DISCRIMINATION OF AGRICULTURAL CREDIT RISKS FROM LOAN APPLICATION DATA

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

Edward Ignatius Reinsel

1963

This is to certify that the

thesis entitled

Discrimination of Agricultural Credit Risks
From Loan Application Data

presented by

Edward Ignatius Reinsel

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Agricultural Economics

John K. Brake
Major professor

Date January 30, 1963

O-169

DISCRIMINATION OF AGRICULTURAL CREDIT RISKS FROM LOAN APPLICATION DATA

By

Edward Ignatius Reinsel

AN ABSTRACT OF A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Agricultural Economics

1963

ABSTRACT

DISCRIMINATION OF AGRICULTURAL CREDIT RISKS FROM LOAN APPLICATION DATA

by Edward Ignatius Reinsel

This study was conducted to accomplish the following objectives: (1) to evaluate the importance of various borrower characteristics in discriminating "successful" from "unsuccessful" loan applicants, (2) to develop a model which can aid in discriminating "successful" from "unsuccessful" loan applicants based on information available at the time the loan is under consideration and (3) to evaluate the effectiveness of present loan applications as sources of data for predicting the outcome of loans.

Changes within agriculture, expanded use of short and intermediate term credit and a need for improved loan arrangements suggest greater emphasis on risk forecasting and the use of more objective risk prediction techniques. Historically little has been done to improve risk prediction by the use of objective methods.

Discriminant analysis was chosen as the method for analyzing the data. The function was of the form $P = a + b_1x_1+b_2x_2 \ldots.+ b_nx_n + \alpha$ where the dependent variable is assigned a value of one or zero according to whether the borrower was classified by the lender as "successful" or "unsuccessful." The functions were solved by least square methods. The estimated values for the dependent variable gave an indication of the "probability" of successful repayment for each borrower. Borrowers for

whom the estimated value of the dependent variable was near one were expected to be successful. Those near zero were to be unsuccessful. The actual discriminating value was set at one-half.

Three offices of agricultural lenders provided data. The lenders were asked to select a dichotomous sample of "successful" and "unsuccessful" borrowers. The Farmers' Home Administration and a Production Credit Association were used as data sources since these lenders have more complete information on their borrowers than other agricultural lenders. Prediction models were developed independently for each of the samples used in the analysis. The form of these functions was much alike although the importance of the different variables did change.

The equations which were selected as being most useful were those which could most accurately predict loan outcomes with relatively few variables and had rather high values for R^2. Loan risk formulae were developed from the prediction equation to aid in applying the results of the study to prospective loans.

The results indicate that lenders generally may have overemphasized the role of debts and assets as risk predictors. These factors did not appear to be important in predicting risk.

Factors which seemed to be important for the PCA borrowers were conventional factors such as: farm

ownership, experience on the particular farm and the relationship between non-real estate debts and total debts. Individuals who were able to make annual gains in their net worth by taking risks appeared to be discriminated against by the PCA.

Analysis of the Ingham County FHA sample produced evidence that the relationship between the firm and the household needs to be given more consideration for these borrowers. Other factors which seemed important were: attitudes toward insurance, the relationship between non-real estate and total debts and planned debt repayment. The ability to make annual increases in net worth prior to the loan seemed in the case of these borrowers to be an indicator that the borrower would succeed.

For the Eaton County FHA sample past level of living was an indicator of potential future capacity of the farm to produce needed income. Factors such as the relationship between debt repayment and income, the relationship between non-real estate debts and total debts and the intensity of the farmers' crop program also appeared to be important.

DISCRIMINATION OF AGRICULTURAL CREDIT RISKS FROM LOAN APPLICATION DATA

By

Edward Ignatius Reinsel

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Agricultural Economics

1963

ACKNOWLEDGMENTS

Thanks are expressed to Dr. John R. Brake, who served as committee Chairman for his help, advice and guidance all of which contributed toward the completion of this thesis. Appreciation is also due to Drs. Raleigh Barlowe, James Bonnen, Leonard Rall, and James Stapleton who served as members of the writer's guidance committee.

Acknowledgment for the financial support of this work is due to the Department of Agricultural Economics and the Agricultural Experiment Station of the Michigan State University. A grant from the Michigan Rural Rehabilitation Corporation and funds provided through Regional North Central 53 Committee also provided monetary support.

Sincere appreciation is due to the lenders who contributed data for this study. Particularly Mr. Lyle Buxton, FHA County Supervisor for Eaton County, Mr. Harold Gates, FHA County Supervisor for Ingham County, and Mr. Arnold Musolf, Secretary-Treasurer of the Lansing PCA are due appreciation.

My wife, Edith Reinsel, deserves special thanks for her moral support and encouragement and for typing earlier drafts of the thesis.

TABLE OF CONTENTS

LIST OF TABLES

Chapter I

INTRODUCTION

The Problem in Brief

Agricultural lenders face many prospective borrowers some of whom lack the ability, the resources or the honesty required for debt repayment. The problem, of course, is to identify these borrowers from all the loan applicants. All lenders have some basis for decisions on granting loans. Often these decisions are based on the experience of the lender and result from the application of "rules of thumb". The question to which this thesis is addressed is whether criteria and methods exist which would allow lenders to establish more accurately the soundness of a loan. This, as will be seen, is an investigation into an area which is not well understood. The problem of selecting appropriate criteria is made difficult by the many interrelationships among risk factors and the large number of factors which may affect debt repayment.[1] Usually factors associated with risk are only partially known; thus loan outcomes are subject to considerable uncertainty. Attempts at prediction where uncertainty from

---

[1]The word risk is used here in the same sense normally used by lenders. It includes both risk and uncertainty as defined by economists.

many factors is present may not produce outstanding results but any improvements in the ability to predict the results of loans will help avoid mistakes in lending.

Rapid changes within agriculture together with other uncertainties faced by agricultural lenders have left lenders with few facts for predicting the consequences of alternative lending practices. As a result of this lack of understanding concerning appropriate factors for differentiating good and poor loans, lenders tend to place considerable emphasis on security. Further, little has been done to develop an objective method for differentiating between "successful" and "unsuccessful" borrowers or for determining the liklihood of successful loan repayment.

This study is concerned with the development of an objective function to aid in differentiating good and poor risks. Although historically the asset position of an individual has been considered by some to be of primary importance in approving loans, the question arises as to whether other factors might either replace or be effectively combined with measures of the borrowers asset position to predict more accurately the outcome of loans. A further important question concerns the relative value of different factors in understanding the results of loans. How important are these factors? Are lenders turning down potentially good loans because of inadequate information on farmer characteristics other than equity?

To summarize, some important problems with which this study will deal are:

1. The identification of borrower characteristics which might aid lenders in differentiating good risks from poor risks.

2. The development of an objective function which might improve lenders' ability to evaluate loans.

3. Determinination of the importance of various borrower characteristics as related to loan repayment.

4. Determining whether lenders could better evaluate loans by emphasizing information which is not stressed on present loan applications.

Need for the Study

Today's agriculture is characterized by rapid expansion in size of businesses, technological change and the increased reliance on purchased inputs. Too, it may be that future agricultural capital requirements and thus credit needs will generally be so much larger per farm unit than has been true in the past that loans based largely on colateral will not provide adequate credit.[2]

Total credit use has expanded so that farmers' debts on January 1, 1962 were $25.8 billion. This has more than doubled since 1950. Recent trends in credit

---

[2]Stanley A. Morrow, "Intermediate-Term Credit in Agriculture," Journal of Farm Economics, Proceedings Issue, Vol. XL, December 1958 pp. 1131-1140.

usage indicate an increase in the use of short and intermediate term credit. Outstanding credit secured by chattel mortgages increased from about $7 billion in 1950 to $13.5 billion in January 1962.[3] Further, additional opportunities appear to exist for lenders to extend their loans to larger groups. This will require a better understanding of the relationships between the various borrower characteristics and loan success. Since risk is an important consideration in establishing interest rates, decreased risk which comes with increased knowledge should make the loan market more efficient.

Finally, individuals who have control of insufficient assets may require credit to enable them to provide their families with satisfactory incomes. Under current lending practices individuals who have little collateral often find that their businesses are inadequately financed. Present standards offer little possibility for more satisfactory loan arrangements for these people. A study of the kind indicated here is needed to aid in providing improved loan arrangements for low equity borrowers.

Objectives of the Study

The specific objectives of this study are:

1. To evaluate the importance of various borrower

---

[3]Economic Research Service, United States Department of Agriculture, The Balance Sheet of Agriculture, 1962, Agriculture Information Bulletin No. 270, pp. 18-22.

characteristics in discriminating "successful" from "unsuccessful" loan applicants.

2. To develop a model which can aid in discriminating "successful" from "unsuccessful" loan applicants based on information available at the time the loan is under consideration.
3. To evaluate the effectiveness of present loan applications as sources of data for predicting the outcome of loans.

The analysis which follows will attempt to determine the critical factors in lending money to farmers. The asset-credit position of the borrower may be of major importance and will probably continue to be regarded as being of considerable concern to lenders. However, it may be useful to place greater emphasis on other factors than has been done in the past.

To use these data for discrimination between "successful" and "unsuccessful" credit risks, an analytical model is needed. Thus, a major objective of this study is to develop a statistical function which will distinguish between the two groups of individuals by use of a linear function of the particular variables. Further it is the objective to develop a function to indicate the liklihood of a given borrower being a good credit risk. A model designed to discriminate between "successful" and "unsuccessful" individuals could also show the relative importance of the various factors and indicate

their contribution to the soundness of prospective farm loans.

Organization of the Report

In this chapter an attempt has been made to explain the general nature of the problem, the need for the study and the objectives of such work.

The second chapter is concerned with an examination of the problem. It presents a summary of earlier research as an aid in understanding the current status of such work. This resume is followed by a section concerned with the scope of this study. In a final section of the chapter various credit factors are identified.

In Chapter III the procedure of the study is described. This includes a discussion of the statistical methods used and a definition of terms. The samples are defined and the procedure for data collection is also discussed.

Chapter IV was written to indicate the potential usefulness of loan application data in risk prediction and show some of the problems inherent in prediction from such data. This chapter notes some of the actual data problems encountered in this study. It indicates several factors which need consideration in risk prediction. It is also concerned with the use of statistical models in risk prediction. Finally criteria are presented for evaluation of the functions to be developed later.

