types have an average of at least four and a half years to default. In contrast, the paid-in-full one-year loans are paid off near maturity as would be expected. The unconditional hazard functions are displayed in Figure 1. If a loan is going to be paid-in-full, it is most likely to happen in the first two years. In contrast, the default hazard is low in the early years and increases as time increases. Eventually, the default hazard overtakes that of

[44] As an example, data were collected on farm ownership loans which can have up to a forty year maturity. But there simply were not enough terminating observations to allow for reliable estimates on this loan type.

[45] The U.S. farm data were obtained at: http://www.ers.usda.gov/Data/farmbalancesheet/fbsdmu.htm. Accessed January 28, 2007.

paid-in-full as would be expected. It should be noted that the hazard functions past two years are based on relatively few loans.

Seven-year loans have a much different duration experience. The mean paid-in-full durations on these loans is less than the seven-year maturity by roughly a year. Figure 2 displays the unconditional hazard functions for the seven-year loans. The hazard for paying in full peaks at seven years. A number of the seven-year loans are paid back earlier. Defaults are unlikely in the first six years but the default hazard increases over time. As time increases, the likelihood of default and paying in full start to converge.

Surprisingly, the mean seven-year loan default times are less than the maturity although this is somewhat misleading. As of November 2004, only nine percent of the one-year loans were still performing whereas 35 percent of the seven-year loans were still performing. As these proportionately more numerous still performing loans are terminated, it is much more probable they will be defaults compared with the loans paid back fully in the sample as of November 2004. Since the latest a loan could be originated in the sample was September 30, 1996, all OL loans should have been paid back. So the remaining performing loans are likely to be problem loans. Consequently the mean default time for the seven-year loans is likely to lengthen considerably.

Table 1 also displays an ultimately limiting aspect of the sample. While the number of FO loans is not small, the proportion of defaulted FO loans is quite small making any estimation in a competing risks framework suspect. Since FO loans can have maturities up to 40 years, it would be expected that few would be paid-in-full. But 43 percent of beginning farmer loans had been paid-in-full by November 30, 2004.

Model Specification

The duration and agricultural loan studies reviewed earlier suggest a number of independent variables for the hazard function. The variable abbreviations and definitions are displayed in Table 2. Two categories of variables are defined. The first category defines core variables. These are loan-to-value (LTV), loan amount (LAMT), repayment capacity (REPAY) and debt-to-asset ratio (DA).[46] Jackson and Kaserman use LTV in both their equity and ability to pay models as do Hakim and Haddad in their analysis. Loan amount follows from an ability to pay theory as does repayment capacity. The debt-to-asset ratio is included given its past importance in agricultural lending and as a measure of borrower solvency.

The second category contains control variables. These include variables used in other studies that are not as strongly justified by theory but have shown some significance in related studies. Thirteen variables fall into this category. Four variables are demographically related and include borrower age (AGE), number of dependents (NUMDEP), whether the loan type is socially disadvantaged (SDA) or beginning farmer (BF). Nominal interest rate (NOMINT) is a loan characteristic and affects debt service. Being a first time borrower (FTIMEBO) indicates less familiarity by FSA with the borrower. Dixon et al. show that past involvement with FSA guaranteed loans is important in whether a borrower exits the FSA direct loan program. This strength of relationship is also measured by the number of OL (NUMOL), FO (NUMFO) and EM (NUMEM) active loans at time of origination. Past financial difficulties also characterize borrower creditworthiness. This variable is measured by whether the borrower had filed for

[46] The loan-to-value variable is defined differently than in the mortgage studies cited earlier. Because of the flexible uses of direct loans—particularly OL loans—and the fact that all direct loans are supposed to be fully collateralized, a broader definition is used. The definition in Table 2 is a measure of net equity available to pay off FSA liabilities.

bankruptcy or been in receivership (PASTBANK). Finally, net profit margin (NETPROFM), current ratio (CR) and total farm sales (TOTALSAL) are considered viable candidate variables in estimation because of the importance of measuring profitability, liquidity and scale of operation on financial viability.

Model Selection Strategy

Model specification proceeded in two steps. First the parametric distribution had to be selected. Then the independent variables for the models had to be chosen. Initially four competing risks models were estimated: two PH and two Weibull. One PH model was for OL loans of one-year maturity and then another PH model for seven-year maturity OL loans. An analogous pair of models was estimated assuming the Weibull distribution. Because of the different loan maturities it would naturally be expected that the coefficients of even identical variables would differ as a function of loan maturity as well as different variables being important as a function of loan type. The statistical significances of all coefficients were examined from the four models. Three variables—NOMINT, PASTBANK and FTIMEBO—were not significant at even the .1 level in any model and were discarded from each model for further consideration. Further eliminations were made by loan maturity. All variables that were not significant at .1 in both the PH and Weibull for a given loan maturity were dropped. This process led to the estimated models displayed in Tables 3 and 4. For the one-year loans this yielded eleven independent variables (not including a constant in the Weibull). The seven-year maturity models had ten independent variables. These reduced parameter models were re-estimated in both PH and Weibull forms.

