To the Graduate Council:

I am submitting herewith a thesis written by Morgan J. C. Scudi entitled "A Discriminant Analysis of Farmers Home Administration Housing Loans." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Agricultural Economics.

B R McManus

B. R. McManus, Major Professor

We have read this thesis
and recommend its acceptance:

Frank O Lentold

M. B. Badenhop

Accepted for the Council:

L Evans Clark

Vice Chancellor
Graduate Studies and Research

Ag-VetMed
Thesis
79
.S283
cop.2

A DISCRIMINANT ANALYSIS OF FARMERS HOME ADMINISTRATION HOUSING LOANS

A Thesis

Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Morgan J. C. Scudi

August 1979

1397219

ACKNOWLEDGMENTS

The Farmers Home Administration of the United States Department of Agriculture cooperated by providing basic information for this research and an hospitable climate in which the field work of this study was conducted.

The author wishes to thank Dr. B. R. McManus for serving as his Graduate Committee Chairman and for his assistance and counsel throughout the project. Thanks are also extended to Dr. Merton B. Badenhop and Dr. Franklin O. Leuthold for their service and aid on the Graduate Committee. Special thanks and appreciation are extended to the members of the author's Ad Hoc Committee, including, but not limited to, Dr. David W. Brown, Dr. Joe A. Martin, Dr. William L. Sanders, Dr. Charles B. Sappington, and Mr. Morgan Grey.

Finally, the author wishes to thank Nancy A. Scudi, his wife and confidante, for her support and understanding throughout this undertaking.

ABSTRACT

The objective of this study was the estimation of discriminant functions for the Section 502 Housing Loan Program of the Farmers Home Administration in Fentress, Grundy, and Haywood counties, Tennessee. The hypothesis of the study was that good and problem borrowers were distinguishable from one another based upon information available on the application form. Two samples were collected from the loan portfolios of each county. One sample represented problem borrowers and the other good borrowers. Data from each application sampled included the socio-economic characteristics of the borrower. The samples were analyzed in three steps: first, in a tabular analysis; second, in stepwise regressions; and third, in discriminant functions.

The tabular analysis revealed that within each county several of the borrower characteristics were related to the loan classification but that the relationships were not consistent among the counties. The stepwise regressions were used to further study the variables. The most predictive variables for each county were then used in the three discriminant models. The models differed reflecting the local employment, creditor, and household types, as well as the financial measures which could be best identified with loan classification in each county.

The discriminant models for Fentress, Grundy, and Haywood counties were tested, found significant, and correctly classified 81, 79, and 81% of the respective loan application samples.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	INTRODUCTION	1
	A. The Problem	1
	B. Counties Under Study	3
	C. Previous Research	6
	D. Objectives	8
II.	PROCEDURE	10
	A. Stepwise Regression	12
	B. Discriminant Analysis	14
III.	ANALYSIS	16
	A. Good Versus Problem Loan Applications	17
	B. Regression Analysis	27
	C. The Discriminant Functions	34
	D. Limitations to the Study	42
IV.	SUMMARY	48
LIST OF REFERENCES		50
APPENDIX		53
VITA		56

LIST OF TABLES

TABLE		PAGE
I.	Borrowers Delinquent by Number and Percentage of FmHA Individual Housing Programs in the United States as of June 30, 1969-1974	2
II.	Rural Housing Units Constructed Under FmHA Section 502 Program in Fentress, Grundy, and Haywood Counties, 1971-1976	4
III.	Characteristics of Housing, Population, Income, and Education for Tennessee and for Fentress, Grundy, and Haywood Counties	5
IV.	Mean Values from Good and Problem Loan Applications, By County, of Selected Sociological Characteristics for the Study Period 1967-1976	18
V.	Mean Values from Good and Problem Loan Applications, By County, of Selected Employment Types and the Number of Jobs and Months of Unemployment in the Previous Five Years for the Study Period 1967-1976	21
VI.	Mean Values from Good and Problem Loan Applications, By County, of Selected Payments and Creditors and of Total Debts and Total Property Owned for the Study Period 1967-1976	23
VII.	Mean Values from Good and Problem Loan Applications, By County, of Selected Income and Expense Measures for the Study Period 1967-1976	25
VIII.	Fentress County Primary Regression Model	28
IX.	Fentress County Secondary Regression Model	30
X.	Grundy County Primary Regression Model	32
XI.	Grundy County Secondary Regression Model	33
XII.	Haywood County Regression Model	35

LIST OF FIGURES

FIGURE		PAGE
1.	Approximate distribution of active FmHA borrowers in Fentress, Grundy, and Haywood counties	44
2.	Distribution of good and problem FmHA borrower samples as classified by the discriminant function	45
3.	Distributions of borrower applicants with and without those rejected for loans	47

CHAPTER I

INTRODUCTION

A. THE PROBLEM

The Section 502 Rural Housing Loan Program of the Farmers Home Administration (FmHA) has grown rapidly in recent years, serving the important function of increasing the availability of decent rural housing (USDA, 1975). Unfortunately, as the program has increased, so too has the problem of delinquencies. As Table I shows, the delinquency rate for the United States grew from 3% in 1969 to 8% in 1974.

A delinquency occurs when a monthly mortgage payment is missed; however, servicing and/or collection procedures normally begin when a borrower is two or three payments delinquent (USDA, 1975). The servicing guidelines of Tennessee's FmHA (Tennessee Bulletin No. 1310, 1976) state that ". . . staffs must . . . become . . . an active participant in a well planned and implemented servicing program that causes borrowers to keep their accounts current and thus retain ownership of their homes. . . ." Clearly, the FmHA goals include continued homeownership and an active role for county supervisors in keeping their borrowers current. Therefore, we can infer that as the delinquencies rise in any given county, the amount of time that county staff must spend in planning and implementing a servicing program will also rise, which must therefore reduce the time available for performing other important services.

TABLE I

BORROWERS DELINQUENT BY NUMBER AND PERCENTAGE OF FmHA INDIVIDUAL HOUSING PROGRAMS IN THE UNITED STATES AS OF JUNE 30, 1969-1974

	Borrowers Delinquent	
Year	Number	Percent
1960	7,472	3
1970	9,353	3
1971	15,016	4
1972	20,313	4
1973	27,252	5
1974	47,820	8

Source: Brief History of the Farmers Home Administration, January 1975, U.S. Printing Office: 1975--667785/48 Region No. 6.

One of the primary goals of the FmHA is to provide "loans to low and moderate income families [to] help them obtain modest homes in suitable rural environments at prices and on terms they can afford" (USDA, 1972). In keeping with this goal, eligibility was increased in 1976 by raising the maximum adjusted family income from $8,500 to $10,000. Further, housing loans are tailored to low-income families by subsidies which reduce the interest rate to as low as 1%.

The FmHA, a government agency, operates its housing program without profit motivation; however, its housing loans must be repaid. The servicing of loans, by expending the time and resources of the local and regional FmHA offices, is an expense that adds to the overall cost of the housing program. Improved methods of screening borrowers in order to reduce the number of delinquencies have the potential of aiding the FmHA provide rural home financing.

B. COUNTIES UNDER STUDY

The Fentress, Grundy, and Haywood county offices were under different FmHA administrative districts and had active FmHA housing programs for low- and moderate-income rural families.

Over the six fiscal year periods of 1971-1976, the Fentress County FmHA office made 496 new housing loans, while Grundy and Haywood counties made 424 and 605 new housing loans, respectively (Table II). Assuming the housing stock of 1970 (Table III) as a base, the new houses built with these loans represent an increase in the housing stock of 14% in Fentress County, 13% in Grundy County, and 11% in Haywood County.

TABLE II

RURAL HOUSING UNITS CONSTRUCTED UNDER FmHA SECTION 502 PROGRAM
IN FENTRESS, GRUNDY, AND HAYWOOD COUNTIES, 1971-1976

Fiscal Year	Fentress County		Grundy County		Haywood County	
	Number	Dollar Value	Number	Dollar Value	Number	Dollar Value
1971	52	612,000	83	667,000	96	1,190,000
1972	79	1,076,000	106	1,064,000	140	1,821,000
1973	80	1,107,000	76	823,000	116	1,654,000
1974	81	1,297,000	37	543,000	61	985,000
1975	101	2,123,500	38	589,000	74	1,354,000
1976	103	1,857,000	84	1,323,000	118	2,275,000
TOTALS	496	7,672,500	424	5,009,000	605	9,279,000

Source: FmHA Tennessee, Summary of Accomplishments, 1971-1976.