The results of the study are shown in Chapter V. The results of each of the samples are presented by discussion of selected equations developed from the data. The chapter is further used to evaluate the effectiveness of the technique for risk prediction. Suggestions are then made for further risk prediction work.

A final chapter was added to briefly explain and summarize the study.

Chapter II

THE PROBLEM OF CREDIT RISK ANALYSIS

A Resume of Previous Investigations

Numerous agricultural credit studies have been published. Often these studies describe past changes in the use of credit and are concerned primarily with policies of lending agencies. Many studies dealing with lending experience have been reported.[4] Jones and Durand summarize several of these and other farm mortgages and income studies and discuss clues as to factors which may be important in selecting credit risks.[5] The second part of their book, a section which deals with farm mortgage distress and individual farm organization, is most closely related to this study. In that section they deal with the question of why some farm units suffer financial distress

---

[4]F. F. Hill, An Analysis of the Loaning Operations of the Federal Land Bank of Springfield from its Organization in March, 1917, to May 31, 1929, Cornell University Agricultural Experiment Station, Bulletin 549, December 1932; Stanley W. Warren, Result of Farm-Mortgage Financing in Eleven Counties in New York State, Cornell University Agricultural Experiment Station, Bulletin 726, December 1939; Charles H. Merchant, Farm Credit in Aroostock County, Maine, University of Maine Agricultural Experiment Station, Bulletin 418, April 1943; Joseph Ackerman and L.J. Norton, Factors Affecting Success of Farm Loans, University of Illinois Agricultural Experiment Station, Bulletin 468 August 1940; Phil S. Eckert and Orlo H. Maughan, Farm Mortgage Loan Experience in Central Montana, Montana State College Agricultural Experiment Station, Bulletin 372, June 1939.

[5]Lawrence A. Jones and David Durand, Mortgage Lending Experience in Agriculture, National Bureau of Economic Research, Princeton University Press, 1954.

even though they are in areas with favorable experience. It should be noted that the data which were collected for these studies were not limited to that available to the lender at the time of the loan. Although their methods may at times better explain what caused difficulties, it may be less useful to lenders for prediction of the outcome of a particular loan. Further, it is useful to note that the above studies and those reviewed by Jones and Durand are dealing with farm land mortgages in contrast to this study which is limited to chattel loans. Appropriate factors for prediction of the outcome of loans may therefore differ. Jones and Durand point out the lack of information concerning personal characteristics of borrowers and farm management techniques used by them. The Jones and Durand book concentrates largely on studies published during the period between World War I and World War II. Present conditions may have changed enough to make the conclusions of these earlier studies inappropriate today.

The literature in fields other than agricultural economics contains several applications of functions for the discrimination between two or more dichotomous classes. Barnard at the suggestion of Fisher used the discriminant function to show a progressive trend in certain cranial measurements of Egyptian skulls.[6] Fisher, whose work is

---

[6] M. M. Barnard, "The Secular Variations of Skull Characteristics in Four Series of Egyptians Skulls," Annals of Eugenics, Vol. 6, London, 1935 pp. 352-371.

regarded as the origin of the discriminant function, first applied the function in an effort to classify plants (Iris) on the basis of measurable characteristics.[7] Fisher later published other articles in which he further developed the function.[8]

Durand, in a research effort initiated by the National Bureau of Economic Research in the area of Consumer Installment Financing, found the discriminant function useful in studying risk elements for consumer credit.[9] This book appears to be the only published study which is both concerned with credit risks and uses the discriminant function as a method of analysis.

Blood and Baker indicate the usefulness of discriminant analysis to agricultural economics research.[10] They demonstrate the use of this analytic technique by its

---

[7]R. A. Fisher, "The Use of Multiple Measurement in Taxonomic Problems," Annals of Eugenics, Vol. 7, London, 1936, pp. 179-188.

[8]R. A. Fisher, "Statistical Utilization of Multiple Measurements," Annals of Eugenics, Vol. 7, 1938, pp. 376-386.

[9]David Durand, "Risk Elements in Consumer Installment Financing," Financial Research Program, Studies in Consumer Installment Financing 8, National Bureau of Economic Research, New York, 1941, p. 125.

[10]Dwight M. Blood and C. B. Baker, "Some Problems of Linear Discrimination," Journal of Farm Economics, August 1958, p. 675-83.

application to the problem of classifying ranches in the Great Plains as either wheat or cattle ranches. The purpose of the classification scheme was to obtain a method for classifying ranches observed at a later date on the basis of the characteristics used in the functions. Blood refers to "discriminant analysis" as "any technique capable of yielding an index which can be used as a critical value for purposes of classification".[11] The article by Blood and Baker was written as an illustration of various techniques of classification and does not attempt to provide research results. Blood and Baker demonstrate three estimating techniques: (1) the linear multiple regression function of the form $Y = a + b_1x_1 + b_2x_2 + \ldots + b_nx_n$, (2) the linear discriminant function of the form $Z = a_1x_1 + a_2x_2 + \ldots + a_nx_n$ where the variables are assigned weights in such a way as to maximize the ratio of the variance of Z between groups to the variance of Z within groups and (3) the "linear probability function," a function of the same form as the linear multiple regression function where when estimating the coefficients of the independent variables the dependent variable is assigned a zero or one according to the classification of the observation into one or the other of the two groups.

---

[11]Blood, Dwight M. "Discriminant Analysis and Farm Management Research," Management Strategies in Great Plains Farming, Great Plains Council Publication No. 19, University of Nebraska, College of Agriculture, Lincoln, August 1961, p. 43.

They point out that with a linear multiple regression problem the dependent variable would need to be quantitatively expressed. This difficulty together with the bias which arises in assigning a discriminating index in the case of the linear multiple regression function leads them to search for more suitable methods for discrimination.

Blood and Baker next investigated the possibilities of Fisher's linear discriminant function. They point out that an optimum[12] index for discrimination can be found by this method.

Blood and Baker next consider the "linear probability function" which they suggest should simplify the computation and add to the value of the results.[13] In their discussion of the "probability" function they indicate that this function is an adaptation of the regression problem obtained by assignment of values of zero and one to the dependent variable. It has many characteristics similar to a regression function. They show that this function gives the same coefficients and the relative

---

[12]Optimum in the sense that it is not possible to find a linear combination of the measurements on the variables in the function which will discriminate more efficiently.

[13]Blood and Baker, op. cit., p. 681.

weights assigned the variables are the same as with the discriminant function. They note that there is a possibility that the "probability function" can provide "probability values" greater than one or less than zero. Since the problem in which we are concerned is of most interest for values of the dependent variable which indicate sensitivity to shifting between the successful and unsuccessful borrower categories, it is unlikely that calculated "probability" values greater than one or less than zero would have serious consequences.

Blood and Baker also conclude that investigation of discriminant analysis would likely prove worthwhile where resources are to be committed on the basis of riskiness. Although they do not specifically mention credit problems, problems of this type appear to be legitimately attacked by the techniques they suggest.

Success in classifying by the use of discriminant analysis and Durand's use of the discriminant function in his credit work suggest that there may be possibilities for transferring similar techniques to problems of discrimination between good and poor agricultural credit risks. Thus far there seems to be a complete absence of such work.

Scope of the Study

This study is economic in nature. The work will be confined to problems in the area of agricultural credit

although similar procedures might yield useful results for other economic questions.

Development of an objective function should aid agricultural lenders in predicting the outcome of loans and should promote more efficient resource allocation by guiding lenders toward providing credit to farmers who can use it most effectively. Success in the study should benefit both the lender and the borrower. The lender might be able to lend to a larger group of clients with similar or decreased risk. The borrower could benefit by finding himself in a group to whom credit was now available whereas in the past credit was impossible for him to find at interest rates and terms which were satisfactory. Presumably, even the person who is denied a loan on the basis of more complete information will benefit for cases where repayment would have been impossible. That is, the denial of a loan would likely be preferred to business failure with its accompanying discouragement and monetary losses.

Data were collected by enlisting the support of three offices of lending agencies. Two of these were offices of the Farmers' Home Administration (FHA) and one was a Production Credit Association (PCA). These sources of information were chosen for several reasons. First, the files of these lending agencies are a source of information concerned with borrowers which was collected

at the time their loans were approved.[14] Thus, it is information which was available for prediction at the time of the loan. Second, if a method of discrimination could be developed with such data, the method would be highly useful to lenders. Third, it was thought that lenders would be best qualified to classify their borrowers according to risk based on the borrower's record since the time of his loan, and finally, completed loan applications should contain much readily available data which would be useful in the study of risks.

At each of the above offices the lender was asked to provide a sample of borrowers who would fall in either of two dichotomous categories. One group was a low risk or successful category; the other group was composed of unsuccessful borrowers. An attempt was made to avoid questionable cases which could not be differentiated until more time had passed or where unusual circumstances made differentiation impractical. Although few lenders are likely to be preoccupied with minor infractions on the loan contract, failure or lack of success are indicated where one or more of the following occurs: excessive delinquency in payment, loan charged off, collection from a cosigner, legal action or an attempt to shift borrowers

---

[14] Myron Wirth, A United States Department of Agriculture employee at the Michigan State University, found in a recent study that rural Michigan bankers typically have incomplete data of the kind needed for this study.

into more liberal lines of credit due to one or more of the above.

The factors that were used for the analysis were those for which data could be obtained from approved loan applications, the original farm plans and other sources available when the loans were approved. Subsequent loans were not considered nor was the information which was collected at the time the subsequent loans were made.

Even though all pertinent data are not contained on loan applications, it is assumed that lenders feel that each item on their loan forms is useful to them in making decisions about their loans. Some of this would be useful for prediction while other items would indicate security. These items need to be tested unless a lack of data or theoretical considerations dictate otherwise. This accounts for the rather large number of variables in some of the functions.

Some of the borrowers who were denied loans could possibly have been successful had their loans been approved. There were insufficient data to develop a function for this group. The function will point out loans that were approved which should not have been granted. Thus, there is a tendency to decrease the total number of approved loans. To overcome this, lenders might wish to compensate by broadening their activities through experimental loans to farmers outside their present experience. They might also look at successes and failures encountered by other

lenders who have in the past followed more liberal lending policies.