The estimates of the reduced parameter models were compared between the PH and the Weibull. For the seven-year maturities the signs of the coefficients of a given variable did not vary between the PH and Weibull models. There was slightly less agreement in signs for the one-year loans. For variables significant in both the PH and Weibull models there were no sign differences. But some variables had conflicting signs if one or both variables were insignificant in the PH or Weibull models. All the variables significant in the seven-year maturity models were significant in both the PH and the Weibull. However, for the one-year maturities there were some variables that were significant in the Weibull but not significant in the PH and vice versa. But for any variable significant in both models, the signs were in agreement.

In what follows the PH model coefficients are discussed and used for analysis. The popularity of the PH and the lack of need to specify the base hazard rate are in its favor as well as ease of interpretation given the exponential form of the PH. But the general flow of the results would not change if the Weibull models were used instead.

The estimated coefficients, respective percentage changes in the hazard for a one unit change in the associated independent variable, and elasticities are presented in Tables 3 and 4. A positive coefficient indicates an increase in the independent variable increases the given hazard rate.

The coefficient signs are generally consistent with expectations. Several variables with significant coefficients at the .1 level or better increase the default hazard on one-year loans. These include NUMDEP and DA. More dependents imply a greater demand for household expenditures and an increased DA ratio implies lower solvency and therefore greater vulnerability. Total sales are also significant and positive. The positive sign is unexpected but might indicate greater risk taking by larger farms. Current involvement with FSA at origination indicates a lower default rate, perhaps indicating better loan making by FSA because of familiarity with the client.

Repayment ratio is negative and significant at 0.01. It also has the largest elasticity reflecting the sensitivity of defaults to the predicted ability to pay. But repayment ability is not significant

paid-in-full. Nor does repayment appear in either of the hazard functions for the seven-year loans.

The default indicators for the seven-year loans differ from the one-year loans. The DA ratio is still significant and positive as expected. Number of EM loans is significant and negative as are the number of OL and FO loans for the one-year loan default hazard. For the seven-year loans, loan-to-value is highly significant with the largest elasticity. This is consistent with past findings for mortgages. Moreover, being a beginning farmer is positive and significant. The percent partial parameter indicates that hazard of defaulting more than doubles for beginning farmers. Beginning farmers are high risk. As observed in Dixon et al., beginning farmers exit FSA direct loan programs at a higher rate than non beginning farmers.

The paid-in-full predictors differ in part from default predictors. The DA ratio is negative and significant for one-year loans but not significant for seven-year loans. Loan amount is significant and negative for both types of loans implying larger loans take longer to pay back. This may also reflect favorable interest rates to borrowers so that there is little incentive to pay loans early. Low interest rates may explain net profit margin having a negative coefficient for one-year loans. Strength of prior FSA relationship is negative for both loan types except for seven-year loans where number of existing FO loans has a positive sign.

The generally negative effect of number of existing FSA loans at origination for both default and paying back is somewhat perplexing. One implication is that loans to borrowers with existing loans generally have longer durations.[47] The difference in coefficient magnitudes for a given variable indicates the stronger effect for the variable. For example, for one-year loans the default coefficient of NUMOL is -0.17 compared with -0.032 for paid-in-full. Thus the net effect of one more outstanding OL loan at origination is to lower the default hazard more than the lowering of the paid-in-full hazard. This indicates an overall greater likelihood of paying the loan in fulldue too increasing NUMOL. This is even stronger for NUMFO for one-year loans given the positive sign on NUMFO for the paid-in-full hazard. For seven-year loans only NUMEM is significant for both default and paid-in-full. The net effect is that higher NUMEM decreases the default likelihood more than paying-in-full.

Higher profit margin farms may have higher internal rates of return and therefore are less likely to pay back one-year loans until due. Larger farms are less likely to pay back one-year loans early, perhaps indicating a greater reliance on farm income to pay back loans. More dependents also indicate less ability to pay one-year loans early.

Additional variables are significant in the hazard function for paying back seven- year loans. Not surprisingly, higher current ratios increase the likelihood of paying in full. Age has a similarly positive impact on paying in full. The two special borrower categories—beginning farmers and SDA farmers—are less likely to pay in full at a given point in time than other borrowers not in these classes. This also implies longer time to pay back. But the lack of SDA being significant in the default hazard implies pay back is longer for SDA loans but there is not a higher default likelihood. Dixon et al. found that having SDA or BF loans did not significantly increase the likelihood of borrowers exiting FSA program early and leaving farming involuntarily.