TABLE III

CHARACTERISTICS OF HOUSING, POPULATION, INCOME, AND EDUCATION FOR TENNESSEE AND FOR FENTRESS, GRUNDY, AND HAYWOOD COUNTIES

Characteristic	Tennessee	Fentress County	Grundy County	Haywood County
Housing--1970				
Number of housing units	1,212,961	3,638	3,175	5,549
House median dollar value	12,755	6,191	4,806	10,805
Units lacking some plumbing (%)	13.0	47.3	37.2	37.0
Units built prior to 1950 (%)	47.5	53.8	55.7	61.8
Housing in one-family units (%)	79.9	93.9	91.2	90.9
Homeowner vacancy rate (%)	1.3	1.1	0.3	0.9
Population--1970				
White	3,297,378	12,587	10,577	8,873
Black	620,636	*	*	10,723
Population per square mile	95	25	30	38
Population change 1960-1970 (%)	10.0	-5.2	-7.7	-16.2
Income--1969				
Median dollar family income	7,447	3,937	4,183	4,474
Number of families	100,692	3,164	2,857	4,634
Number of familes below poverty level	18,326	1,343	1,148	1,879
Families below poverty level (%)	18.2	42.4	40.2	40.5
Education--1970				
Persons over 25 years of age with less than five years' education (%)	9.5	22.3	16.9	19.5
Persons with four years of high school or more (%)	41.8	18.8	18.6	25.3

*None reported.

Source: Tennessee Statistical Abstract, 1977, Center for Business and Economic Research, College of Business Administration, The University of Tennessee, Knoxville, July 1977.

Table III shows that in 1969 all three counties had low-median family incomes, with over 40% of their families below the poverty level, and that they each experienced declines in population from 1960-1970. In 1970 the median housing value for each county was low and well below the state average. Over one-third of all housing in Grundy and Haywood counties and almost half in Fentress County lacked some plumbing, while over 90% of all housing was single family. All three counties had low homeowner vacancy rates, and more than half the housing was built before 1950. In short, Fentress, Grundy, and Haywood counties had large proportions of their populations living in substandard housing, earning low-median family incomes, and having low levels of educational achievement.

C. PREVIOUS RESEARCH

Several studies have been made using methods of discriminant analysis to estimate functions useful in the credit screening of agricultural loan applicants. The idea is put forth that there are key characteristics of applicants which can be used to predict loan success, and that the identification of these characteristics is possible by studying the records of past borrowers. Therefore, data are collected from the applications of loans that are classified as good or successful and from loans that are classified as problem or unsuccessful. The data from the two groups are compared for differences related to the loan classifications by means of stepwise regression or stepwise discriminant analysis. The key characteristics of borrowers are in this way determined, and a final discriminant function is calculated.

In the Bauer and Jordon study (1971) of Production Credit Association (PCA) in Tennessee, eight borrower characteristics were found to be important and were used in the discriminant function. They were the debt to asset ratio, farm value, total liabilities, marital status, family expense as a percent of total expenses, current asset to current liability ratio, number of dependents, and expected income as a percent of the previous year's income. The discriminant function correctly classified 85% of the sample loans included in the study.

Johnson and Hagan (1973), based on PCA loans in Missouri, found six important borrower characteristics that related to loan classification. They were the debt to asset ratio, current asset to current liability ratio, a repayment index, debt to net worth ratio, gross income to current debts ratio, and net worth to PCA commitment ratio.

The study by Dunn and Frey (1976), using Illinois PCA loans, found four important borrower characteristics in its discriminant analysis. They were the total liabilities to total assets ratio, amount of credit life insurance, amount of the note divided by cash farm income, and number of acres owned. The discriminant function correctly classified 75% of the sample loans tested.

The results of these studies show that the financial position of the borrower as determined by ratios using debts, liabilities, and their relationships to income and net worth make up most of the important borrower characteristics. In fact, of the characteristics identified as important, only marital status and the number of dependents are not wealth related.

The loans studied in this analysis differ from those used in the previous studies because of two major distinctions. First, housing loans differ from farm loans because housing is not income producing and, therefore, cannot be analyzed in the type of feasibility studies available with various farm enterprises. Also, housing loans, unlike farm loans, are not in any way repaid out of income produced from the effective use or investment of the borrowed funds, but are repaid from income allocated to meet living expenses. Second, unlike the PCA agricultural loan program, the FmHA housing loan program has income restrictions which eliminate many potentially safe loans and only allows lending money to those borrowers who are unable to secure home financing elsewhere.

D. OBJECTIVES

The objective of this study was to develop discriminant models which would differentiate between good and problem borrowers currently enrolled in the FmHA Section 502 Rural Housing Program in three rural Tennessee counties. This objective was accomplished through three steps:

1. A brief descriptive analysis of the socioeconomic characteristics of good and problem borrowers in Fentress, Grundy, and Haywood counties.
2. A stepwise regression to select variables which were statistically related to the loan classification of good and problem borrowers for each county from those variables which had a priori theoretical relationships

to loan repayment and which were available from the data on the FmHA loan applications.

3. The derivation and testing of a separate discriminant model for each of the three counties, built upon the variables selected by the stepwise regression.

This study was an attempt at the discrimination of good from problem borrowers in the FmHA housing loan program in Tennessee. The discriminant model building was based on the premise that only good and problem borrowers currently in the program need be sampled (Borrowers who were neither clearly good or problem were excluded from the sampling.). If these two groups of borrowers could be discriminated based upon information available at the time of their loan application, then quantifiable predictions of loan success could be made for new applicants at the time of application.

CHAPTER II

PROCEDURE

The procedure for this analysis was built around information available on the FmHA loan applications, It was assumed that the applications were designed to collect information that was useful in the loan evaluation process. For this reason, all of the socioeconomic information available on the application was collected.

Two samples were made from each of the loan portfolios in Fentress, Grundy, and Haywood counties. One sample of approximately 50 applications represented good loans, and the other of a like number represented problem loans. Good loans were defined for the purpose of this study as loan accounts which were not delinquent, i.e., always paid when due, over a two-year period, 1974-1976. Problem loans were defined as loan accounts which were found delinquent by three or more months at any time during the same two-year period. The three-month delinquency was chosen to define problem loans because a State Administrative Directive (Tennessee Bulletin No. 1310, 1975, p. 3) called for active loan servicing on the part of the county supervisor when a borrower was two or three payments (months) delinquent on his or her mortgage.

The data collected from the loan applications were evaluated in terms of their relationship to the good and problem loan classifications. The available information was categorized under four broad headings: sociological, employment, payments and debts, and income and expenses.

The information which was classified under the heading of sociological characteristics included the applicant's age, years of education, race, months in residence, and military experience. Also included under the same heading were the total household number and family type, which was divided into four dummy variables: male-headed with spouse present, male-headed without spouse present, female-headed with spouse present, and female-headed without spouse present. These sociological characteristics were included in the analyses because they were expected to reflect part of the borrower's attitudinal values towards home ownership and debt repayment. As a result, the group of good borrowers might differ from problem borrowers in the values for these variables, which in turn might be useful in discriminating between the groups.

The employment characteristics collected for each sample borrower included the number of jobs in the last five years for both spouses as present in the household, the number of months of unemployment in the last five yeras for both spouses, and the employment type at the time of application for both spouses. The employment type was broken down into 12 dummy variables which included: service, sales, farm, housewife, unemployed, craftsman, machine operator, retired, laborer, and miscellaneous. These employment variables were used in the study because home mortgage loans are repaid out of income, which in turn is dependent upon employment and its reliability in generating a minimum level of disposable income.

The characteristics included under payments and debts that were used in this analysis included the borrower's monthly payments, total debts, total assets, and insurance. Monthly payments were divided into

those made on real estate and those made on personal property. Three discrete variables were used to measure present creditors: banks, credit cards, and small loan companies. All of these variables were included in the analysis because they are used to measure financial position, debt experiences, and repayment capacity.

The information which was classified under income and expense characteristics included income, planned income, expenses, and planned expenses. Income and planned income were divided into that of the male spouse, the female spouse, and the total household. The expense and planned expense figures were for the entire household. This income and expense information is used by the FmHA in determining borrower eligibility, interest credit, and lending limits, but was of importance in this study as a measure of financial liquidity and discretionary spending.

A. STEPWISE REGRESSION

The identification of borrower characteristics used in the evaluation of loan applications in the discriminant analysis was accomplished by a stepwise multiple regression. In general terms, the multiple regression model is considered to be:

$$y_t = b_o + b_1 x_{1t} + b_2 x_{2t} + \ldots + b_k x_{kt} + u_t, \; t = 1, \ldots, n$$

where:

y_t = the t^{th} observation on the dependent variable

x_{it} = the t^{th} observation on the i^{th} independent variable (or regressor), $i = 1, 2, \ldots k$ if we have k independent variables

u_t = the t^{th} value of the disturbance term

b_i = the coefficient of the i^{th} independent variable

(b_o represents the constant term)

and is estimated in the form:

$$\hat{y}_t = \hat{b}_o + \hat{b}_1 x_{1t} + \ldots + \hat{b}_k x_{kt}$$

and the estimator of the disturbance term is:

$$\begin{aligned}\hat{u}_t &= y_t - \hat{y}_t \\ &= y_t - \hat{b}_o - \hat{b}_1 x_{1t} - \ldots - \hat{b}_k x_{kt}\end{aligned}$$

Thus,

$$\begin{aligned}y_t &= \hat{y}_t + \hat{u}_t \\ y_t &= \hat{b}_o + b_1 x_{1t} + \ldots + \hat{b}_k x_{kt} + \hat{u}_t\end{aligned}$$

(Kelejian and Oates, 1976, pp. 121-124).