The analytical techniques used in this study will demonstrate the use of discriminant analysis as an aid in solving problems in credit risk discrimination. It should provide a method which will apply over broad areas of credit work. Although no single technique will be a panacea for all credit problems, new techniques for lenders may be just as important as new technology which is financed for farmers by agricultural lenders. Possibilities exist for extended use of this type of analysis. Some of the factors which appear important in this study will need to be examined again over broader areas for different classes of lenders and for other types of farms. Since the data on which this study is based come from the records of a PCA and the FHA and are concerned with Central Michigan dairy farms, it is expected that the particular results would apply best under like circumstances. The method used, however, should not have similar limitations.

This study deals with success in loan repayment. It is not within its scope to consider important related problems such as those incurred in collecting delinquent accounts. It is recognized that changes in collection policies may change the desirability of particular credit policies.

Some note should be made of the fact that all loan distress may not be evidenced by delinquency or failure in debt repayment. Some families may be able to avoid delinquency yet suffer deprivation. The scope of this study does not include problems which may arise under these conditions since lenders were asked to differentiate between successful and unsuccessful borrowers on the basis of their success in debt repayment.

Misuse of the function or errors in its development could cause incorrect conclusions to be drawn. Difficulties of this kind would be hard to recognize until some time had elapsed. It is thus not expected--nor would it be desirable--that the results obtained here should replace current lending techniques until they have withstood actual tests in use. It is hoped however that some of the implications can be used to supplement and improve current lending procedures.

Identification of Credit Risk Factors

One of the major problems to be dealt with in this study is the identification of borrower characteristics which might aid lenders in differentiating between good and poor risks. The scope of the study will be further indicated in this section by a discussion of factors which, in theory, could have a bearing on the borrowers creditability. Some of these factors will later be shown to be unimportant as risk predictors. Others cannot be tested

due to a lack of data. Successful development of a statistical function for prediction requires a knowledge of these factors. Further, judgments need to be made regarding appropriate variables for consideration in the equations.

The factors to be considered will be classified under two headings, (1) those related to the general economy or other non-firm influences and (2) those under the control of or influenced by the farm firm.

Non-Firm Credit Risk Factors

Factors external to the firm could play a role in loan risk analysis. For example, the general price level or the parity ratio may have an influence on the outcome of loans made at different stages of the business cycle. Specific prices might be used in risk analysis within an industry. Within the dairy industry milk and feed prices would be relevant. In addition to the absolute level of product and factor prices, the changes in these which occur in the period following the loan are probably important. These price changes may be of theoretical concern even though loan applications could not be expected to provide such information. Differences could exist in the effect of such factors due to the type of farming, the nature and length of the loan and the source of credit.

To the extent that monetary and fiscal policies along with income transfers and price support programs of the Federal Government are effective in controlling

severe agricultural depression, it appears that non-firm forces such as these could, in the future, be a less important guide to debt repayment ability than they may have been in the past. This suggests less reliance on security and assets and assignment of more importance to income and the ability to repay.

Credit Risk Factors On the Firm Level

Of the credit factors concerning the firm there are those dealing with the borrower's resources, his ability and his character.[15] Definable characteristics of the firm are usually some combination of these three. An attempt will be made to classify the various factors even though it is recognized that some factors could be classified in different categories.

Resources--As is customary in economics, resources are sub-classified into land, labor and capital. The borrower's physical resources are among those which can be most readily quantified. In addition to being easier to quantify, resources are sometimes regarded--perhaps inaccurately--as an indication of both past financial acumen and current capacity for debt repayment. Too, the borrower's resources or assets can serve as security in

---

[15]These factors are similar to the three C's of credit -- Character, Capacity and Collateral -- as used by lenders.

the event that the borrower is unsuccessful.

An important resource with which the farmer-borrower works is land. Land or other factors alone would not be expected to be valuable risk indicators. The primary concern is probably the relationship between his resources. Is his management such that he is or will be combining the land resources with labor and capital so as to maximize debt repayment ability?

Often land must be treated as a fixed factor since it is not worthwhile, or it may not be possible, to vary it. When changes occur in the land factor, loan risk may also change.

If it is possible, the examination of land needs to consider both quantity and quality. When available, information on the soil type may provide the lender with an important view of the real worth of land resources. This would be most valuable in areas with heterogenous soils. Also soils need to be considered in relation to their use and the cultural practices applied. Other factors related to the soil type are its variability and depth as well as the fertility and the topography of the farm. The effect of topography may be shown in high tillage costs or slow adoption of mechanization. Soils and topography would be expected to have more pronounced effects under intensive crop programs than would be true where livestock enterprises utilize forage crops and pasturage. Possible expressions of the land factor include soil

classification, land use classification, acreage of tillable land or total acreage of land. One could also study the percent of land in crops or acreage of certain crops to indicate the intensity of land use. Fertilization practices could show the adequacy of the farmer's soil maintainence program.

Another important factor which needs to be considered is the labor available on the farm. This should include both family labor and hired labor. The total amount of labor available would not be expected to be as important as the efficiency with which it is used. Too, consideration needs to be given to the quality of the available labor. Normally farmers with superior intelligence and higher levels of education and training would be expected to be more successful.

Limitations to the physical or mental health of the farmer or his family could affect the business in two ways. First, it is expected that such limitations would cause him to be less able to operate his business and, secondly, he could incur added labor and medical expenses, each of which would drain off needed funds. When planning the labor supply, lenders usually consider potential contributions to the farming operation of the various members of the farmer's household. Large families may provide sufficient labor to make success more likely. But whether or not family members supply labor, a large family would usually require more income for living expenses.

Capital is an important resource in loan risk analysis to the extent that it provides an income base. If it does not add to income and debt repayment ability, it may not be important for prediction of risk. It also needs to be noted that an individual can have the ability to pay debts without having a large amount of capital. The important factor may be his ability to produce income. If he has the needed skills and management ability or control of assets owned by others, he should be able to produce the needed income and should be a good risk.

Availability of capital and other non-human resources is apparent from a financial statement and inventory which are normally obtained at the time of the loan and periodically thereafter. The adequacy of buildings and equipment are generally considered important.

Once the assets have been listed, lenders are generally also interested in the debts of the borrower. These include previous financial commitments such as other loans, unpaid bills and mortgages which might be impediments to successful repayment of newly contracted debts. Logically these commitments should be considered in relation to the assets of real estate, machinery, livestock, debts owed to the borrower by others and income. The size of the debts may not show the financial position as accurately as would measures which reflect the relationship of total debts to short term debts. Along

with the value and composition of his debts the number of sources of credit may also indicate financial management ability.

The repayment schedule may be influenced by the equity of the borrower. Lenders customarily feel that low equity loans are more risky than those for which the borrower also stands to lose significantly should the farm business become defunct.

Most agricultural loans are secured loans. Acceptable collateral may be real estate or chattels. Borrowers are sometimes asked to compensate for the uncertainties involved in their loans due to a lack of security by having a third party to the contract, a cosigner. Other things equal, this should decrease the risk for the lender. It has, however, been found in previous studies that in cases where several cosigners were required the degree of risk involved actually increased. This probably occured since the additional cosigners were used to help compensate for known poor risks and not because the cosigners caused the risk to increase.[16]

Ability--Insufficient emphasis has historically been placed on the ability of the applicant to repay his loan. If successful loan repayment is the goal of lenders,

---

[16]John M. Chapman and associates, "Commercial Banks and Consumer Installment Credit," Financial Research Program, National Bureau Economic Research, 1940, p. 134.

then the ability to repay a loan may be the most important aspect of lending. Without the ability to repay, none of the other factors can make a loan successful. In this sense ability presupposes that the farmer's resources are adequate for repayment although they need not provide a high level of security. Given ability to repay plus the required strength of character, collateral becomes unimportant.[17]

Debt repayment must generally be made from current and future income. Thus it would be reasonable to give attention to some measures of the size and stability of income in relation to loan size. These together with measures of expenses are indicators of his ability to repay and should aid in attempting to predict the creditworthiness of an individual. It has been found in one study that delinquency on farm equipment notes varied inversely with net farm income.[18]

Consideration of income should not exclude income from non-farm sources since this may be applied to agricultural debts in lieu of income from the farm. Sources of non-farm income such as part-time employment, the

---

[17]Ability to repay debts could be defined in such a way that it includes character. This would reduce the necessary factors to the one important factor, ability.

[18]Howard G. Diesslin, "Agriculture Equipment Financing" National Bureau of Economic Research, Occasional Paper #50, Chapter 6, 1955, p. 74.

wife's wages in non-farm employment and income from natural resources may thus materially affect the outcome of a loan. Consideration of all income sources and expenditures may be particularly useful in agriculture since the firm and the household are so closely related.

The rate of repayment may be quite important in considering the liklihood of an individual being able to meet his credit contract. Lenders generally associate higher risks with extended repayment plans. Differences may exist between the individual who assumes short term credit or rapid repayment by choice and others who are forced to accept such conditions to obtain credit.[19] Within a given lender-creditor relationship, several different lengths of term may exist.

Important differences could exist in the risk involved in loans which are largely short term loans. Heavy short term commitments in relations to total debts suggests that large payments would be needed early in the loan period. This may adversely affect the ability of the farmer to repay.

For goods which depreciate appreciably repayment will normally be arranged so that the value of the property will be greater than the remaining debt. Many loans are

---

[19]David Durand, Risk Elements in Consumer Installment Financing, Financial Research Program, Studies in Consumer Installment Financing 8, National Bureau of Economic Research, New York, 1941, p. 55-56.

not paid off as originally planned but are refinanced after the principle has been reduced to some extent. It may be that these refinanced loans bear different risks than new loans.

Given the close relationship between agricultural firms and households, farmers with large families may have more difficulty in repaying debts. This theory is in harmony with a conclusion reached by Brake and Holm who note that "there is strong evidence that families tend to maintain a given level of family living based on family size."[20] They found that family expenditures generally were higher for the larger farm families.

A look at present living standards or those provided by the farm in the past should serve as an indicator of future possibilities. It would need to be recognized that different sets of economic conditions and family characteristics will also play an important role.

It may also be worthwhile to examine the influence of various insurance programs on the farm firm. The fact that an individual has insurance may give an insight into his ability as a manager. It may show that he is concerned about income variability or heavy losses and that he is prepared to take responsibility for his actions.