[47] The survivor function, S(t), for the PH is a baseline S(t) estimated with $\mathbf{x} = \mathbf{0}$. With $\mathbf{x}$ fixed over time the base line S(t) is exponentiated to $\exp(\boldsymbol{\beta}'\mathbf{x})$ (Kalbfleisch and Prentice, pg. 97.). So an increase in a component of $\mathbf{x}$ associated with a negative coefficient will increase the duration of a loan.

Hazard Ratios

In a competing risks model, deriving the net impact of a change in an independent variable common to both hazards requires examining how the ratio of hazards changes. This was explored earlier for the impact of the number of existing direct loans at time of origination. This analysis is extended to the four core variables. Figures 3-6 plot the ratio of the default hazard to the paid-in-full hazard as a function of these four variables, respectively. Figures 3-5 show that the default to paid-in-full hazard ratio is uniformly lower for one-year loans than seven-year loans. This is indicated by one-year loan graphs being below their respective seven-year loan graphs. All the graphs in the figures also demonstrate that the likelihood of paying in full is much greater than defaulting.

Figure 3 shows the importance of loan amount, particularly for seven-year loans. As the loan amount rises, the likelihood of default increases and it is much more steeply sloped for seven-year loans. While it would certainly be expected that default likelihood would increase for loan amount—particularly controlling for other variables—it is surprising that the increase is so steep for seven-year loans. So FSA borrowers have experienced more difficulty when loan amounts are high, ceteris paribus.

Figure 4 shows the hazard ratios as functions of the loan-to-value ratio. As with loan amount, seven-year loans are much more sensitive to loan-to-value than one-year loans. Loan-to value is measured in this study as the value FSA liabilities divided by farm assets net of non FSA liabilities. The seven-year loan graph indicates the high sensitivity of default relative to paying-in-full for borrowers who have inadequate net worth to collateralize loans. Interestingly, one-year loans show little sensitivity to this variable indicating the comparatively greater risk for seven-year loans.

The hazard ratios as a function of debt-to-asset are displayed in Figure 5. The graphs' slopes are modest compared with those in Figures 3 and 4. Given the general significance and elasticities of the debt-to-asset coefficients in Tables 3 and 4, this is somewhat surprising. But the graphs show that decreasing solvency leads to a greater chance of default indicating the risks involved in lending to less solvent borrowers.

Figure 6 shows the hazard ratio sensitivity as a function of the repayment ratio for one-year loans.[48] The severely negatively sloped curve demonstrates a high sensitivity to the outcome as a function of repayment. For example, an increase in REPAY from .75 to 1.0 reduces the hazard ratio by 38 percent. But the graph also indicates that as REPAY increases, higher levels of repayment ability are subject to diminishing impacts on the hazard ratio. Once a strong capacity is established to repay, further repayment capability is not likely to greatly improve default odds.

Conclusions

Competing risk, Cox proportional hazard models of FSA operating loans (OL) are estimated using data from a survey of all OL loans originated in federal fiscal years 1994-1996. The OL loans have maturities of one or seven years so that one set of models is estimated for each maturity.

The proportional hazards model allows for a more robust analysis than the more conventionally used binary choice models such as probit and logit. In addition to indicating the relative likelihood of default to paying a loan in full, the duration model enables estimating how long a

[48] A graph is not shown for seven-year loans. The REPAY variable was deleted in the seven-year model because it lacked significance in the first stage of model specification.

loan will perform until default or payback. Furthermore, observations on loans still performing can be utilized in estimation by treating them as censored observations as opposed to omitting them, or creating a nebulous third category or waiting until all loans have been terminated.

Our analysis shows that many independent variables significantly influence the hazards of default or paying-in-full. Measures of solvency, liquidity, repayment ability and profitability are important. These findings are similar to those from past agricultural credit worthiness analyses. Moreover, these variables are significant in predicting loan durations.

The results indicate the importance of the loan amount. In the estimated models the default risk relative to paying-in-full rises steeply as loan principal increases. This is especially true for seven-year loans. Since this is a ceteris paribus result, it raises the issue of why larger loans have considerably more risk. It may be that larger loans are made to those whose creditworthiness is diminished in ways not reflected in FSA approval methods. That the effect is so much less for one-year loans is puzzling since borrowers can average good and bad years over a seven-year loan. This phenomenon bears further investigation.