The dependent variable ($\hat{y}_t$) in this analysis was the loan classification which was either good or problem and was, therefore, a zero-one dependent variable. This factor alone is important, in that regression techniques predict a continuous linear equation, while the sample dependent variable ($\hat{y}_t$) is discontinuous and results in an increased disturbance term and, among other things, a lower R-square than would otherwise be expected.

The independent variables (x_{kt}) are borrower characteristics taken from the loan applications. The same independent variables were considered in all three counties. All calculations for the stepwise

regression were made using the SAS 76 maximum R-square improvement technique (Barr, 1976, pp. 256-256) at the University of Tennessee Computer Center.

B. DISCRIMINANT ANALYSIS

The variables or borrower characteristics identified in the regression analyses were then used to develop a mathematical model which separates the classes of good from problem loans by using a two-group discriminant analysis. The discriminant function, which resembles and is closely related to the regression model (Tatsuoka, 1972, pp. 170-173), is made up of the same dependent variable set equal to the independent variables and their coefficients, called the discriminant function coefficients. These coefficients form the basis for the mathematical model, in that any loan application can be classified by calculating the sum of the applications' important characteristics times their respective discriminant function coefficients. The calculated value is then compared to the cutoff number of the respective county; ones with larger values are likely to be problem loans, and ones with smaller values are likely to be good loans.

Mathematically, the discriminant function is the linear equation:

$$y = ax_a + bx_b + \ldots kx_k$$

The mean values of the first group, good loans, become

$$\bar{y}_1 = a\bar{x}_{a1} + b\bar{x}_{b1} + \ldots k\bar{x}_{k1}$$

and those for the second group, problem loans, are then

$$\bar{y}_2 = a\bar{x}_{a2} + b\bar{x}_{b2} + \ldots k\bar{x}_{k2}$$

and the difference between the means is

$$\bar{y}_1 - \bar{y}_2 = D;\ \bar{x}_{a1} - \bar{x}_{a2} = d_a;\ \ldots\ \bar{x}_{k1} - \bar{x}_{k2} = d_k$$

so that the discriminant function is

$$D = ad_a + bd_b + \ldots kd_k$$

(Tippett, pp. 294-299). This function is computed and has the desirable characteristic of being based not only on the characteristics of the two groups, but also on the distance between their mean values. All calculations were made by using the discriminant analysis technique of SAS 76 (Barr, 1976) and of BMD (Dixon, 1963) and computed by the University of Tennessee Computer Center.

CHAPTER III

ANALYSIS

The analysis is divided into three parts. The first part, based on the sample mean values, considers differences between the good and problem loan samples, both within each county and among the three counties. This first step is important because it helps identify possible relationships between the variables and the loan classification, and, more importantly, it shows that many of these relationships differ in direction in one or more of the counties.

The second part of the analysis reports the results of two regression models calculatd for each of the three counties by using the maximum R-square regression technique. The primary model was constructed by using only those variables were were: 1) believed to be related to the loan classification; and 2) consistently reported on the applications, i.e., had a low number of missing values in the samples. The second model was constructed by using those variables which were believed to be related to loan classifications but which had a low number of sample observations. The regression technique eliminated observations with missing values for any of the independent variables, making inclusion of these variables in the primary model impossible.

The discriminant analysis makes up the third part of this chapter. The discriminant function for each county was constructed by using the 10 variables selected in the regression analysis as having the strongest explanatory powers in predicting the loan classification. Each county

model was then tested for homogeneity of their within covariance matrices, and a pooled covariance matrix was used in the calculation of functions. The predictive strength of each model was tested by classifying the sample application data and comparing the results with the actual classifications.

A. GOOD VERSUS PROBLEM LOAN APPLICATIONS

The mean values of the sample variables easily characterize the two classes of borrowers and add insight into the borrower characteristics used in the regression and discriminant analyses that follow.

All mean values were calculated directly from the information contained in the sample application forms. The values of continuous variables such as age, incomes, and expenses were accepted as accurate. In the cases of actual values for property, incomes, or debts, blank spaces on the sample applications were considered zeroes; but in the cases of variables such as age, months in residence, and planned income and expense figures, blank spaces were considered missing data. The values of dummy variables, such as employment types and creditors, were coded based on a determination by the researcher as to the appropriate categories under which the data were to be included.

Sociological Characteristics

Good and problem borrowers can be analyzed in terms of several sociological characteristics. Table IV displays the mean values of these and shows that on the average good borrowers were older and had fewer household members than did problem borrowers in all three counties.

TABLE IV

MEAN VALUES FROM GOOD AND PROBLEM LOAN APPLICATIONS, BY COUNTY, OF SELECTED SOCIOLOGICAL CHARACTERISTICS FOR THE STUDY PERIOD 1967-1976

	Fentress County				Grundy County				Haywood County			
	Problem		Good		Problem		Good		Problem		Good	
Variables	Mean	(N)[a]	Mean	(N)	Mean	(N)	Mean	(N)	Mean	(N)	Mean	(N)
Applicant Information												
Age	30.42	(57)	37.58	(50)	30.68	(54)	37.47	(51)	33.98	(67)	41.65	(55)
Education	10.18	(56)	9.32	(47)	8.98	(54)	9.02	(50)	9.45	(64)	9.04	(52)
Race[b]	1.00	(57)	1.00	(50)	1.00	(54)	1.00	(51)	0.25	(67)	0.42	(53)
Veteran	0.23	(57)	0.20	(50)	0.20	(54)	0.25	(51)	0.12	(65)	0.14	(55)
Months in residence	15.93	(45)	30.22	(32)	20.53	(38)	14.31	(32)	26.09	(56)	33.02	(46)
Household Number	3.67	(57)	2.29	(50)	3.81	(54)	3.33	(51)	4.46	(67)	3.51	(55)
Household Type[c]		(57)		(50)		(54)		(51)		(67)		(55)
Male-headed												
With spouse	0.79		0.72		0.81		0.72		0.70		0.65	
Without spouse	0.09		0.08		0.04		0.00		0.03		0.05	
Female-headed												
With spouse	0.00		0.06		0.02		0.02		0.00		0.04	
Without spouse	0.12		0.14		0.13		0.25		0.27		0.25	

[a](N) denotes number of observations.

[b]Discrete variable where 1 denoted Caucasian, 0 denoted Negro.

[c]Mean values of the dummy variables represent the fraction of the sample within that classification.

Years of education, on the other hand, do not appear to show strong or consistent relationships to loan repayment classifications. The variable months in residence appears to have a strong relationship with loan classifications within each county but was inconsistent among counties. The fact that problem borrowers have spent more months in residence than good borrowers have in Grundy County is perhaps due to unique features within the county.

Household type, race, and military experience were measured as zero-one variables, so that the mean values in Table IV can be converted to percents (multiply by 100) to aid in their interpretation. Further, household type is broken down into four dummy variables, so that for any sample the sum of the four household categories should equal one (100%). All households with both spouses present were classified as male-headed unless the female was the sole income earner for the household, in which case a female-headed classification was given.

Three observations about household types can be made about the sample means:

1. The traditional family type, male-headed with spouse present, was the most common in all samples.
2. The second most common household type was female-headed without the spouse present.
3. The good loan samples had consistently lower percentages of male-headed households with the spouse present than did the problem loan samples.

Employment Characteristics

On good loan applications, both men and women from all three counties held, on the average, fewer jobs and spent fewer months unemployed over the five years prior to the application than did those on problem loan applications (Table V). While this relationship appears strong and consistent among the counties, it suffers in sample size (N) due to missing values. Job classifications for the applicant and spouse reflect only the jobs at the time of the application and were classified by means of dummy variables. The most frequent job classifications are shown in Table V; they are not consistent among counties but do appear, in some cases, to be related to loan classifications. This relationship may reflect strengths of particular local employers or other unique county situations.