---

[20] J. R. Brake and C. R. Holm, "The Influence of Household Size and Income On Farm Family Expenditures In Michigan, 1960," Quarterly Bulletin, Michigan State University, February 1962, pp. 541-553.

There is also a possibility that an individual would wish to avoid all risks. He might do some of this by using both formal and informal insurance schemes. In so doing, although his income might be more stable, it would be lower by the amount of the premium or in the case of informal insurance by the additional returns foregone.

Any event would be closely related to the ability to repay debts which could require a large quantity of money at an unpredictable time. The same reasoning applies to events which would seriously affect earning power and thus threaten the well-being of the farm business.

Character - Lenders generally recognize the importance of the character of a particular borrower although they often have little accurate information about him.

Earlier loan experience is generally considered to be rather important by lenders. This may be due to the indications such experience gives about the borrower's character. Borrowers who have demonstrated that they have sufficient resources as well as the probity and ability to repay a loan would likely be allowed to borrow again. Conversely, unsuccessful borrowers would not be expected to be given an opportunity to fail a second time. It is expected that unknown borrowers, particularly new residents in an area, would be accepted with more

reluctance on the part of lenders even though the borrowers may actually be good credit risks.

There may be other useful and measurable indicators of character. Some lenders feel that knowledge of the family and the personal background of the individual are important. Such factors as age, experience and farm ownership may give helpful clues in understanding attitudes and judging characters under certain lending situations. In other cases they may not help.

Not only the character of the farmer but also that of his family may be important. Some lenders feel that many successes and failures of borrowers can be explained by attitudes and are due to the influence of the farmer's wife. It is interesting also to note that many non-farm firms, where firm-household relationships are usually felt to be less interdependent, find the views of their employees' wives to be important to the success of their workers.

The age of an individual could be important. It may be that this factor would need to be considered in relation to his assets and the type of loan for which he had applied. Conceptually, both the very young and inexperienced and the aged that lack resources may be poor risks.

Agricultural credit is by definition mainly for farmers. Differences in borrower occupation are thus less pronounced than in urban areas. Farmers might, however, be classified by dominant enterprise. Lenders

generally question non-typical enterprises. Too, certain enterprises might be charged higher interest rates or have smaller loans available to them. Also, consideration needs to be given to the effect of part time off-farm employment on credit risks. Stability of residence and type of farming as well as experience may also be crucial factors in judging agricultural loans.

Combinations of some of the factors noted above may also give indications of character. For example, the ability of the farmer to accumulate wealth during his early working years may indicate how well he will do in the future.

Adequate Records--Information on many of the previously discussed factors may be difficult to obtain unless adequate records are available. In accurately predicting future income and debt repayment ability some knowledge of past achievements should be worthwhile. Accurate records of the past may give a rather clear expectation of the following period. A lack of adequate records would allow false statements by the borrower or at least allow him to be ignorant of facts which are relevant to his success as a loan applicant. If, for example, it was found that individuals with well kept records were also good credit risks, then lenders might find it worthwhile to give more recognition to this fact by requiring borrowers to keep records. Further, records may indicate a deeper regard for the business aspects of farming which,

in itself, could influence the farmer's success in meeting his obligations.

Farm Management and Risk Analysis

In essence farm management is reaction to uncertainty. It is the process by which the decision maker (1) observes, (2) analyzes, (3) decides, (4) takes action and (5) bears responsibility. The farmer may use the process of management in decisions of either a technical or financial nature. Generally these would be closely related.

Financial Management--The use of the decision making process for reacting to uncertainty involved in dealing with money is financial management. This may concern accumulation of capital, the use of credit or personal money management. The skilled manager will generally apply the decision making process to financial matters. Some of the credit factors which have been discussed are concerned with the effective use of management in financial affairs.

Resources, ability and character each influence financial management. Resource ownership suggests a need to make correct decisions and implies an ability to take whatever action is necessary to maintain the resources. It could indicate that past decisions have been effective for accumulating resources. Credit factors which indicate debt repayment ability should also be rather important demonstrators of capacity for financial management. Ability to repay debts comes from both the potential of the

resources controlled and from the mental and physical capabilities of the borrower. Character is needed to be sure that the borrower uses his resources and his management ability. Probably financial management depends more on character than does management of the technical production processes.

Since management is adjustment to uncertainty, lenders likely feel that a good manager actually decreases the uncertainty which they face. Logically they would be inclined to loan more money given less uncertainty. In the case of the FHA and a growing number of commercial banks and insurance companies attempts are made to provide some management skills to borrowers. This should decrease the lenders uncertainty.

Summary

The purpose of this chapter was to examine the problem of this thesis in more detail than was done in the introductory chapter. In the resume of earlier research an attempt was made to bring together previous work which is concerned with credit and that which is related to discriminant analysis. The scope of the study shows in some detail areas to be considered by this study. To allow an examination of the particular factors which may be important in credit risk analysis individual credit factors were discussed.

CHAPTER III

THE PROCEDURE

The Methodology

In studying prospective loans, decisions need to be made on whether a particular characteristic is important in determining risk and whether an individual will be a good or poor risk.

Conceivably, these decisions could be made without reference to borrower characteristics or previous loan experience. As an extreme example all borrowers could be accepted. If the percentage of potentiality unsuccessful loans in the total population were known, that percentage of individuals could be assigned to the high risk category by chance. Given no better information borrowers necessarily would be selected by such naive models.

Generally lenders do have more information. Conventional procedures for estimating the soundness of a loan are based largely on the experience of the lender. He can sometimes rather accurately predict the outcome of a loan by applying his experience and judgment to clues which are available in a particular case. Factors which are considered may not be well defined. Risk prediction could in this case be classed as an art.

The more usual case would be where the lender selects a few factors which aid him in his decision making. He uses his experience and that of others to determine which factors are important. Under this system difficulties

are apparent where lenders lack needed guidelines and the elements which cause or indicate risk are unclear.

To get better indicators of future loan results borrowers might be divided into two dichotomous groups on the basis of their past performance. Lenders could then identify characteristics which appear to be important in either group. Statistical tests might then be used to indicate whether groups separated on the basis of their past record of repayment were in fact significantly different with respect to other characteristics. One common test in such a situation involves determining whether the means of the two samples with respect to some characteristic are significantly different. If the means are found to be significantly different for a characteristic such as age or family size, the characteristic could then be associated for the purpose of prediction with the populations from which the samples were drawn. Future loans could be made with this additional knowledge.

A statistic which could be used to test the independence of success among borrowers and some characteristic such as age is Chi-Square. The hypothesis would be that success and age are not related. If the null hypothesis is rejected for the Chi-Square test, the characteristic in question would appear to be related to risk.

Problems are immediately apparent when a statistical method is used to classify borrowers into two dichotomous groups on the basis of some one factor. In

particular, such a technique presupposes knowledge of the dominant element and also that one factor will adequately predict the outcome of a given loan.[1] In credit analysis it is likely that no single characteristic would be as valuable a predicter when used alone as would several variables used together. Not only is it likely that several factors are important but interrelationships are probably important. For example, the amount of income available for debt repayment and family size both influence risk. These factors may not be independent. Larger families logically would require more money for subsistence and would thus decrease the amount available for business purposes and hence decrease income.

The Statistical Analysis[2]

To overcome some of the difficulties inherent in the just previously discussed decision making methods a technique known as discriminant analysis was used and seems to be appropriate for this study. Although discriminant analysis was found to be more satisfactory from some standpoints use of this technique did result in some additional problems. The populations of successful and

---

[1]It would of course be possible to repeat the technique for many variables.

[2]Appreciation is expressed to Dr. James Stapleton for his assistance in developing this section.

unsuccessful borrowers are to be distinguished depending on the values of N independent observation variables. A brief discussion of the theory underlying the use of discriminant analysis for two populations follows.

For each population it is assumed that the independent variables have a probability distribution called the multivariate normal distribution. The rectangular array of variances and covariances of these variables is called the covariance matrix. It is assumed that the covariance matrices for the two populations are identical. Since a multivariate normal distribution is determined uniquely by its covariance matrix and its "mean vector," that is, its column of means, the two populations differ only in that the means for the independent variables may differ. In the case that the common covariance matrix Σ is known, it can be shown that the discriminant function should have the form.[3]

(*) $X' \Sigma^{-1}(\mu^{(1)} - \mu^{(2)}) - \frac{1}{2}(\mu^{(1)} + \mu^{(2)})' \Sigma^{-1}(\mu^{(1)} - \mu^{(2)})$,

where $\mu^{(1)}$ and $\mu^{(2)}$ are the mean vectors and X is the column vector of independent variables. This is the linear function of the independent variables which "best" distinguishes between the populations in a probability sense.

---

[3] T. W. Anderson, An Introduction to Multivariate Statistical Analysis, Wiley 1958, Sections 6.1 - 6.5, pp. 126-42.

In practice the matrix Σ is unknown and it must be estimated by using pooled estimates of the variances and covariances based on the observations which were made. The discriminant function is then given by substituting the estimate S for Σ .

In order to use the technique available in regression analysis programs on MYSTIC[4] the following method was used to calculate (*) with S substituted for Σ . The dependent variable was coded as one for a successful borrower and zero for an unsuccessful borrower. The usual multiple regression technique were then used to find an expression: $P = a + b_1 x_1 + b_2 x_2 + \ldots + b_n X_n$. The resulting linear function in the x values can be shown to be identically equal to the expression given by (*) with S substituted for Σ . The P values given by the linear function obtained may be loosely interpreted as probabilities of success. That is, values close to one indicate a high chance that the borrower will be successful; values close to zero indicate a small chance that he will be successful. Since P values can be greater than one or less than zero, references to "probability" will be shown in quotation marks.

The values of P are to be used to discriminate between prospective successful and unsuccessful borrowers.

---

[4]MYSTIC is the electronic computer in use at the Michigan State University.

The value of one-half was chosen rather arbitrarily as the point of distinction. Calculated values of P less than one-half indicate individuals classified by the function as unsuccessful and conversely those greater than one-half would be classified as successful.