The loan-to-value (LTV) variable measures the effect of the ratio of FSA liabilities to assets net of non-FSA liabilities on default and payback capabilities.
While LTV variation is of minor importance for one-year loans, seven-year loans are highly sensitive. This suggests that FSA should perhaps add weight to the overall FSA debt burden on borrowers when making seven-year FSA loans relative to assets available to pay off FSA debts.

Strength of involvement with FSA is important. The results show that borrowers with currently active loans are generally less likely to default. The estimated models indicate that borrowers involved with FSA at loan origination likely default less and take longer to payback than those who have no active direct FSA loans at origination. Perhaps repeat borrowers are financially weaker and need numerous direct loans because of inability to get commercial loans. But the default rates are less because FSA has greater knowledge of the borrower effects not observable in the variables measured. This measure of borrower credibility deserves consideration in the loan approval process.

Table 1. Mean Duration Times by Loan Type and Loan Termination for FY 1994-1996 Sample

	OLREG 1yr	OLREG 7yr	OLBF 1yr	OLBF 7yr	OLBF SDA 1yr	OLBF SDA 7yrs	OLSDA 1yr	OLSDA 7yr	FOREG	FOBF	FOBFSDA	FOSDA	EM	Total
Defaulted (n)	53	38	14	20	4	15	15	14	2	4	2	2	28	211
Duration (Years)	4.87	6.00	4.76	4.85	6.78	4.76	4.63	4.99	5.70	5.59	4.53	5.92	4.98	68.36
	(0.31)	(0.37)	(0.60)	(0.50)	(1.60)	(0.82)	(0.82)	(0.85)	(1.60)	(1.13)	(2.26)	(2.26)	(0.44)	
Paid-in-full (n)	449	244	96	67	24	15	85	58	23	68	10	7	224	1370
Duration (Years)	1.18	5.48	1.07	5.79	1.14	6.40	1.06	5.28	7.42	6.49	7.18	7.67	5.37	61.53
	(0.08)	(0.11)	(0.18)	(0.21)	(0.5)	(0.59)	(0.25)	(0.31)	(0.36)	(0.21)	(0.67)	(0.92)	(0.12)	
Still Performing (n)	56	155	11	33	3	10	6	26	44	87	9	9	118	567
Duration (Years)	9.60	9.57	9.57	9.55	9.63	9.46	9.35	9.57	9.73	9.26	9.59	9.75	9.34	123.97
	(0.12)	(0.74)	(0.28)	(0.16)	(0.5)	(0.3)	(0.36)	(0.18)	(0.14)	(0.1)	(0.3)	(0.3)	(0.08)	

There are 567 observations considered as missing. Figures in parentheses are standard errors.

Table 2. Variable Definitions

Dependent Variable

DUR Number of years from origination to observed status (default, paid-in-full or still performing)

Independent Variables

Core Variables

LTV Loan-to-value measured as sum of loan amount and FSA farm liabilities divided by the sum of loan amount and farm assets minus the difference between farm liabilities and FSA farm liabilities.

LAMT Loan amount in thousands of dollars.

REPAY Income net of expenses divided by annual debt service.

DA Total borrower debts divided by total borrower assets.

Control Variables

AGE Operator age in years at time of application.

NUMDEP Number of dependents in household including borrower.

SDA Equals 1 if loan has socially disadvantaged assistance code, 0 otherwise.

BF Equals 1 if loan has beginning farmer assistance code, 0 otherwise.

NOMINT Nominal interest rate of loan at origination in percent.

FTIMEBO Equals 1 if borrower has never obtained an FSA direct or guaranteed loan, 0 otherwise.

NUMOL Number of active direct OL loans at time of loan application.

NUMFO Number of active direct FO loans at time of loan application.

NUMEM Number of active direct EM loans at time of loan application.

PASTBANK Equals 1 if borrower had been in receivership, discharged in bankruptcy, or petitioned for reorganization in bankruptcy prior to loan application.

NETPROFM Net farm income less living expense, this difference divided by total cash farm income.

CR Current liabilities divided by current assets.

TOTALSAL Total cash farm income in thousands of dollars.Table 3. Estimates of PH Coefficients and Their Partial Effects for the One-year Loans