Payment and Debt Characteristics

Monthly payment figures from the samples (Table VI) appear to be related to loan classification, as shown by large differences between the group means, but are inconsistent among counties. Likewise, the mean values for total debts, total property owned, and present creditors are also inconsistent because the problem loan sample means are not larger than the good loan sample means for all counties. The sample means indicate that in Fentress and Grundy counties, good loan applicants on the average owned more property, owed more in debts, and had higher monthly payments on real estate owned than problem loan applicants; the opposite was true in Haywood County. In Fentress and Haywood counties, good loan applicants were more likely to owe money to banks

TABLE V

MEAN VALUES FROM GOOD AND PROBLEM LOAN APPLICATIONS, BY COUNTY, OF SELECTED EMPLOYMENT TYPES AND THE NUMBER OF JOBS AND MONTHS OF UNEMPLOYMENT IN THE PREVIOUS FIVE YEARS FOR THE STUDY PERIOD 1967-1976

	Fentress County				Grundy County				Haywood County			
	Problem		Good		Problem		Good		Problem		Good	
Variables	Mean	(N)[a]	Mean	(N)	Mean	(N)	Mean	(N)	Mean	(N)	Mean	(N)
Number of Jobs												
Male	2.46	(43)	1.79	(38)	2.51	(41)	2.23	(41)	1.87	(47)	1.60	(35)
Female	0.86	(43)	0.70	(37)	1.17	(47)	1.11	(44)	1.54	(54)	1.41	(41)
Months of Unemployment												
Male	2.26	(39)	0.36	(25)	1.55	(20)	1.27	(15)	0.54	(22)	0.16	(25)
Female	0.31	(39)	0.00	(28)	0.50	(24)	0.00	(11)	2.12	(34)	0.80	(30)
Male Employment Type[b]		(57)		(50)		(54)		(51)		(67)		(55)
Self-employed	0.04		0.08		0.00		0.00		0.00		0.02	
Machine operator	0.21		0.20		0.26		0.20		0.28		0.18	
Laborer	0.09		0.18		0.31		0.20		0.13		0.11	
Unemployed	0.04		0.00		0.02		0.00		0.01		0.00	
Service	0.18		0.12		0.13		0.02		0.12		0.07	
Professional	0.02		0.04		0.00		0.04		0.01		0.04	
Craftsman	0.14		0.10		0.11		0.16		0.07		0.05	
Sales	0.10		0.02		0.00		0.00		0.03		0.02	
All other types	0.18		0.26		0.17		0.38		0.35		0.51	

TABLE V (Continued)

Variables	Fentress County				Grundy County				Haywood County			
	Problem		Good		Problem		Good		Problem		Good	
	Mean	(N)	Mean	(N)	Mean	(N)	Mean	(N)	Mean	(N)	Mean	(N)
Female Employment Type[b]												
Retired	0.00		0.04		0.02		0.18		0.01		0.07	
Laborer	0.02		0.00		0.00		0.02		0.09		0.05	
Service	0.14		0.10		0.07		0.06		0.11		0.25	
Housewife	0.56		0.50		0.74		0.59		0.31		0.33	
All other types	0.28		0.36		0.17		0.15		0.48		0.30	

[a](N) denotes number of observations.

[b]Mean values of the dummy variables represent the fraction of the sample with that classification.

TABLE VI

MEAN VALUES FROM GOOD AND PROBLEM LOAN APPLICATIONS, BY COUNTY, OF SELECTED PAYMENTS AND CREDITORS AND OF TOTAL DEBTS AND TOTAL PROPERTY OWNED FOR THE STUDY PERIOD 1967-1976

	Fentress County				Grundy County				Haywood County			
	Problem		Good		Problem		Good		Problem		Good	
Variables	Mean	(N)[a]	Mean	(N)	Mean	(N)	Mean	(N)	Mean	(N)	Mean	(N)
Monthly Payments												
Real estate ($)	3.04	(57)	7.53	(49)	3.96	(54)	6.27	(51)	2.83	(65)	2.37	(54)
Personal property ($)	72.02	(57)	73.86	(50)	76.04	(53)	65.37	(51)	84.06	(66)	62.29	(55)
Creditors[b]		(57)		(50)		(54)		(51)		(67)		(55)
Banks	0.51		0.62		0.70		0.51		0.31		0.42	
Credit cards	0.19		0.14		0.15		0.14		0.24		0.09	
Loan companies	0.35		0.04		0.15		0.18		0.31		0.16	
Total Debts ($)	1,842.82		2,390.94		1.764.65		1.870.76		1,335.59		1,237.69	
Total Property Owned ($)	4,851.84		6,299.92		3,813.24		4,501.72		3,198.43		2,743.65	

[a](N) denotes number of observations.

[b]Mean values of the discrete variables represent the fraction of the sample with that classification.

and less likely to owe money to small loan companies than were problem loan applicants; the opposite was true in Grundy County. Further, in Grundy and Haywood counties good loan applicants had smaller average monthly payments on personal property than did the problem loan applicants; the opposite of this was true in Fentress County.

Income and Expense Characteristics

The mean values of the income and expense figures for the samples (Table VII) were calculated from the figures in the applications. Since all the samples were collected over the same period from applications made during the same years, the mean values are comparable among the counties and between the classifications. However, the individual observations of income and expense for different points in time are not comparable due to inflation and were, therefore, put into ratios for use in the regression and discriminant analyses.

Only three consistent relationships exist for the sample mean values of income and expenses as related to loan classification for all three counties. These are that male incomes, planned household incomes, and planned household expenses are all larger on problem loan applications than they are on good ones. Perhaps problem borrowers are more dependent upon male incomes and have greater expectations for the future relative to the past than do good borrowers; but in view of the other income variables (Table VII), these three relationships are not enough in themselves to make strong statements about the relationship of income and expenses to loan classification for all three counties.

In Fentress County, problem loan applications had higher average incomes and expenses in all but one of the eight categories listed in

TABLE VII

MEAN VALUES FROM GOOD AND PROBLEM LOAN APPLICATIONS, BY COUNTY, OF SELECTED INCOME AND EXPENSE MEASURES FOR THE STUDY PERIOD 1967-1976

	Fentress County				Grundy County				Haywood County			
	Problem		Good		Problem		Good		Problem		Good	
Variables ($)	Mean	(N)*	Mean	(N)	Mean	(N)	Mean	(N)	Mean	(N)	Mean	(N)
Income												
Male	4,312.40	(57)	3,738.26	(50)	4,518.67	(54)	4,063.14	(51)	2,924.27	(67)	1,975.09	(55)
Female	1,360.23	(57)	1,452.58	(50)	700.72	(54)	886.98	(51)	1,255.86	(67)	976.98	(55)
Household	6,060.71	(56)	5,920.00	(50)	5,581.48	(54)	6,013.72	(51)	5,442.59	(54)	5,459.46	(37)
Planned Income												
Male	5,374.00	(50)	4,363.64	(44)	5,483.33	(48)	6,330.30	(33)	5,057.78	(45)	4,407.89	(38)
Female	1,652.73	(55)	1,606.67	(45)	1,251.61	(31)	1,179.17	(24)	2,706.81	(44)	2,116.33	(49)
Household	6,726.32	(57)	6,188.00	(50)	6,040.74	(54)	5,858.82	(51)	6,501.59	(63)	6,429.09	(55)
Household Expenses	4,624.68	(57)	3,802.22	(50)	5,568.52	(54)	5,784.33	(51)	2,622.75	(67)	2,258.65	(55)
Planned Expenses	5,739.22	(51)	5,276.92	(39)	6,076.47	(51)	5,700.83	(48)	5,594.44	(54)	5,211.11	(45)

*(N) denotes number of observations.

Table VII. The exception was female income, which may have been affected by the large number of female-headed households in the good classification for Fentress County (Table IV).

The samples in Grundy County reveal principle differences between the incomes and expenses actually earned and paid and those planned. Problem loan applicants had, on the average, higher male incomes but lower female incomes, household incomes, and household expenses than did good loan applicants. However, the planned figures show exactly the reverse relationships by loan classification.

The Haywood County samples reveal that, as in Fentress County, problem loan applicants had higher average incomes and expenses in all but one of the income and expense categories (Table VII) than did the good loan sample. The exception in Haywood County was that total household incomes were, on the average, slightly higher for the good loan sample.

The fact that no clear-cut relationships appear between income or expenses and loan classification may be explained by the income limitations of the loan program itself. Borrowers in the sample were limited to maximum adjusted family incomes of $8,500 (USDA, 1972) and minimum incomes that allowed for loan repayment over 33 years at 1%. Further, only borrowers unable to obtain financing elsewhere were eligible for the program. The result appears to be that borrowers, both good and problem, are in the same relative financial positions and that local conditions within the counties influence the significance of these borrower characteristics.

B. REGRESSION ANALYSIS*

The purpose of the regression analysis was to identify the important borrower characteristics that relate to loan classifications in order to use them in the generation of discriminant models for each county. These calculations were possible because borrower characteristics differ between the good and problem loan samples for each county.

The regression models consider loan classification as the dependent variable and regress the applicant's age, employment, household type, borrowing characteristics, expense-income ratios, and other socioeconomic characteristics. The significance level for all models and variables is 10%.