The coefficients a_1, b_1, b_2 ..., b_n are of course subject to randomness and it would be desirable to have some measure of their variability. Even in the case that the above assumptions are all met, however, the standard errors of the coefficients have problems associated with them for small samples. If the samples are large, it can be shown that the standard errors of the coefficients are given approximately by the standard errors reported in the corresponding regression analysis problem described above.

The assumptions concerning multivariate normal distributions are of course not met in the present case. This is most obvious when it is noted that some of the independent variables take only the values zero and one. Of course, the assumption of normality in the case of variables such as age and family size are also doubtful. Thus the justification of the method used here must be based on intuitive grounds, and on the fact that it seems to work. Calling the values given by the linear function "probabilities," giving values zero to unsuccessful borrowers, and one to successful borrowers and using regression analysis has intuitive appeal. The results to be reported will demonstrate that the functions obtained are useful. The

multivariate correlation coefficient R obtained by the usual multiple regression techniques for the dependent variable coded with ones and zeros is a valuable measure of the "goodness" of the functions obtained and the value of R^2 will be reported for each discriminant function presented.

Since the assumptions which would allow accurate statistical tests are invalid, the results of this analysis will be descriptive in nature, that is, no tests of hypothesis will be made.

Since it is not known how accurately the standard errors of the coefficients in the discriminant function are estimated it was decided to use the usual standard errors which are calculated with the regression program by the computer as indices of the variability of these coefficients. Thus the variability coefficient reported with each coefficient of the independent variables is an index of the variability of that coefficient and is not the estimate of the standard error. There is good reason to believe that the ratio of a coefficient of an independent variable to its variability coefficient should give a good indication of the worth of that particular variable when used in connection with the other variables in the same linear function. It should be noted that a certain variable may seem to have an important part in the discriminant function when used with certain other variables, but it

may not have such an effect when used with another collection of variables. As a rule of thumb a coefficient might be thought of as important if the ratio of the coefficient to its variability coefficient exceeds two.

Errors in classification are apparent when an individual is classified as being from the population of successful borrowers when in reality he is from the population of unsuccessful borrowers, or is not classified as successful when he should be so classified. Theoretically it would be desirable to consider the cost of such errors in classification. At present little basis exists for making the comparisons needed for the assignment of cost although an individual lender might wish to do so. He could, for example, consider losses from unsuccessful borrowers to be twice as important on the average as the gains foregone by failure to lend to a potentially successful borrower. If this were done an attempt could be made to minimize costs of misclassification. For this study no attempt will be made to differentiate between the costs of the two types of errors. Thus it is assumed that these costs are equal.

Definition of Terms

In order to clarify the discussion which follows and to interpret better the results in a later chapter, a brief definition of terms will be given. R^2 is the proportion of the variation accounted for by the prediction equation.

A coefficient of an independent variable is an estimate of the amount by which the dependent variable changes when the particular independent variable changes by one unit.

A variability coefficient is an index of the variability of the coefficient of an independent variable.

The Sample and Data Collection

A sample of fifty-six creditor records was obtained from the FHA office which serves Ingham County and is located at Mason, Michigan. A second FHA sample included twenty-eight records of borrowers from the Eaton County office of the FHA at Charlotte, Michigan. The Lansing Production Credit Association provided the third sample of forty cases. This association serves six contiguous counties in the lower peninsula of Michigan. The association has its main office in Lansing, Michigan and maintains several branch offices in nearby towns.

At each of the above offices the lender was asked to provide a sample of borrowers who would fall in either of two dichotomous categories. One group was to be a low risk or successful group, the other was to be a high risk, unsuccessful category. An attempt was made to avoid cases which could not be differentiated until more time had passed or where unusual circumstances made differentiation impractical. Essentially, all unsuccessful individuals for whom records were available and who met the other criteria

of first applicant, non-real estate, dairy loans were used. Additional records for individuals in the successful category could have been obtained. This was not done since equal numbers in each of the two groups were felt to be desirable.

The successful cases were those who have progressed in operating their farm businesses and have repaid or have been satisfactorily repaying their debts. Unsuccessful borrowers are those who have been delinquent in their payments and continue to have serious financial difficulty or have quit farming due to financial problems. It is recognized that this classification is dichotomous. Some who were in a given category were "better" risks than others in the same category. No attempt was made to rank the individuals within either category.

Each of the three samples appears to have some characteristics which are unlike the others. Tables 1 and 2 show some of the similarities and differences among the samples.

The PCA Six County Sample--A sample of forty loan records was obtained from the Lansing PCA. One-half were borrowers who were considered to be successful. The other half included those whose performance caused the lender to classify them as "poor" risks. The sample size was limited by the number of loans which met the established criteria.

TABLE 1. A Comparison of Sample Means for Borrower Characteristics by Success Category and Lender

Characteristic	PCA		FHA (Ingham)		FHA (Eaton)	
	Success-ful	Unsuccess-ful	Success-ful	Unsuccess-ful	Success-ful	Unsuccess-ful
Age	38.6	32.4	31.1	36.5	34	28
Family size	4.6	4.4	3.6	4.4	4.2	4.0
% Owners	95	55	32	39	57	50
% Holding life insurance	85	80	93	54	79	86
Amount of life insurance	8738	6469	----	----	----	----
Years operated farm	12.1	4.7	----	----	----	----
% households with running water	----	----	96	78	100	71
Size of original loan	10918	3338	5631	4901	4897	4243
Net worth	76246	19518	8003	9803	14472	10742
Assets	105032	28802	13750	15726	25981	18579
Debts	28786	9284	5747	5923	11509	7837
Non-real estate debt	9292	3583	2305	3100	3764	3484
Number of debts	3.5	3.5	3.6	4.9	4.9	5.5
Planned debt repayment first year	----	----	1970	2079	2635	2434
Expected gross income	----	----	7032	6767	8579	7065
Expected cash operating expense	----	----	2903	2731	4126	3035
Expected living expense	----	----	1845	1837	1774	1620
Total acres	368	189	219	172	211	198
Crop acres	----	----	166	135	163	146
Dairy herd size	$15015[a]	$5485[a]	17	17	19	18
Corn acreage	----	----	49	32	39	58
Number of cases	14	14	28	28	20	20

[a]Value of cattle

TABLE 2. A Comparison of Sample Medians for Borrower Characteristics by Success Category and Lender

Characteristic	PCA		FHA (Ingham)		FHA (Eaton)	
	Successful	Unsuccessful	Successful	Unsuccessful	Successful	Unsuccessful
Age	38.5	31.5	29	36.5	32	27.5
Family size	4	5	3	3.5	4	4
Years operated farm	24.5	29.5	----	----	----	----
Size of original loan	6332	3149	5685	4860	4250	2775
Net worth	52695	12576	7621	5680	11138	5499
Assets	67260	19415	11745	8200	22307	14410
Debts	18588	5463	3455	3239	8310	5562
Non-real estate debt	4363	3174	1100	2028	2253	2260
Number of debts	3	3	3	4.5	4.5	6
Planned debt repayment first year	----	----	1463	1602	1965	1980
Expected gross income	----	----	6733	5518	7190	5700
Expected operating expense	----	----	2900	1966	3449	2563
Expected living expense	----	----	1647	1669	1582	1656
Total acres	355	189	200	164	180	180
Crop acres	----	----	150	133	130	145
Dairy herd size	$4000[a]	$12050[a]	19	17.5	16	15
Corn acreage	----	----	48	30	34	50
Number of cases	20	20	28	28	14	14

[a]Value of cattle

Among the successful cases in this sample, there were no cases which have resulted in actual losses to the association. In fact, there have been almost no losses in the association for several years. Borrowers were classified as unsuccessful by the lender where serious delinquency existed along with a general lack of financial progress. No limits were set on established real estate loans from other sources or on such loans being approved at a later date. Such loans were always made through another lender since PCA loans are, as a matter of policy, secured by chattel mortgages.

Data obtained at the PCA office were in many ways similar to those obtained at the FHA offices. Earlier PCA loan applications were less detailed. Apparently due to the strong equity position of some PCA borrowers the applications were sometimes not entirely completed. For example, borrowers with a relatively small debt load in relation to assets were not always required to furnish information needed to complete the income and expense section of the application.

The Ingham County FHA Sample--The largest sample was the Ingham County FHA sample. Fifty-six cases were included of which twenty-eight were successful and twenty-eight were unsuccessful. Of the unsuccessful group it was found that many individuals were no longer farming. In some cases they have quit farming at the suggestion of the

Farmers' Home Administration County Supervisor. Losses were actually incurred in more than one-third of the unsuccessful cases. In other cases creditors other than the Farmers' Home Administration suffered losses. In still others there is little doubt but that future losses could easily result. In none of the unsuccessful cases who are still farming is there reason to think that financial progress is being made even though the Farmers' Home Administration may have sufficient security to avoid financial loss.

A few of the farmers classified as good risks have quit farming for reasons such as health, age or off-farm opportunities.

In some major respects individuals within this sample have similar characteristics. First, each of the borrowers is primarily a dairy farmer although several were not dairymen when they established their loans. Some sought FHA assistance in getting established in farming. Secondly, all of the original loans are of either short of intermediate term, thus no real estate loans are included in the sample. This neither precludes subsequent loans for real estate nor the existence of real estate loans from other lenders. All of the loans were secured by crop and chattel mortgages.

Most of the farmers included in this sample were residents of Ingham County at the time of their initial Farmers' Home Administration loan. Six had their first

loan in another county. One-half of these were made by the Ingham County supervisor while he was responsible for their counties due to shifting of the areas of responsibility for the county supervisor. Of those coming from other counties three were successful and three were unsuccessful.

The Eaton County FHA Sample--A second sample of FHA loans was obtained at the Charlotte, Michigan office of the FHA. The County Supervisor was asked to provide a sample which would meet the same criteria as that used in the Ingham County FHA sample. Twenty-eight records were obtained. Again cases which the county supervisor could not satisfactorily classify were not used. All of the borrower cases for which records were available and which were considered by the county supervisor to be poor risks were used except those who failed to meet the other criteria of first applicant, non-real estate dairy loans which were closed between 1950 and 1960.

Unlike the other FHA sample unsuccessful cases in this sample did not cause actual losses to the FHA. This is true since few losses have been incurred at this office. Rather they were loans which were delinquent on farms that were making little or no financial progress or appeared in the county supervisor's judgment to be in financial trouble.