Variable[1]	Coefficient Estimate	Percent Partial[2]	Elasticity[3]
NUMDEP_d	0.10152*	10.7	0.322
LAMT_d	0.00234	0.2	0.124
LTV_d	0.33345	39.6	0.290
NUMOL_d	-0.16920**	-15.6	-0.371
NUMFO_d	-0.68930***	-49.8	-0.375
NUMEM_d	0.06304	6.5	0.035
DA_d	0.42374***	52.8	0.323
CR_d	-0.10424	-9.9	-0.203
REPAY_d	-1.79501***	-83.4	-1.889
TOTALSAL_d	0.0004809*	0	0.075
NETPROFM_d	0.23216	26.1	0.009
CON_p	1.89778**	567.1	1.898
NUMDEP_p	-0.06427***	-6.2	-0.204
LAMT_p	-0.00416***	-0.4	-0.221
LTV_p	-0.00114	-0.1	-0.001
NUMOL_p	-0.03169**	-3.1	-0.069
NUMFO_p	0.07210*	7.5	0.039
NUMEM_p	0.02643	2.7	0.015
DA_p	-0.20573**	-18.6	-0.157
CR_p	0.00379	0.4	0.007
REPAY_p	0.02208	2.2	0.023
TOTALSAL_p	-0.000852**	-0.1	-0.133
NETPROFM_p	-0.27580**	-24.1	-0.011

[1] CON indicates a constant term for the paid-in-full model. The _d and _p indicate coefficients of variables in the default and paid-in-full hazards, respectively.

[2] Estimated percentage change in duration for a one unit change in the corresponding independent variable.

[3] Elasticities evaluated at the sample independent variable means.

Table 4. Estimates of PH Coefficients, Partial Effects and Elasticities for the Seven-year Loans

Variable[1]	Coefficient Estimate	Percent Partial[2]	Elasticity[3]
AGE_d	0.00554	0.6	0.226
SDA_d	0.04122	4.2	na
BF_d	0.83182***	129.7	na
LAMT_d	-0.00157	-0.2	-0.069
LTV_d	1.21191***	236	0.802
NUMOL_d	-0.09583	-9.1	-0.137
NUMFO_d	-0.19843	-18	-0.111
NUMEM_d	-0.49488*	-39	-0.167
DA_d	0.41586*	51.6	0.269
CR_d	-0.00451	-0.5	-0.011
CON_p	3.67540***	3846.4	na
AGE_p	0.00791*	0.8	0.323
SDA_p	-0.45856***	-36.8	na
BF_p	-0.26575*	-23.3	na
LAMT_p	-0.01056***	-1.1	-0.463
LTV_p	-0.36064***	-30.3	-0.239
NUMOL_p	-0.15394***	-14.3	-0.220
NUMFO_p	0.13722***	14.7	0.077
NUMEM_p	-0.26741***	-23.5	-0.090
DA_p	0.13219	14.1	0.086
CR_p	0.02066***	2.1	0.049

[1] CON indicates a constant term for the paid-in-full model. The _d and _p indicate coefficients of variables in the default and paid-in-full hazards, respectively.

[2] Estimated percentage change in duration for a one unit change in the corresponding independent variable.

[3] Elasticities evaluated at the sample independent variable means.