Fentress County

The primary regression model for Fentress County (Table VIII) uses three continuous and seven discrete variables to explain the sample's loan classification. The three continuous variables are age (X_1),

---

*A note of caution is needed in viewing the R-square results of the regression analysis. Because all calculations were done by using the ordinary least squares method, the zero-one dependent variable, loan classification, causes problems with heteroscedasticity and a positively biased disturbance term. The effect is a lowering of the R-square (For a full discussion, see Kmenta, 1971, pp. 425-428.). In other words, a regression equation predicted by means of the ordinary least squares method predicts values for the dependent variable outside the range of zero-one. What is important here is that this problem is basic and accepted in the use of a discriminant function and that a low R-square in the regression, if it only reflects inefficiencies in the technique, will not diminish its discriminating capacity. Note here also that the three primary model R-squares are 0.39, 0.32, and 0.51, but the same discriminant functions correctly classified 81%, 79%, and 81%. To further emphasize this point, a heteroscedastic regression was run using the Fentress County primary model and resulted in an R-square of 0.58 versus the OLS 0.39, but it discriminated the samples no better than did this study's discriminant function.

TABLE VIII

FENTRESS COUNTY PRIMARY REGRESSION MODEL

Variable		B Value	F	Prob > F
Intercept		0.445		
X_1	(age of applicant)	0.012	13.92	0.0003
X_2	(debts on credit cards)	-0.193	2.82	0.0963
X_3	(debts to loan companies)	-0.537	27.74	0.0001
X_4	(female-headed household with spouse present)	0.505	4.04	0.0473
X_5	(self-employed)	0.258	2.03	0.1578
X_6	(machine operator)	0.173	2.67	0.1074
X_7	$(\frac{\text{planned household income}}{\text{total household income}})$	-0.404	7.61	0.0070
X_8	$(\frac{\text{total monthly payments - subsidy}}{\text{total household income}})$	4.017	2.26	0.1357
X_9	(laborer)	0.296	5.53	0.0208
X_{10}	(unemployed)	-0.649	4.51	0.0363

Note: R-square = 0.393; F = 6.09; Prob > F = 0.0001.

planned household income divided by total household income (X_7), and total monthly payments less subsidy divided by total household income (X_8). The seven discrete variables represent only three borrower characteristics because two dummy variables are for source of debts, credit card debt (X_2), and small loan company debt (X_3), and four dummy variables are for male employment, self-employed (X_5), machine operator (X_6), laborer (X_7), and unemployed (X_{10}). The last characteristic, household type, is represented by the lone dummy variable female-headed household with male spouse present (X_4).

Seven of the ten variables are significant at the 10% level. Of these, four are negatively related to good loan classifications, which means that credit card debts, small loan debts, large planned household income ratios, and unemployed male workers are associated with bad loan classifications (Y = 0). The remaining three variables--age, female-headed households, and male laborers--are positively related to good loan classifications.

The Fentress County secondary regression model (Table IX) is made up of three variables which could not be used in the primary model due to missing values. Months in residence (X_2) is significant and positively related to good loan classification. Months of male unemployment (X_3) is also significant at the 10% level but is negatively related to a good classification.

Both the primary and secondary regression models for Fentress County are significant at the 10% level. The secondary model points out that both months in residence and months of male unemployment, which were calculated from the employment history of the last five years, are two variables which may have improved the primary model.

TABLE IX

FENTRESS COUNTY SECONDARY REGRESSION MODEL

Variable	B Value	F	Prob > F
Intercept	0.214		
X_1 (monthly rent)	0.002	0.71	0.4049
X_2 (months in residence)	0.005	4.04	0.0517
X_3 (months of unemployment)	-0.054	3.71	0.0619

Note: R-square = 0.168; F = 2.49; Prob > F = 0.0750.

Grundy County

The primary regression model for Grundy County (Table X) is made up of three continuous and seven discrete variables. The continuous variables are female income divided by male income (X_7), total monthly payments less subsidy divided by household income (X_2), and X_2 divided by household number (X_{10}). The discrete variables, all zero-one dummy variables, are small loan company debts (X_1); male-headed households without spouse present (X_3); two female job types, female retired (X_4) and female laborer (X_8); and three male job types, service (X_5), professional (X_6), and draftsmen (X_9). All are significant at the 10% level except female laborer (X_8).

Of the significant variables, three are negatively related to good loan classifications--male-headed households without spouse, male service employment, and large total payments as a proportion of income. The remaining six variables are positively related to good loan classification.

The secondary regression model (Table XI) shows that although there are severe problems caused by the small sample, months in residence (X_2) may be a variable that, if reported, could improve the primary model for Grundy County.

The primary regression model is significant at the 10% level, but the secondary model is not.

Haywood County

The primary regression model for Haywood County is made up of four continuous and six discrete variables. Age (X_1), household number (X_4), amount of life insurance (X_{10}), and mortgage payment less subsidy divided

TABLE X

GRUNDY COUNTY PRIMARY REGRESSION MODEL

Variable		B Value	F	Prob > F
Intercept		0.571		
X_1	(debts to loan companies)	0.238	3.71	0.0573
X_2	$(\frac{\text{total monthly payments - subsidy}}{\text{total household income}})$	-14.219	11.38	0.0011
X_3	(male-headed household without spouse present)	-0.563	3.13	0.0800
X_4	(female retired)	0.424	7.43	0.0077
X_5	(service employment)	-0.311	3.57	0.0619
X_6	(professional)	0.650	4.23	0.0425
X_7	$(\frac{\text{female income}}{\text{male income}})$	0.0001	6.09	0.0154
X_8	(female laborer)	0.560	1.61	0.2073
X_9	(craftsman)	0.220	2.77	0.0994
X_{10}	$(\frac{X_2}{\text{household number}})$	19.615	5.54	0.0206

Note: R-square = 0.32; F = 4.38; Prob > F = 0.0001.

TABLE XI

GRUNDY COUNTY SECONDARY REGRESSION MODEL

Variable	B Value	F	Prob > F
Intercept	0.513		
X_1 (monthly rent)	-0.004	1.50	0.2388
X_2 (months in residence)	-0.007	3.71	0.0733

Note: R-square = 0.24; F = 2.37; Prob > F = 0.1278

by household income (X_7) are the continuous variables of which household number and life insurance are negatively related to good loan classifications. The discrete variables, male-headed households without spouse present (X_2), male professional employment (X_3), and males sales employment (X_9) are also negatively related to the good classification, while bank debt (X_5), female service employment (X_6), and housewife (X_8) are positively related (Table XII).

All of the ten variables are significant at the 10% level, as well as the regression equation as a whole. No significant secondary variables or regression equations were found for Haywood County.

C. THE DISCRIMINANT FUNCTIONS

The discriminant functions for Fentress, Grundy, and Haywood counties are designed to meet the objective of this study by forming mathematical models which separate good from problem loan applications. The variables used in these three models were selected by the previous regression analysis and are the same ones used in the primary regression models. The discriminant functions do lack the constant values found in the regression models because the discriminant functions are relative measures designed only to sort the observations into two groups, good or problem.

In theory, the coefficients of the primary regression models and those of the corresponding discriminant model are proportional (Tatsuoka, 1971, pp. 170-177). The calculated coefficients are, however, only roughly proportional for two reasons:

TABLE XII

HAYWOOD COUNTY REGRESSION MODEL

Variable		B Value	F	Prob > F
Intercept		-0.333		
X_1	(age of applicant)	0.015	22.65	0.0001
X_2	(male-headed household without spouse present)	-0.766	7.33	0.0085
X_3	(professional)	-1.410	6.34	0.0141
X_4	(household number)	-0.054	7.20	0.0091
X_5	(debts to banks)	0.362	15.81	0.0002
X_6	(female service employment)	0.318	8.14	0.0057
X_7	$(\frac{\text{monthly mortgage payment - subsidy}}{\text{total household income}})$	13.065	10.64	0.0017
X_8	(housewife)	0.241	6.32	0.0142
X_9	(sales)	-0.568	4.23	0.0434
X_{10}	(life insurance)	-0.00001	4.52	0.0371

Note: R-square = 0.515; F = 7.57; Prob > F = 0.0001.

1. The continuous variables were carried for eight decimal places in the discriminant analysis but for 16 places in the calculated regression equations.
2. In Haywood County, the discriminant analysis included three more observations than the regression analysis did, and therefore the degrees of freedom and sums of squares are different.

Fentress County

The discriminant function calculated for Fentress County is:

$$\hat{Y} = -0.00077X_1 + 0.01214X_2 + 0.03382X_3 - 0.03186X_4 - 0.01629X_5 - 0.0109X_6 + 0.0255X_7 - 0.25322X_8 - 0.01865X_9 + 0.04093X_{10}$$

where:

$\hat{Y}$ = the loan classification value

X_1 = the age of the applicant

X_2 = debt on credit cards (0 or 1 value)

X_3 = debt to loan companies (0 or 1 value)

X_4 = female-headed household with spouse present (0 or 1 value)

X_5 = self-employed male (0 or 1 value)

X_6 = machine operator male (0 or 1 value)

$X_7 = \frac{\text{planned household income}}{\text{total household income}}$

$X_8 = \frac{\text{total monthly payments - subsidy}}{\text{total household income}}$

X_9 = laborer, male (0 or 1 value)

X_{10} = unemployed male (0 or 1 value)

The variables associated with a good loan classification have negative discriminant coefficients because the calculated mean dependent value, $\bar{Y}$, for the good loan sample of 49 observations is -0.01456, $\bar{Y}_G$, and it is 0.01021, $\bar{Y}_P$, for the problem loan sample of 56 observations.