Again the fact that an individual had decided to quit farming was not used as an indicator of success or failure.

The sample consisted entirely of dairy loans of short or intermediate term secured by crop and chattel mortgages. Although real estate loans are not included, no restrictions prohibited real estate loans from other sources or such loans having been made at a later time by the FHA.

Three of the loans in this sample were made in other counties. One of these was made by the Eaton County supervisor while he was responsible for another county. Of these three loans the two made in other counties by different county supervisors were found to be unsuccessful by the Eaton County supervisor.

Summary

This chapter has described the procedure followed in collecting and analyzing the data. Each of the samples was described. The following chapter will include a discussion of the problems of predicting risks from loan application data and the chapter following that will show the results of the study.

CHAPTER IV

PREDICTION FROM LOAN APPLICATION DATA

Some Data Problems Encountered

Selection of Variables--When one attempts to develop a function such as the one described, two questions arise concerning the variables. First, has the variable in question sufficient theoretical basis? It is known that, given a large number of variables, even though randomly chosen, the apparent explained variation can be made very high. One might also select from a very large number of variables a few which would appear to be good predictors. These could occur by chance. In this study each variable which was used was first studied to determine if there were valid theoretical arguments for including it in the analysis.

The second question which must be answered concerns the availability of the data. Theoretically important variables for which information is unavailable may help rationalize the unexplained relationships but can hardly be useful otherwise. Unfortunately several variables which logically seem to be important were necessarily omitted for lack of data. Information that reflects on ability and character--which may be the key to risk prediction--is difficult to set down in an orderly manner, is not easily summarized and is often scanty. It is, therefore, not well represented in the analysis.

Nature of Agricultural Credit--The nature of agricultural credit restricted this study. First, individual agricultural loan offices do a rather small volume of business, therefore, they have relatively few cases. Further, they typically have had low losses and thus few unsuccessful borrowers in the post World War II years. In addition within a given area farms tend to be homogeneous with respect to enterprises. Each of these factors tends to decrease the sample size.

Further Problems in Studying Agricultural Credit--In this study efforts were made to find first applicant loans. Actually a large percentage of short-term agricultural credit is not for new loans but rather as a continuing line of credit to individuals who are well known and who have already demonstrated their capabilities. It is expected that risk prediction among known borrowers should not be as difficult as it would be for new borrowers. Difficult questions can result, however, even when loans are based on a present loan situation which either progresses or deteriorates. PCA loans in particular are frequently made for one year at a time with the understanding that they can be renewed when due. Under these circumstances individual loans are paid up annually. Renewal of the loan may be accompanied by increased loan size. Even where loans have deteriorated, lenders may sometimes be forced to make further advances to prevent a total collapse at some inopportune time.

Possibly methods such as interviewing borrowers could be a more suitable data collection method. Yet, serious limitations are likely in obtaining information from borrowers through an ex post interview. Among the cases in this study, it was found that several of the borrowers could have no longer been contacted as they had moved without leaving a forwarding address or were deceased. At least one unsuccessful borrower was confined to a penal institution. Still other problems could be expected in attempts to obtain information from borrowers due to the rather long period of time since the loans were approved. It is likely that much information which might have been obtained by interview at the time of the initial loan would have been forgotten by the borrower. Too, borrowers who have actually defaulted may refuse to discuss the loan due to fear of collection attempts.

A study based on data from applications of credit agencies does not represent all income classes. This is true since samples of farmers who have successfully obtained loans from lenders will not include "unacceptable" credit risks. Very low income farmers would undoubtedly fall in this "unacceptable" category. Since loans are usually to be paid out of income, seriously underemployed and marginal farmers do not have credit available from today's agricultural lenders.

Lenders probably classified their borrowers for this study as successful or unsuccessful chiefly on the

basis of their experience and intuition. It seems unrealistic to assume that lenders from different areas serving different classes of farmers, under different lending policies would agree on what factors would cause them to classify a farmer in one category or another. It may be true that the better prospective borrowers in the FHA samples would receive poor classifications from other lenders.

One's first inclination in an effort to overcome some of these data problems is to broaden the sample to include many lenders. This needs to be a further step after first determining whether discriminant analysis is useful in discriminating borrowers for particular lenders. Combining information from many lenders was thought to be unwise at this time given the incompleteness of the data and the limited experience in risk discrimination. It was felt that more fruitful results could be obtained by more intensive analysis of a small number of lenders.

Inadequacies in Loan Applications--Data problems encountered in this study point to inadequacies in loan forms which were in use at the time these loans were made. No doubt today's improved application forms, if complete, would provide some information which was unavailable earlier. Yet it is also likely that major improvements are still possible in loan applications. They might be made to serve as better risk prediction tools.

More serious problems exist for the many agricultural lenders who obtain almost no information about their borrowers. The lack of data has resulted in the omission of commercial bankers from the sample. This is an important matter which needs to be considered if better data become available.

Among the factors of interest to lenders at the firm level, physical resources have historically been given the most emphasis on loan forms. Evaluation of land, buildings, equipment and livestock as currently done may serve fairly well as an indicator of security. Theoretically it can be argued that such a method lacks much as a useful part of the system for prediction of the ability to repay loans. Standard loan forms could easily err by assigning inappropriate values to asset inventories. For example, a modern self propelled combine might be valued at $5,000 even though it was located on a small farm where it added little to income. Such items need to be valued with respect to possible use unless they are to be sold. Loan forms assume that each piece of equipment can be valued alone. Often farm machines are near perfect compliments and hence need to be considered in relation to the complete set. Further, unless the operation is large enough to use both efficiently, duplication of items of equipment which are substitutes adds to a net worth statement but not necessarily to income from which debts are repaid.

Buildings also need to be considered in relation to their productivity. Inventory values may not reflect actual values to the farming operation. Even though well maintained and structurally sound a barn could be inefficient or unneeded, have high maintenance costs and perhaps increase taxes. Loan forms which contain more detail in physical terms concerning livestock capacity of buildings could better aid in determining the potential size and efficiency of the business.

Lenders typically have made little use of soil maps in loan analysis. Some information which is available today was not available earlier. The use of information about soils supposes that the individual is a permanent resident. This was not always true for the individuals in the study.

Ability to repay is crucial; yet present loan forms lack effective indicators of ability. How the farmer will repay is not always noted. Potential levels of production for crops, animals or animal products and measures of efficiency are generally not shown on the loan forms. Labor productivity and income per man are unavailable. Loan forms often do not clearly show how the gross income is to be used. Questions exist about how much will be needed for operating and living expenses and what will be left for debt retirement.

Character is almost neglected in present loan analysis and is probably the most difficult to estimate of

Buildings also need to be considered in relation to their productivity. Inventory values may not reflect actual values to the farming operation. Even though well maintained and structurally sound a barn could be inefficient or unneeded, have high maintenance costs and perhaps increase taxes. Loan forms which contain more detail in physical terms concerning livestock capacity of buildings could better aid in determining the potential size and efficiency of the business.

Lenders typically have made little use of soil maps in loan analysis. Some information which is available today was not available earlier. The use of information about soils supposes that the individual is a permanent resident. This was not always true for the individuals in the study.

Ability to repay is crucial; yet present loan forms lack effective indicators of ability. How the farmer will repay is not always noted. Potential levels of production for crops, animals or animal products and measures of efficiency are generally not shown on the loan forms. Labor productivity and income per man are unavailable. Loan forms often do not clearly show how the gross income is to be used. Questions exist about how much will be needed for operating and living expenses and what will be left for debt retirement.

Character is almost neglected in present loan analysis and is probably the most difficult to estimate of

the factors which are closely related to the firm. What information is available is general and subjective. This limited information is probably more useful to lenders than for statistical problems since loan application notes concerning character often are recorded differently for each individual. Almost nothing of value to this study was found on the loan forms concerning character.

The Use of Statistical Models for Risk Prediction

An aim of this study was to formulate the problem of risk prediction in terms of a statistical model and then to test the model with respect to actual data. Acceptability of the model for prediction needs to be determined. Criteria need to be established to determine whether the objectives were successfully accomplished. To what extent is the function capable of predicting loan outcomes better than naive or less sophisticated models? Can it predict better than lenders who have access to loan applications plus the more subjectively evaluated clues obtained by interviews or by personal visits to the farms of the borrowers?

Statistical techniques are useful to the extent that they aid in interpretation of data. Useful results rely on selection of the appropriate statistical model. Discriminant analysis appears to be an appropriate technique for this problem.

In this study discriminant analysis was used in an attempt to predict the value of one variable, the loan outcome, from the values of given variables. It can also be used to determine the effects of the other variables on the one being estimated.

Criteria for Evaluation of the Functions

Ideally R^2 should be large. For the functions used in this study it was expected that R^2 would be lower than would have been anticipated for a normal multiple regression problem due to the discontinuous nature of the dependent variables.

The methods used in this study may also be judged by the accuracy with which individuals are placed in the appropriate categories. Another criterion for judging the function is the improvement in prediction over the common procedures used in lending or compared to use of some single factor for making the decisions about lending.

Summary

This chapter has described some of the questions which need to be resolved in predicting risks from loan application data. These include the problems which arise in selecting variables to be tested and the lack of the kinds of data needed. Sections were written to point out considerations in the use of statistical models in risk prediction. Finally, criteria were established for evaluating the function.

CHAPTER V

RESULTS OF THE RISK PREDICTION MODELS

Analysis of the PCA Data

Of the equations which were fitted to the PCA data several are presented in Table 3. An attempt was made to study all of the factors which might affect loan risk and for which adequate information was available.

An indication of the importance of the several variables in the functions can be seen by observing the affect on the proportion of the variation accounted for by the prediction equation of deleting particular variables. These values are presented in Table 4 for the functions shown in Table 3. In four of the five functions deletion of the variable for total acres decreased R^2 more than deletion of any other variable. In the one exception, farm ownership appeared to be a more important contributor.

Farm Size--Farm size as measured by total acres was found to be related positively to success. Interestingly, total acres appear to be a better risk predictor than total crop acres. This may be indicative of the economic value of pasture land on dairy farms.

Farm Ownership--Among PCA borrowers owners appeared to be better risks than renters. Ownership could indicate greater financial strength and more involvement in the business. Owners have more to lose should the farming operation fail.