Figure 1. Unconditional Hazard Rates for One-Year Operating Loans

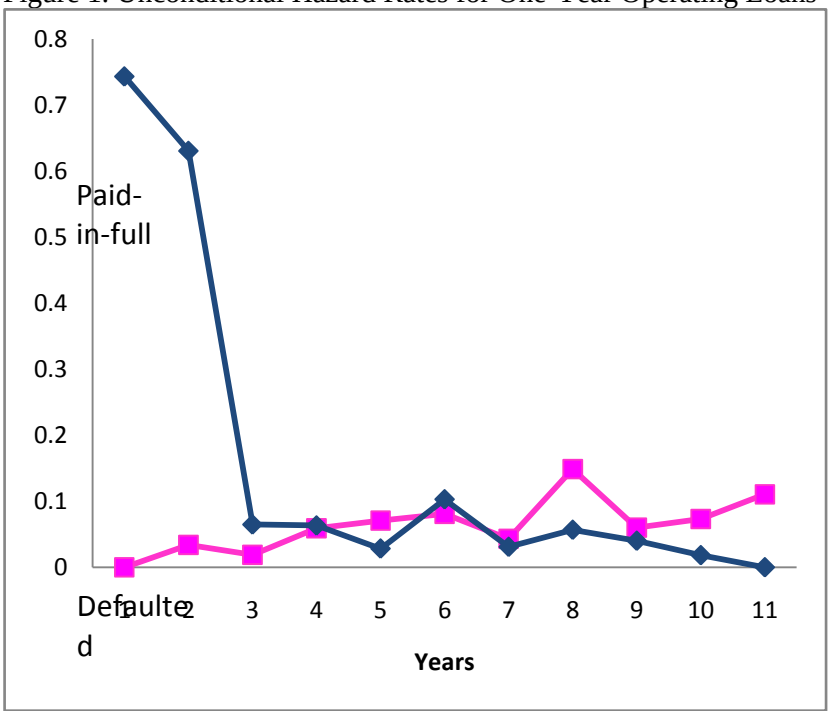


Figure 2. Unconditional Hazard Rates for Seven-Year Operating Loans

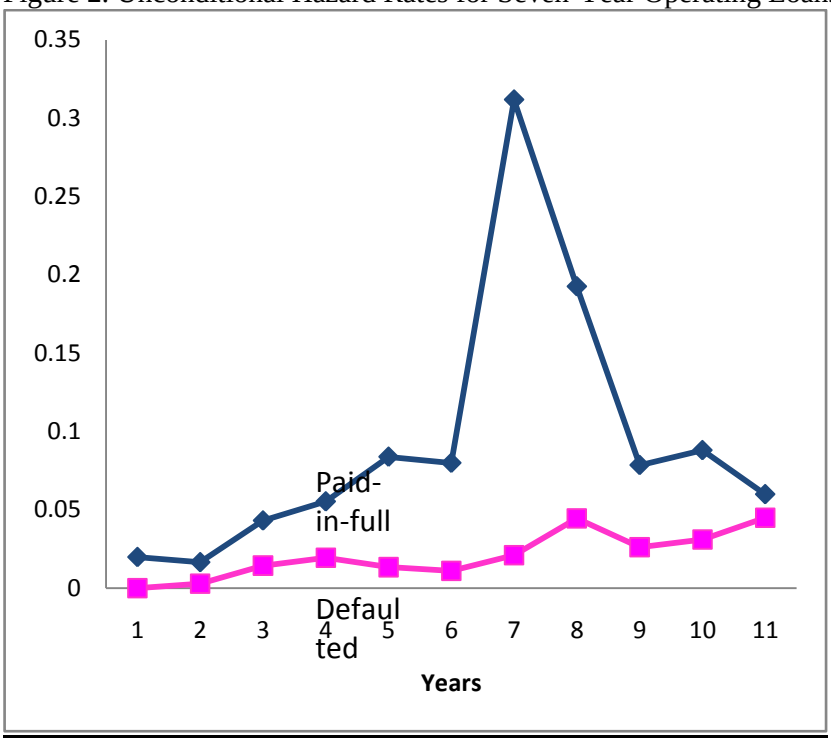


Figure 3. Ratio of Default Hazard to Paid-in-full Hazard as a Function of Loan Amount

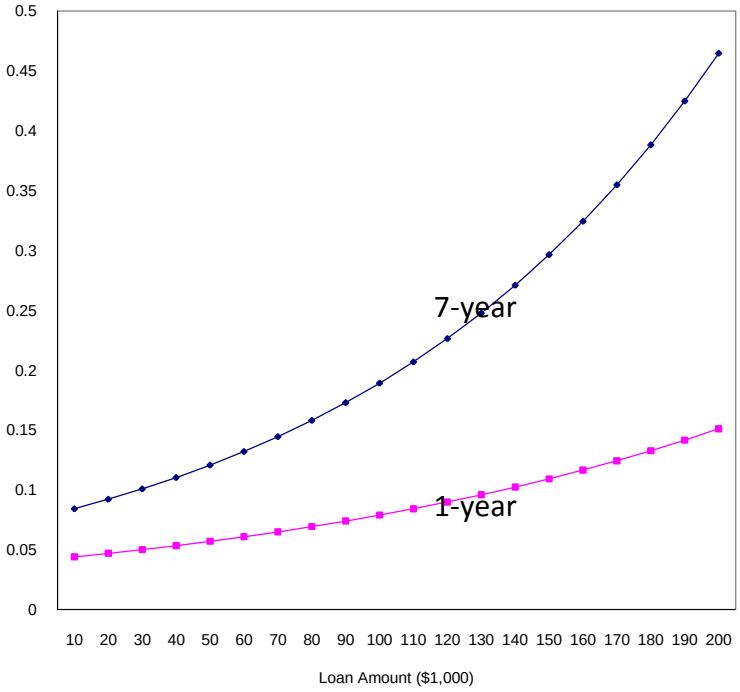


Figure 4. Ratio of Default Hazard to PIF Hazard as a Function of Loan-to-Value Ratio