A cutoff value, $\bar{Y}_C$, was estimated for the Fentress County discriminant function by using the formula:

$$\bar{Y}_C = \frac{\sigma_G \bar{Y}_P + \sigma_P \bar{Y}_G}{\sigma_G + \sigma_P}$$

where:

σ_G = standard deviation of the good loan sample = 0.01444

σ_P = standard deviation of the problem loan sample = 0.01638

$\bar{Y}_G$ = the mean calculated dependent value for the good loan sample = -0.01456

$\bar{Y}_P$ = the mean calculated dependent value for the problem loan sample = 0.01021

By substituting the appropriate values, the equation becomes:

$$\bar{Y}_C = \frac{(0.01444)\ (0.01021) + (0.01638)\ (-0.01456)}{(0.01444) + (0.01638)} = 0.00295458$$

The cutoff value is the dividing point between good and problem loan classifications, so that applications can be evaluated by calculating $\hat{Y}$ and comparing it to 0.00295458, $\bar{Y}_C$. Applications with a calculated $\hat{Y}$ larger than the $\bar{Y}_C$ are classified as likely problem loans, and those smaller as likely good loans.

An F-test was performed on the Fentress County discriminant functions, and the value F(10,94) = 6.08502 was calculated and found significant at the specified 10% level.

The discriminant function was then used to classify the Fentress County good and problem loan samples, and it correctly classified 82% of the problem loan applications and 80% of the good loan applications.

Grundy County

The discriminant function based on the loan samples from Grundy County is:

$$\hat{Y} = -0.01345X_1 + 0.80494X_2 + 0.031816X_3 = 0.0240X_4 + 0.0176X_5 - 0.03682X_6 - 0.00001X_7 - 0.3169X_8 - 0.01245X_9 - 1.11041X_{10}$$

where:

$\hat{Y}$ = the loan classification value

X_1 = debt to loan companies (0 or 1 value)

$X_2 = \frac{\text{total monthly payments - subsidy}}{\text{total household income}}$

X_3 = male-headed households without spouse present (0 or 1 value)

X_4 = female retired (0 or 1 value)

X_5 = male service employed (0 or 1 value)

X_6 = male professional (0 or 1 value)

$X_7 = \frac{\text{female income}}{\text{male income}}$

X_8 = female laborer (0 or 1 value)

X_9 = male craftsman (0 or 1 value)

$$X_{10} = \frac{X_2}{\text{household number}}$$

The calculated mean dependent value from the problem loan sample, $\overline{Y}_P$, of 53 observations is 0.01348, while $\overline{Y}_G$ from the good sample of 51 observations is -0.00466. Therefore, variables with negative discriminant function coefficients are associated with good loan classifications, and variables with positive discriminant function coefficients are associated with problem loan classifications.

The Grundy County discriminant model cutoff value, $\overline{Y}_C$, was estimated by using the formula:

$$\overline{Y}_C = \frac{\sigma_G \overline{Y}_P + \sigma_P \overline{Y}_G}{\sigma_G + \sigma_P}$$

where:

σ_G = standard deviation of the good loan sample = 0.01522

σ_P = standard deviation of the problem loan sample = 0.01122

$\overline{Y}_G$ = the mean calculated dependent value for the good loan sample = -0.00466

$\overline{Y}_P$ = the mean calculated dependent value for the problem loan sample = 0.01348

By substituting the appropriate values, the equation becomes:

$$\overline{Y}_C = \frac{(0.01522)\ (0.01348) + (0.01122)\ (-0.00466)}{(0.01522) + (0.01122)} = 0.00295458$$

This cutoff value divides good from problem loan classifications, so that Grundy County loan applications can be evaluated by calculating the $\hat{Y}$ for the discriminant function and comparing the value to the cutoff. Applications with smaller values are classified as likely good loans, and those with larger values are classified as likely problem loans.

An F-test was performed on the discriminant function, and the value F(10,93) = 4.38317 was calculated and found significant at the specified 10% level.

The discriminant function was then used to classify the Grundy County good and problem loan samples, and it correctly classified 79% of the problem loan and 78% of the good loan applications.

Haywood County

The Haywood County discriminant function is:

$$\hat{Y} = -0.00135X_1 + 0.05491X_2 + 0.00626X_3 + 0.00529X_4 - 0.02444X_5 - 0.03128X_6 - 0.53771X_7 - 0.02078X_8 + 0.04719X_9$$

where:

$\hat{Y}$ = the loan classification value

X_1 = the age of the applicant

X_2 = male-headed household without spouse present (0 or 1 value)

X_3 = male professional (0 or 1 value)

X_4 = household number

X_5 = debt to banks (0 or 1 value)

X_6 = female service employment (0 or 1 value)

$X_7 = \frac{\text{monthly mortgage payment - subsidy}}{\text{total household income}}$

X_8 = housewife (0 or 1 value)

X_9 = male sales employment (0 or 1 value)

The tenth variable, amount of life insurance, dropped out of the function because its coefficient was less than 0.00001.

The calculated mean dependent value, $\overline{Y}_P$, from the problem loan sample is -0.03625, and for the good loan sample, $\overline{Y}_G$ is -0.07836. Therefore, variables with negative discriminant function coefficients are associated with good loan applications, and ones with positive coefficients are associated with problem loan applications.

The cutoff value, $\overline{Y}_C$, for the Haywood County model was estimated by using the formula:

$$\overline{Y}_C = \frac{\sigma_G \overline{Y}_P + \sigma_P \overline{Y}_G}{\sigma_G + \sigma_P}$$

where:

σ_G = standard deviation of the good loan sample = 0.01861

σ_P = standard deviation of the problem loan sample = 0.02500

$\overline{Y}_G$ = the mean calculated dependent value for the good loan sample = -0.07836

$\overline{Y}_P$ = the mean calculated dependent value for the problem loan sample = -0.03625

and by substituting the appropriate values, the equation equals:

$$\bar{Y}_C = \frac{(0.01861)\ (-0.03625) + (0.025)\ (-0.07836)}{(0.01861) + (0.025)} = -0.06039011$$

Applications can be classified as likely problem loans if their calculated $\hat{Y}$'s are larger than $\bar{Y}_C$ and as likely good loans if smaller.

An F-test was performed in the Haywood County discriminant function, and the value F(10,74) = 6.46799 was calculated and found significant at the specified 10% level.

The discriminant function was then used to classify the Haywood County loan samples, and it correctly classified 84% of the problem loan and 78% of the good loan applications.

D. LIMITATIONS TO THE STUDY

The two group discriminant functions of this study are sorting equations by which a loan application can be categorized as either likely problem or likely good. The advantage to the Farmers Home Administration of such predictions, if accurate, would be a decline in delinquencies and increased efficiency for their home mortgage program. However, several problems stand in the way of the application of these functions to mortgage applicants in Fentress, Grundy, and Haywood counties.

Approximately 70% of the active loans in Fentress, Grundy, and Haywood counties were not classified either good or problem because they were found to have been delinquent, but never delinquent by three months or more. If these borrowers who were, in actuality, neither good nor problem were to be categorized by the discriminant functions of this study, they would have to be classified as good or problem.

Graphically, the problem may be explained by looking at a normal distribution of borrowers (Figure 1) centered around the average number of months delinquent. Good and problem borrowers make up the two tails of the distribution. From these two extreme groups, samples were taken, and a discriminant function was estimated (Figure 2). A cutoff value was estimated to separate the two groups for classification purposes. Assuming a normal distribution of all borrowers and an equal number of good and problem borrowers with the same variance, the calculated cutoff value for a homogeneous discriminant function with equal prior probabilities will divide the total population of borrowers (Figure 1) exactly in half.

The cutoff value can, however, be shifted by means of adjustments in the prior probabilities. As the cutoff ($\overline{Y}_C$) is shifted to the left (Figure 2), the number of good borrowers misclassified as problem will decrease, but, consequently, the number of problem borrowers misclassified as good will increase. This problem is analogous to the shifting probabilities of Type I and Type II errors in hypothesis testing. The shifting of the cutoff to the left will affect the classification of the total population of borrowers (Figure 1) by increasing the number classified as good and decreasing the number classified as problem. Conversely, a shift of the cutoff to the right would have the opposite effect.