Unfortunately this factor may have influenced the lender in his decision about whether the borrower was successful or unsuccessful. This is suggested as a possibility although no evidence is offered as proof. Even though this may be true, the function could still aid the lender in knowing the likelihood of each borrower being in the category of successful as he has defined it. More importantly, it should indicate to the lender factors which could compensate for a lack of ownership.

Experience--In each of the PCA equations the variables for years of experience on the particular farm had a positive coefficient. This is as would be expected and indicates that the more experience a borrower has had on a particular farm the more likely he is to be successful. This variable also appears to reflect the importance of a degree of stability. It gives some indication that well established farmers are unlikely to default on loans. Occupational and residential stability were both found to be important risk factors by Durand.[1]

Number of Sources of Credit--It appears that individuals who come to a PCA with many sources of credit should be looked on as being riskier borrowers. This seems reasonable since numerous credit sources probably reflect poor past financial management. Obtaining an unsupervised loan would not usually alter the borrower's

---

[1]Durand, op. cit., pp. 65-67.

TABLE 3. Selected Discriminant Functions for PCA Borrowers[a]

I_a $P = .14927 + \underset{(.00719)}{.01671}\ X_1 - \underset{(.02558)}{.03848}\ X_2 + \underset{(.19100)}{.21158}\ X_3 + \underset{(.00039)}{.00161}\ X_5 - \underset{(.00001)}{.00002}\ X_7 - \underset{(.21881)}{.42333}\ X_8$

$R^2 = .63$ (classified 38 of 40 correctly)

I_b $P = -.15747 + \underset{(.00757)}{.01766}\ X_1 - \underset{(.02431)}{.05601}\ X_2 + \underset{(.1430)}{.41725}\ X_3 + \underset{(.00038)}{.00141}\ X_5$

$R^2 = .55$ (classified 35 to 40 correctly)

I_c $P = -.11898 + \underset{(.00795)}{.01676}\ X_1 - \underset{(.02837)}{.05004}\ X_2 + \underset{(.16342)}{.44931}\ X_3 + \underset{(.00042)}{.00134}\ X_5 - \underset{(.48896)}{.20645}\ X_6$

$R^2 = .55$ (classified 35 of 40 correctly)

I_d $P = -.15759 + \underset{(.00743)}{.01517}\ X_1 - \underset{(.02365)}{.06109}\ X_2 + \underset{(.14121)}{.47135}\ X_3 + \underset{(.00041)}{.00172}\ X_5 - \underset{(.00001)}{.00002}\ X_7$

$R^2 = .59$ (classified 34 of 40 correctly)

I_e $P = .28682 + \underset{(.00786)}{.01958}\ X_1 - \underset{(.02677)}{.04573}\ X_2 + \underset{(.14466)}{.43564}\ X_3 - \underset{(.000002)}{.000002}\ X_4 + \underset{(.00075)}{.00201}\ X_5$

$R^2 = .56$ (classified 34 of 40 correctly)

[a]The variability coefficient is noted in parentheses for each estimated parameter. R^2 is shown for each function. The variables are described below.

P = An estimated value for the dependent variable. The "probability" that a particular individual will be successful. Individuals for whom values of P are estimated to be greater than .5 are those who are expected to be successful. Those for whom values of P are estimated to be less than .5 are expected to be unsuccessful.

X_1= The number of years that the farmer has operated the present farm.

TABLE 3. (Continued)....

X_2 = The number of sources of credit which were disclosed by the prospective borrower.
X_3 = A one for borrowers who own their farms and zero for renters.
X_4 = The dollar value of the borrowers assets.
X_5 = The size of the farm as measured by total acres of land.
X_6 = The ratio of debts to assets.
X_7 = The average annual increase in net worth after the borrower is 20 years old.
X_8 = The percent of the total debts which are non-real estate debts.

[a] Additional functions are shown in the appendix.

TABLE 4. R^2 Values Resulting from Deletion of Each Variable in the PCA Prediction Equations[a]

Function Number	R^2 for the function	R^2 Value with Each Variable Deleted							
		X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
I_a	.63	.57	.61	.62		.45		.59	.59
I_b	.55	.48	.48	.44		.38			
I_c	.55	.50	.51	.45		.42	.55		
I_d	.59	.54	.51	.46		.38		.55	
I_e	.56	.48	.52	.45	.55	.47			

[a] The value shown in the column "R^2 for the function" shows the value of R^2 for the functions shown in Table 3. The remaining columns show the values obtained when the functions were changed by deletion of any one of the variables.

money management practices. It may be that individuals with a large number of debts at the time of their first PCA loan will continue to obtain credit from many sources. This could occur even when the lender consolidates many of the borrower's previous debts.

An exception may exist to the above reasoning. That is, several sources of credit might also indicate that credit has been inadequate for the needs of the particular farmer. Given this situation the better financial manager might be forced to establish many lines of credit.

Net Worth Increases--Another variable which appeared to aid in risk prediction was one which shows the ratio of net worth to productive years of work. It was expected that this variable would have a positive coefficient. As is shown on Table 3 the coefficient was negative. This raises a question about why individuals who were able to accumulate net worth would be poorer risks. As an explanation it is suggested that these farmers were in fact able to make greater net worth gains by taking sizeable risks. If these borrowers accepted the risks of heavy indebtness as a means of growth, the lender may have been inclined to discriminate against such borrowers. These loans may in fact be more risky and the lender may be justified in discriminating against risk takers but they would at the same time be discriminating against those who have been able to make the greatest financial gains.

The Ratio of Non-Real Estate Debts to Total Debts--Large non-real estate debts or short term commitments are apparently a detriment to effective repayment. As a variable the ratio of non-real estate debts to total debts resulted in negative coefficients. This is what would be expected theoretically. Debts of this type require heavy payments early in the loan period. Some farm firms may not be able to meet such payments even though they could pay off long-term debts.

For the case of an individual with very low total debts where all of his debts were non-real estate debts, this ratio would be high (equal to 1) just as it would be high for another individual who had very large non-real estate debts. This probably affected the importance of this particular variable.

Total Debts--This factor alone did not appear to contribute importantly to the function. R^2 was essentially unchanged when the variable was deleted. The sign of the coefficient was negative as expected.

Total Assets--The amount of the farmer's assets also appears to be unimportant as a variable in the function. In this case the sign of the coefficient was negative rather than positive as expected. The reason for this is not clear. One possibility that seems plausible is based on the high positive correlation between debts and assets. It is probably more risky to have large assets and large debts than it is to have small assets and small debts.

The Ratio of Debts to Assets--It seems logical that information about debts and assets should be considered together. However, this variable appeared to be rather unimportant even though the sign of the coefficient was negative as expected.

Life Insurance--The coefficient for the variable concerning life insurance was positive as expected. Life insurance appeared not to be a valuable risk indicator for the PCA sample. One explanation may be the fact that PCA's offered credit life insurance to their member borrowers. They may encourage those borrowers who appear more risky at the time that they apply for a loan and have little or no life insurance to purchase either the PCA credit life insurance or to obtain life insurance from other sources.

Personal and Family Background--The data on which this study of PCA loans is based contain few indicators of personal and family background. The few possible variables such as age and family size appeared to add little to the prediction function. It may be that for highly commercial farms such as found in the PCA sample, firm-household relationships do not affect the firm as strongly as for smaller units. It is also possible that more complete data would show relationships not apparent with present data.

Choice of a Function and Borrower Classification

The function which was chosen to represent the PCA data was function Ia. A factor which contributed to its selection was the relatively high value of R^2 compared to many of the functions. For this function R^2 was .63. The function also was quite successful in classifying the borrowers with a relatively small number of variables. Thirty-eight of the forty borrowers in the sample were classified correctly. Of the two who were misclassified one would have been expected to be unsuccessful based on the function while in reality he was successful the other was expected to be successful but was unsuccessful. The function seems to have been quite effective when compared with naive models, one variable model, or the lenders scheme for approving borrowers.[2] The classification of borrowers is more apparent if the range of "probability" values is divided in half at the mean "probability" value of the sample (.5) and then each of these halves is subdivided at the mean of the successful (.82) and unsuccessful (.18) classes. This has been done in Table 5.

By looking at the data for the two borrowers who were misclassified, it is possible to suggest reasons for

---

[2]This makes the tenable assumption that the lender would not have approved loans for some of the unsuccessful borrowers who were correctly classified by the function had he known the outcome of their loans.

TABLE 5. Classification of the PCA Sample According to Estimated "Probability" Values (Function Ia)

	Number of Borrowers				Total
Successful	0	1	8	11	20
Unsuccessful	10	9	1	0	20
"Probability of success"	.8	.5	.82		

their misclassification. Farmer I, although an owner, was situated on a rather small farm compared to the rest of the sample and lacked experience yet was successful. Farmer II was an owner with considerable experience and no debts but was unsuccessful. It should also be noted that these two cases are still active. This allows for the possibility that they could at some later date be classified differently by the lender.

TABLE 6. Characteristics of the Misclassified Farmers in the PCA Sample

Characteristic	Sample Mean	Farmer I	Farmer II
Years operated farm	8.4	2	14
Number of debts	3.5	3	6
Farm Ownership	75%	owner	owner
Total acres of land	279	160	220
$\frac{\text{Non-real estate debt}}{\text{Total debt}}$	.48	.47	0
$1000 average annual change in net worth per year after age 20	4.03	6.04	2.82

PCA Credit Rating Formula

Using function Ia it is possible to develop an index which can be used to classify new observations. These new observations would be classified on the basis of the kinds of data available from this sample. Also the index would be based on the assumption that the successful and unsuccessful classes are defined in the same way that they were defined when the lender classified his borrowers. Although a particular function was chosen for developing the index, other indexes could also be based on similar functions.

The formula should predict most accurately for values near the sample mean. These mean values are given in Table 6.

The formula considers the six factors listed in Table 6 to rate a prospective borrower. The weights assigned the factors are as follows:

Years Operated Farm--Add .017 for each year the borrower operated the present farm.

Number of debts at time of loan application--For each debt subtract .042.

Farm ownership--Add .31 for farm ownership.

Total acres--Add .017 for each 10 acres of land.

The ratio of non-real estate debts to total debts--The largest possible value for this factor is 1. If the value recorded for an individual is 1, subtract .3. If it is less than 1, subtract a

proportionally smaller value. For example, if the value is .5, subtract .15.