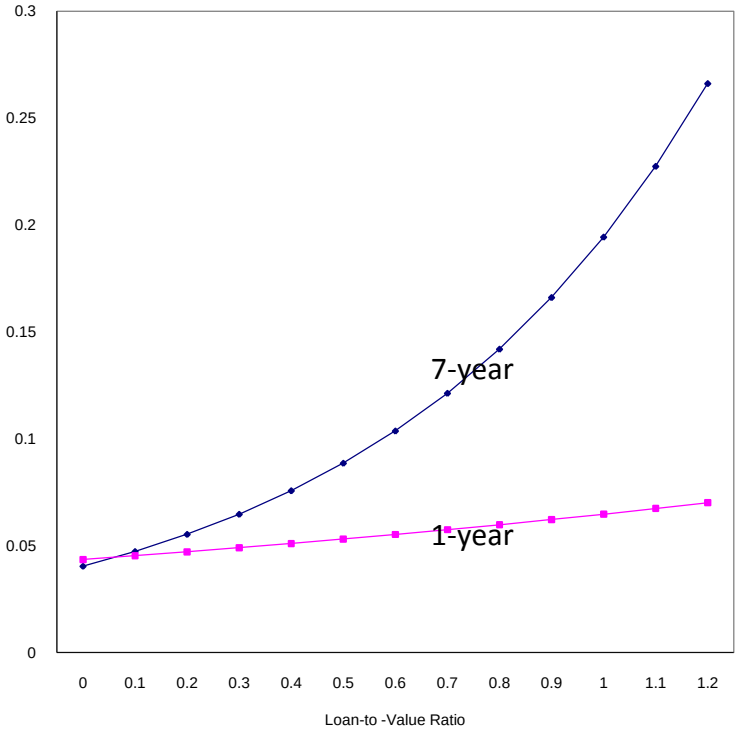


Figure 5. Ratio of Default Hazard to PIF Hazard as a Function of Debt-to-Asset Ratio

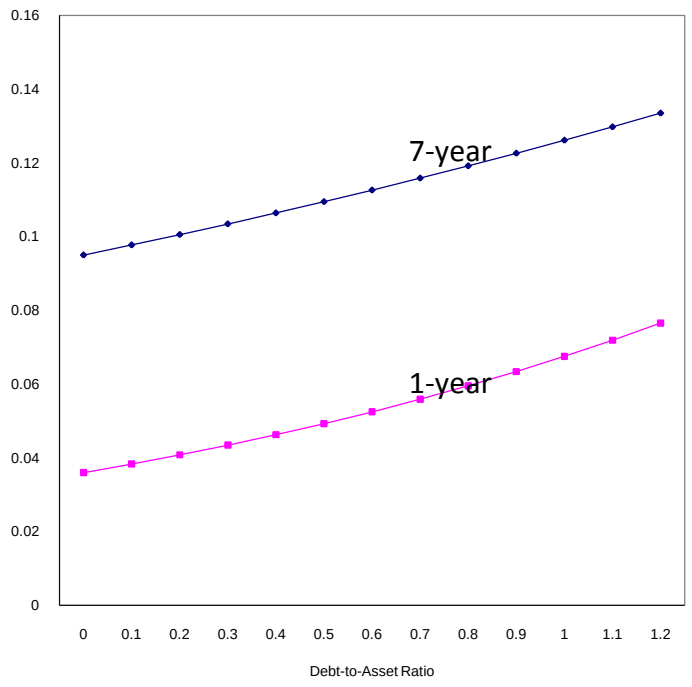
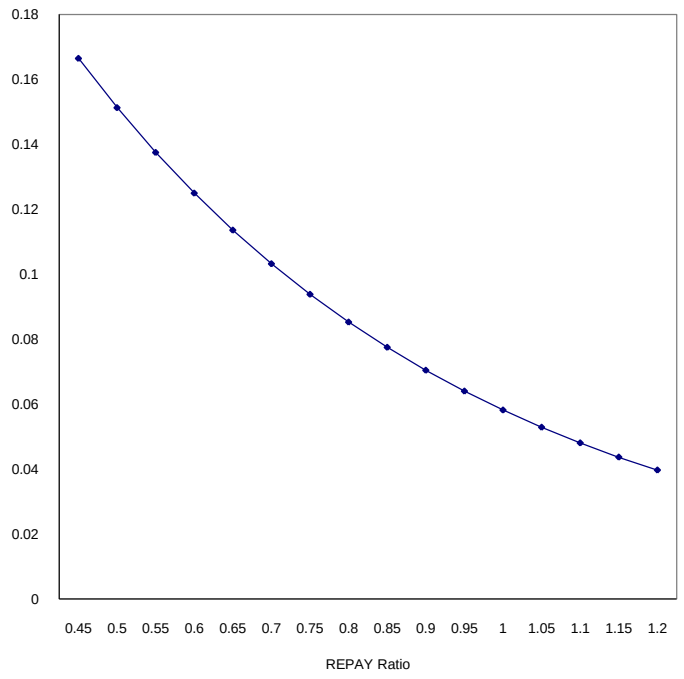


Figure 6. Ratio of Default Hazard to PIF Hazard as a Function of REPAY Ratio for One-Year Loans



References

Ambrose, B. W., and C. A. Capone. "The Hazard Rates of First and Second Defaults." *Journal of Real Estate Finance and Economics*. 20, 3(2000):275-293.