The question of where to set the cutoff has not been answered by this study. Rather, the cutoff used was calculated with prior probabilities proportional to the sample sizes, so that the discriminant function could be tested for its accuracy in discriminating between the two

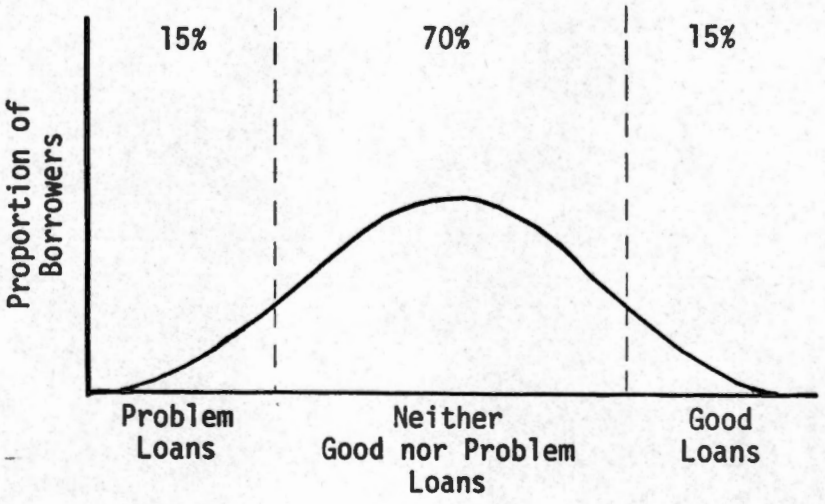


Figure 1. Approximate distribution of active FmHA borrowers in Fentress, Grundy, and Haywood counties.

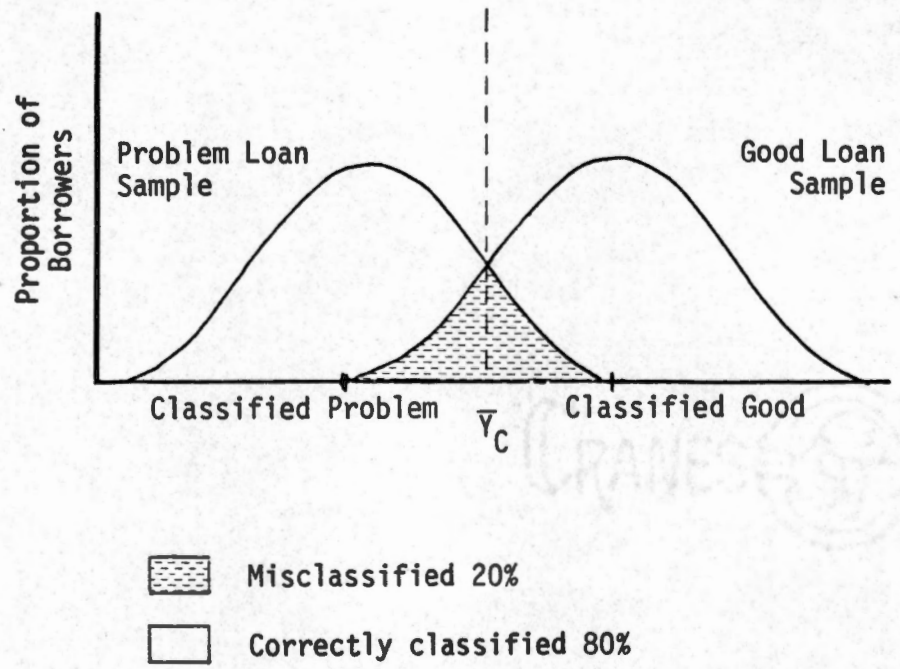


Figure 2. Distribution of good and problem FmHA borrower samples as classified by the discriminant function.

sample groups. Additional information should be gathered before any attempt is made to arrive at the prior probabilities appropriate for calculating a cutoff value for all loan applications.

The areas that need additional study include the rejected applicants because they may shift the total population distribution of all FmHA mortgage applicants to the left (Figure 3). The rejected applicants are a portion of the initial applicant population and, as such, need to be considered in setting the prior probabilities of borrowers being good or problem. The current borrowers that are in between the good and problem loan samples should also be studied before any assumptions about their characteristics are made, and, finally, a policy decision should be made as to whether it is more desirable to avoid problem loans at the expense of some good loans, or to include good loans at the cost of additional problem loans.

The discriminant functions for Fentress, Grundy, and Haywood counties are homogeneous and significant, but appropriate prior possibilities must be assigned before a cutoff value is used in the evaluation of new borrower applications.

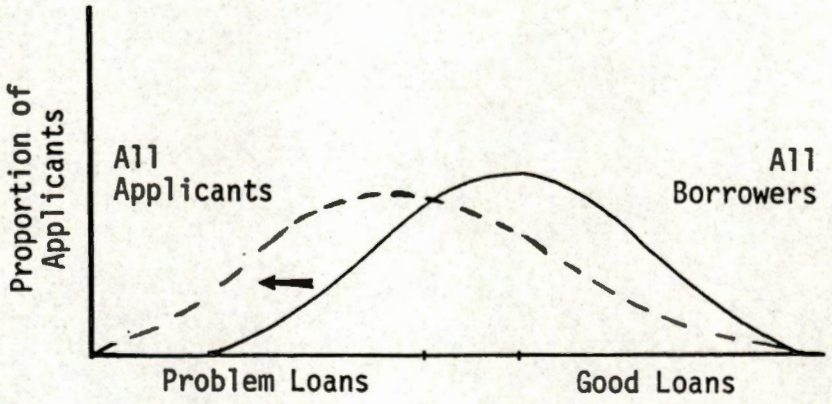


Figure 3. Distributions of borrower applicants with and without those rejected for loans.

CHAPTER IV

SUMMARY

The thesis of this study was that good and problem borrowers in the FmHA rural housing loan program are distinguishable from each other based upon combinations of their socioeconomic characteristics. Samples of good and problem loan applications were collected from the files of three FmHA county offices, and the characteristics of good and problem borrowers were compared and analyzed. Significant differences were found between the data for good versus problem borrowers for each of the three counties, but the differences were not consistent among the counties. A discriminant function was derived for each county based on nine borrower characteristics, four of which were further broken down into several dummy variables. The discriminant functions calculated for each county are unique, reflecting the particular employment types, household types, and creditor types which were found, through a stepwise regression, to be related to loan classifications in those counties. The mathematical models were tested, found significant, and correctly classified approximately 80% of the sample applications.

This study has gone beyond previous research using the discriminant analysis technique to analyze loan classifications for agricultural borrowers by taking three separate samples within the loan program. Each sample in this study was restricted to a rural county, and thus geographic, administrative, and community boundaries were not crossed. The tabular analysis and derived discriminant models indicate that a broad state-wide discriminant model faces severe problems in handling

peculiarities that local economic communities produce. In other words, a socioeconomic characteristic which might be positively linked with repayment capacity in one community is often negatively related in another.

This finding is not inconsistent with previous research, although neglected, because the earlier studies sampled across geographic and administrative boundaries and then aggregated the data for analysis. The significant borrower characteristics of these studies were found, generally, to be income or wealth related. A similar finding might well have been made in this study if, indeed, the samples had been aggregate, and if the loan program studies had not been geared to low and lower middle income borrowers.

LIST OF REFERENCES

LIST OF REFERENCES

Annual Report Tennessee 1973, Farmers Home Administration, Nashville, 1974.

Barr, Anthony J., et al. A User's Guide to SAS 76, Raleigh: SAS Institute, Inc., 1976.

Bauer, Larry L. and John P. Jordan. A Statistical Technique for Classifying Loan Applications, Knoxville: University of Tennessee Agricultural Experiment Station, 1971.

Brief History of the Farmers Home Administration, January, 1975, USDA, Washington: U.S. Printing Office, 1975.

Dixon, W. J. BMD Biomedical Computer Programs, Berkeley: The University of California Press, 1968.

Dunn, Daniel J. and Thomas L. Frey. "Discriminant Analysis of Loans for Cash-Grain Farms," Agricultural Finance Review, 30:60-66, 1976.

Farmers Home Administration Tennessee 1971 Annual Report, Farmers Home Administration, Nashville, 1974.

Johnson, Russell B. and Albert R. Hagan. "Agricultural Loan Evaluation with Discriminant Analysis," Southern Journal of Agricultural Economics, December 1973.

Jordan, John P. "A Discriminant Analysis of Production Credit Association Loans in East Tennessee," M.S. Thesis, Knoxville: University of Tennessee, 1970.

Kelejian, Harry H. and Wallace E. Oates. Introduction to Econometrics Principles and Applications, New York: Harper and Row, 1974.

Kmenta, Jan. Elements of Econometrics. New York: The Macmillan Company, 1971.

Summary of Accomplishments, Farmers Home Administration, Tennessee, June, 1972, Farmers Home Administration, Nashville, 1971.

Tatsuoka, Maurice M. Multionials Analysis: Techniques for Educational and Psychological Research, New York: John Wiley and Sons, Inc., 1971.

Tennessee Bulletin No. 1310, Nashville: Farmers Home Administration, 1976.

Tennessee Statistical Abstract, 1977, Knoxville: University of Tennessee, 1977.

This is FHA, USDA, Washington: U.S. Government Printing Office, 1972.

Tippett, L. H. C. The Methods of Statistics, Fourth Edition, New York: John Wiley and Sons, 1952.