Cox, D. R. and D. Oakes. *Analysis of Survival Data*. Chapman and Hall, London,1984.

Dodson, C. B., and S. R. Koenig. "Farm Service Agency Credit Delivery to Different Classes of Borrowers." In C. G. Turvey. *Proceedings of the 46th Agricultural Finance Conference: The Changing Nature of Agricultural Risks*, University of Guelph, Ontario Canada, 2000.

Dixon, B. L., B. L. Ahrendsen, O. John Nwoha, S. J. Hamm, D. D. Danforth. "FSA Direct Farm Loan Program Graduation Rates and Reasons for Exiting." *Journal of Agricultural and Applied Economics*. In press.

Escalante, C. L., R. L. Brooks, J. E. Epperson and F. E. Stegelin. "Credit Risk Assessment and Racial Minority Lending at the Farm Service Agency." *Journal of Agricultural and Applied Economics.* 38, 1(April 2006):61-75.

Featherstone, A. M., C. A. Wilson, T. L. Kastens and JH. D. Jones. "Factors Affecting the Agricultural Loan Decision-Making Process." *Agricultural Finance Review*. 67, 1(2007):13-33.

Gallagher R. L. "Distinguishing Characteristics of Unsuccessful Versus Successful Agribusiness Loans." *Agricultural Finance Review*. 61, 1(2001):19-35.

Hakim, S. R., and M. Haddad. "Borrower Attributes and the Risk of Default of Conventional Mortgages." *Atlantic Economic Journal*. 27, 2(1999): 210-220.

Huffman, W. and T. Feridhanusetyawan. "Migration, Fixed Costs, and Location-Specific Amenities: A Hazard Analysis for a Panel of Males." *American Journal of Agricultural Economics*. 89, 2(2007):368-382.

Jackson, J. R., and D. L. Kaserman. "Default Risk on Home Mortgage Loans: A Test of Competing Hypotheses." *Journal of Risk and Insurance.* 47, 4(1980):678–90.

Kalbfleisch, J. D. and R. L. Prentice. *The Statistical Analysis of Failure Time Data*, second ed. Wiley-Interscience, Hoboken NJ, 2002.

Key, N. and M. J. Roberts. "Government Payments and Farm Business Survival." *American Journal of Agricultural Economics*. 88, 2(2006):382-392.

Lambrecht, B., W. Perraudin, and S . Satchell. "Time to default in the UK mortgage market." *Economic Modelling*. 14(1997):485-499.

Limsombunchai,V., C. Gan, and M. Lee. "An Analysis of Credit Scoring for Agricultural Loans in Thailand." *American Journal of Applied Sciences*. 2, 8(2005):1198-1205.

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

Nwoha, J., B. L. Ahrendsen, B. L. Dixon, E. C. Chavez, S. J. Hamm, D. M. Settlage, and D. Danforth. *Farm Service Agency Direct Farm Loan Program Effectiveness Study*. Fayetteville, Arkansas: Arkansas Agricultural Experiment Station, Division of Agriculture, University of Arkansas System Research Report #977,2005. At: http://www.uark.edu/depts/agripub/Publications/bulletins/977.pdf.

Splett, N. S., P. J. Barry, B. L. Dixon and P. N. Ellinger. "A Joint Experience and Statistical Approach to Credit Scoring." *Agricultural Finance Review*. 54(1994):39-54.

Turvey, C. G. and A. Weersink. "Credit Risk and the Demand for Agricultural Loans." *Canadian Journal of Agricultural Economics*. 45, 3(1997):201-217.

Turvey, C. G. "Credit Scoring for Agricultural Loans: A Review with Applications. *Agricultural Finance Review*. 51(1991):47-57.

U.S. Department of Agriculture, Economic Research Service. *Agricultural Income and Finance, Situation and Outlook Report*. AFO-48, February 1993.

U.S. Department of Agriculture, Farm Service Agency. "Direct Loan Program." Avalaible at: http://www.fsa.usda.gov/Internet/FSA_File/flp_direct_farm_loans.pdf. Accessed November 18, 2007.

U.S. Department of Agriculture, Farm Service Agency. "Farm Loans Program Funding." Available at: http://www.fsa.usda.gov/dafl/funding.htm. Accessed June 30, 2004a.

U.S. Department of Agriculture, Farm Service Agency. "Fact Sheet, February, 2004." Available at: http://www.fsa.usda.gov/pas/publications/facts/html/farmloaninfo02.htm. Accessed July 15, 2004 b.

U.S. Department of Agriculture, National Agricultural Statistics Service. *1997 Census of Agriculture United States Summary and State Data*. Washington, D.C., March 1999.