1974 Summary of Accomplishments, Farmers Home Administration, Nashville, 1975.

1975 Summary of Accomplishments, Farmers Home Administration, Nashville, 1976.

1976 Summary of Accomplishments, Farmers Home Administration, Nashville, 1977.

APPENDIX

EXAMPLE OF LOAN APPLICATION FORM

Position 3

FORM APPROVED.
BUDGET BUREAU NO. 40–R3656

APPLICATION FOR RURAL HOUSING LOANS
(NON FARM TRACT)

(Please Print or Write Plainly)

For County Office Use Only

RACE (*Ethnic*) Code–Check One
W ☐ N ☐ O ☐
AI ☐ S ☐

1. APPLICANTS NAME (FIRST) (MIDDLE) (LAST) | SOCIAL SECURITY NO. | NICKNAME OR KNOWN AS

2. SPOUSE'S NAME (FIRST) (MIDDLE) (MAIDEN) | SOCIAL SECURITY NO. | TELEPHONE NO.

3. COMPLETE ADDRESS (ROUTE, POST OFFICE, STATE, ZIP CODE, COUNTY)

5. HIGHEST GRADE OF SCHOOL COMPLETED BY APPLICANT (CIRCLE APPROPRIATE YEAR)
ELEMENTARY: 1 2 3 4 5 6 7 8
HIGH SCHOOL: 1 2 3 4
COLLEGE: 1 2 3 4 5 6 OR MORE

4. AGES OF PERSONS IN HOUSEHOLD	HUSBAND	WIFE	SONS	DAUGHTERS	OTHERS

6. ARE YOU A CITIZEN OF THE UNITED STATES? ..... ☐ YES ☐ NO
IS YOUR SPOUSE A CITIZEN OF THE UNITED STATES? ..... ☐ YES ☐ NO

7. ARE YOU A VETERAN? ..... ☐ YES ☐ NO IF "YES" INDICATE SERVICE FROM _____ TO _____

8. LOCATION OF HOME TO BE BUILT, PURCHASED OR IMPROVED | SIZE OF LOT

9. IF AN OWNER, WHEN DID YOU ACQUIRE THE HOUSE OR LOT _____, 19____
HOW? (PURCHASED, INHERITED, ETC.)
FROM WHOM?
IF PURCHASED, GIVE PURCHASE PRICE ..... $ _____

10. HAVE YOU EVER OBTAINED A LOAN FROM FARMERS HOME ADMINISTRATION ☐ YES ☐ NO
IF YES, WAS THE LOAN PAID IN FULL? ☐ YES ☐ NO

11. AMOUNT OF LIFE INSURANCE ..... $ _____

12. IF YOU RENT GIVE NAME AND ADDRESS OF LANDLORD
HOW LONG HAVE YOU LIVED IN THIS HOUSE
MONTHLY RENT ..... $ _____

13. NAME AND ADDRESS OF BANK WITH WHICH YOU HAVE A CHECKING OR SAVING ACCOUNT

14. NAME AND ADDRESS OF LENDER WITH WHICH YOU LAST HAD A LOAN

15. EMPLOYMENT RECORD OF APPLICANT. IF SPOUSE IS EMPLOYED, ALSO INCLUDE EMPLOYMENT RECORD. LIST PRESENT EMPLOYERS FIRST AND GO BACK FIVE YEARS FROM DATE OF APPLICATION. (IDENTIFY APPLICANT AS "A" AND SPOUSE AS "S".) APPLICANT'S WORKING HOURS _____ TO _____

DATE OF EMPLOYMENT (FROM – TO)	NAME OF EMPLOYER	ADDRESS OF EMPLOYER	KIND OF WORK	ANNUAL GROSS INCOME	REASON FOR CHANGE

16. CASH FAMILY INCOME	RECEIVED LAST YEAR	PLANNED THIS YEAR	17. CASH FAMILY EXPENSES	SPENT LAST YEAR	PLANNED THIS YEAR
HUSBAND'S TOTAL EARNINGS	$	$	FAMILY LIVING: (FOOD, CLOTHING, UTILITIES, ETC.)	$	$
WIFE'S TOTAL EARNINGS (IF EMPLOYED)	$	$	TAXES PAID	$	$
OTHER NON BUSINESS INCOME (SOCIAL SECURITY, PENSIONS) WELFARE ETC. IDENTIFY)	$	$	CAPITAL GOODS BOUGHT FOR CASH (FURNITURE, TV, CAR, ETC.)	$	$
NET BUSINESS INCOME (GROSS INCOME LESS BUSINESS EXPENSE ATTACH LATEST ANNUAL OPER. STATEMENT)	$	$	ALL OTHER PAYMENTS	$	$
TOTAL	$	$	TOTAL	$	$

FmHA 410-4 (Rev. 9-27-74)

18. FINANCIAL STATEMENT AS OF DATE OF APPLICATION

(Show property owned and debts owed by applicant and spouse)

PROPERTY OWNED		DEBTS OWED					
REAL ESTATE:	CURRENT VALUE	NAME AND ADDRESS OF CREDITOR	FINAL DUE DATE	INTEREST RATE	MONTHLY PAYMENT	AMOUNT DELINQ.	UNPAID BALANCE
	$	DEBTS ON REAL ESTATE:					
					$	$	$
TOTAL REAL ESTATE	$						
AUTO AND EQUIPMENT:							
AUTO	$	TOTAL DEBTS ON REAL ESTATE			$	$	$
OTHER EQUIPMENT (DESCRIBE)		DEBTS ON PERSONAL PROPERTY (SUCH AS AUTO, FURNITURE, TV.)					
					$	$	$
		TOTAL DEBTS ON PERSONAL PROPERTY:			$	$	$
TOTAL AUTO & EQUIPMENT	$	JUDGMENTS:					
					$	$	$
OTHER PERSONAL PROPERTY:							
HOUSEHOLD GOODS	$	ALL OTHER DEBTS (SUCH AS DOCTOR, STORE, ETC. — IDENTIFY)					
CASH ON HAND							
BONDS AND INVESTMENTS							
ACCTS. OWED US—COLLECTIBLE							
OTHER ASSETS (DESCRIBE)							
TOTAL OTHER PERSONAL PROPERTY	$					TOTAL OTHER DEBTS:	$
1. TOTAL PROPERTY OWNED	$	TAXES DUE: REAL ESTATE $_____ PERSONAL $_____					
2. TOTAL ALL DEBTS	$	INCOME & SOCIAL SECURITY $ _____ TOTAL TAXES DUE:					$
3. NET WORTH (1 MINUS 2)	$					TOTAL ALL DEBTS	$

19. PLANNED USE OF CASH ON HAND

20. FOR WHAT PURPOSE DO YOU NEED A RURAL HOUSING LOAN AND HOW MUCH DO YOU NEED TO BORROW?

21. I AM UNABLE TO PROVIDE THE HOUSING I NEED ON MY OWN ACCOUNT, AND I AM UNABLE TO SECURE THE CREDIT NECESSARY FOR THIS PURPOSE FROM OTHER SOURCES UPON TERMS AND CONDITIONS WHICH I CAN REASONABLY FULFILL. I CERTIFY THAT THE STATEMENTS MADE BY ME IN THIS APPLICATION ARE TRUE, COMPLETE AND CORRECT TO THE BEST OF MY KNOWLEDGE AND BELIEF, AND ARE MADE IN GOOD FAITH TO OBTAIN A LOAN.

WARNING: Section 1001 of Title 18, United States Code provides: "Whoever, in any matter within the jurisdiction of any department or agency of the United States knowingly and willfully falsifies, conceals or covers up . . . a material fact, or makes any false, fictitious or fraudulent statements or representations, or makes or uses any false writing or document knowing the same to contain any false, fictitious or fraudulent statement or entry, shall be fined not more than $10,000 or imprisoned not more than 5 years, or both."

DATE _____, 19 ____ _____ SIGNATURE OF APPLICANT

22. FOR COUNTY OFFICE USE:

☆ U. S. GPO: 1975-0-665-664/1755

VITA

Morgan J. C. Scudi was born August 1, 1950, in New York City. He was graduated from Seton Hall Preparatory School, South Orange, New Jersey, in 1968. He attended classes at Bard College, Anandale-on-Hudson, New York; Rutgers University, Newark, New Jersey; and the University of Maryland, College Park, Maryland; and he was a partner in the construction firm of 421 Contractors and Company of Washington, D.C. before enrolling at The University of Tennessee in January 1974. He received his Bachelor of Science degree in Agricultural Economics in March 1975 and entered the graduate program in that same month. He left graduate school in August 1977 to work with Kingston Construction Company, Knoxville, Tennessee, and re-entered graduate school in September 1978 to fulfill the necessary requirements for the Master of Science degree. He is presently enrolled at the Washington College of Law of The American University, Washington, D.C., in pursuit of a Juris Doctor degree.

He is married to the former Nancy Anne Runion of Knoxville, Tennessee.