Average annual gain in net worth--For each $1000 average annual increase in net worth after age 20 subtract .023. (See page 61 for the reason this is subtracted rather than added.)

Factor Substitution Rates

Another point of interest is the rate at which the different variables in the function substitute for each of the other variables in the function. These values are shown in Table 7. It can be seen that the values on the upper right side of the table and those on the lower left are reciprocals. The sign of the coefficients of the variables is shown in each case and needs to be taken into account. Directions for using the table are provided at the bottom of the table.

Analysis of the Ingham County FHA Data

Several equations which will be used to explain the results of the work with the Ingham County FHA sample are shown in Table 8. This analysis should aid in risk prediction under circumstances similar to those encountered by these borrowers.

The effect on R^2 of deleting each of the variables in turn is shown in Table 9. In each case deletion of the variable concerned with life insurance affects R^2

TABLE 7. Rates of Substitution Between Variables Used in the PCA Functions (Function Ia)[a]

Units of Variables	One year experience on this farm	Each debt	10 acres land	An increase in the ratio: non-real estate debt / Total debt of .10	$1000 average annual change net worth	Farm Ownership
One year experience on this farm	1	2.54	10.3	1.80	1.35	18.71
	+	-	+	-	-	+
Each debt	.393	1	4.04	.708	.529	7.35
	-	+	-	+	+	-
One acre land	9.74	24.8	1	17.6	13.1	182.2
	+	-	+	-	-	+
An increase in the ratio: non-real estate debt / total debt of .10.	.55	1.41	5.7	1	.75	10.38
	-	+	-	+	+	-
$1000 average annual change net worth	.74	1.89	7.6	1.34	1	13.89
	-	+	-	-	+	-
Farm ownership	.053	.136	0.55	.096	.072	1
	+	-	+	-	-	+

[a]Read across the table, then up. For example 2.54 years of experience is approximately equivalent to one debt or, as a second example, 9.7 acres of land is approximately equivalent to one year of experience.

TABLE 8. Selected Discriminant Functions for Ingham County FHA Borrowers[a]

II_a $P = .64848 - \underset{(.03236)}{.06457}\ X_1 + \underset{(.16924)}{.30069}\ X_2 + \underset{(.12675)}{.48395}\ X_3 - \underset{(.12075)}{.27624}\ X_4 - \underset{(.15401)}{.29341}\ X_5 - \underset{(.20884)}{.40597}\ X_6 + \underset{(.00003)}{.00006}\ X_8$

$R^2 = .44$ (classified 45 of 56 correctly)

II_b $P = .93644 - \underset{(.03266)}{.07344}\ X_1 + \underset{(.12894)}{.50498}\ X_3 - \underset{(.12329)}{.28431}\ X_4 - \underset{(.15606)}{.32862}\ X_5 - \underset{(.21141)}{.35570}\ X_6 + \underset{(.00003)}{.00006}\ X_8$

$R^2 = .41$ (classified 43 of 56 correctly)

II_c $P = .10598 - \underset{(.03311)}{.05156}\ X_1 + \underset{(.18070)}{.27029}\ X_2 + \underset{(.13229)}{.53962}\ X_3 - \underset{(.42193)}{.44826}\ X_{10} + \underset{(.57154)}{.86036}\ X_7 - \underset{(.12756)}{.26913}\ X_4$

$R^2 = .38$ (classified 44 of 46 correctly)

II_d $P = .40247 - \underset{(.03290)}{.06302}\ X_1 + \underset{(.17664)}{.24830}\ X_2 + \underset{(.12897)}{.55381}\ X_3 - \underset{(.12694)}{.31304}\ X_4 + \underset{(.55028)}{.85784}\ X_9 - \underset{(.15575)}{.25704}\ X_5 - \underset{(.20966)}{.29519}\ X_6$

$R^2 = .42$ (classified 46 of 56 correctly)

II_e $P = .51925 - \underset{(.03245)}{.06101}\ X_1 + \underset{(.17542)}{.28164}\ X_2 + \underset{(.12753)}{.56988}\ X_3 - \underset{(.12794)}{.27154}\ X_4 - \underset{(.04256)}{.06609}\ X_{10} + \underset{(.54629)}{.95920}\ X_9 - \underset{(.15815)}{.31608}\ X_5 - \underset{(.20703)}{.31461}\ X_6$

$R^2 = .45$ (classified 48 of 56 correctly)

Each variable used in the functions listed in this table is described below. The variability coefficient for each estimated parameter is noted in parenthesis. R^2 is also shown for each function.

TABLE 8. (Continued)...

P = An estimated value for the dependent variable. The "probability" that a particular individual will be successful. Individuals for whom values of P are estimated to be greater than .5, are those who are expected to be successful. Those for whom values of P are estimated to be less than .5 are expected to be unsuccessful.

X_1= The size of the borrower's family.

X_2= One for borrowers with running water in their homes and zero for those without running water

X_3= One for individuals who have life insurance and zero for those without such insurance.

X_4= One for borrowers with health insurance and zero for those without health insurance.

X_5= The percent of the total debts which are non-real estate debts.

X_6= The percent of the expected net cash income which will be used for debt repayment in the first year.

X_7= The percent of the expected gross cash income which will be used for debt repayment in the first year.

X_8= The average annual increase in net worth per year of productive work.

X_9= The percent of the total cropland in corn or to be planted to corn.

X_{10}=Planned fertilizer expenditure per crop acre in dollars.

[a] Additional functions are shown in the appendix (Table 3).

TABLE 9. R^2 Values Resulting from Deletion of Each Variable in the Ingham County FHA Prediction Equations[a]

Function Number	R^2 for the function	R^2 Value with each variable deleted									
		X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
II_a	.44	.40	.41	.27	.38	.40	.40		.39		
II_b	.41	.34		.22	.34	.35	.37		.36		
II_c	.38	.34	.35	.16	.32			.35			.36
II_d	.42	.38	.40	.20	.35	.39	.40			.39	
II_e	.45	.41	.42	.22	.40	.41	.42			.42	.42

[a] The value shown under the column heading "R^2 for the function" shows the values of R^2 for the functions shown in Table 3. The remaining columns show the values obtained if each of the functions were changed by deletion of any one of the variables.

more than deletion of any of the other variables.

Life Insurance--The FHA sample borrowers who hold life insurance appear to possess some attribute which makes them more likely to be good credit risks. This variable appeared to be rather important as an aid in risk prediction. This factor was also found to be important by Durand in his consumer installment credit study. As has been noted, a similar relationship was not found to exist for the PCA sample. A question exists concerning the real nature of this characteristic. It may reflect on general financial responsibility and attitudes. Even though the exact nature of the variable is unknown, it aids in prediction and it would be relatively easy to check this factor when considering prospective borrowers. For this sample the data told only whether or not the borrower was covered by insurance. The PCA data allowed tests of the affect of holding insurance as well as the affect of the size of the policies. In the PCA functions neither seemed important. As was suggested earlier this may be due to the fact that life insurance is available to borrowers at PCAs.

Health Insurance--One variable which was expected to have a positive coefficient but did not was health insurance. This deserves an explanation. Three possibilities are offered other than chance. First the data were not as good in the case of this variable as for most other variables. It was sometimes necessary to refer in the lenders

files to data collected after the original loan to determine whether the family had health insurance. This leaves the data on health insurance open to question. In addition while data were being collected at the FHA office in this county, the office stenographer stated that she often advised the uninsured farm families to obtain health insurance. If by chance she saw a greater need for insurance among those with lower incomes, she may have induced some of them to purchase insurance. This could make it appear as if more low income (and high risk) borrowers had health insurance than would otherwise be true. A third possible explanation might be that people with poor health would tend to purchase health insurance. If this were true, poor health could be the contributing cause of failure and the variable merely a reflection of this fact.

Family size--In this sample in contrast to the results of the PCA sample it appeared that family size was related to risk. Larger families seemed to have more difficulties in debt repayment. Possibly the generally low incomes of this group of borrowers made family support more difficult. Then too, given the close relationship between these agricultural firms and households, farmers with low incomes may be forced to consume present income and curtail investments. The fact that family size appeared to be unimportant in the PCA sample even though families were slightly larger suggests that PCA borrowers are receiving incomes somewhat above the minimum required for their

families. The PCA borrowers also were generally in a stronger financial position as shown by their large asset holdings.

Explanation models might show even stronger relationships between family size and debt repayment. Many of the families no doubt became larger during the period before the loans were repaid, The information on family size used in this study is limited to that available when the loan application was made.

The Ratio of Non-Real Estate Debts to Total Debts--This variable used the ratio of non-real estate debts to total debts to determine the effect a high proportion of short term commitments had on the ability to make satisfactory repayments.

As expected the coefficient for this variable was negative. It appears that large non-real estate debts have some adverse effects on debt repayment. Many lenders as a matter of policy attempt to correct heavy concentrations of short term debts by refinancing to give borrowers more realistic payments. This policy appears to be justified and logically would include attempts by the lender to prevent the borrower from again obtaining excessive short term debts. Lenders might give further consideration to use of additional debt consolidation measures.

Net Worth Increases--A variable which appears to aid in predicting risk was one which shows the ratio of net worth to productive years of work. This variable was

expected to have a positive coefficient which would indicate that individuals who have in the past been able to add to their net worth from year to year would be preferred risks. The results of the investigation of the PCA data contradicted this hypothesis. The explanation offered was that these gains in net worth were made by the borrower assuming heavy risks. It was suggested that the PCA may discriminate against these risk takers.

In this sample the sign of the coefficient was positive. This may show that the FHA supervisor was less concerned with risk taking and more interested in gains in net worth.

Level of Living--A variable, for which data were not available and which was therefore not examined for the PCA sample, deals with the household conveniences available on the farm. This factor appears to add to the ability of the function to predict loan outcomes. This particular variable is based on whether or not running water was available in the house at the time of the loan. Recognition of this as a variable reflects the importance of past income. If previous farm income was not sufficient to provide such bare essentials, it appears likely that the same farm will not provide high incomes in the future. This factor may also reflect on the level of motivation of the particular farm family. It may be that farm families who borrow from the FHA succeed or fail largely on the strength of their desires for improved living conditions.

The use of the variable concerned with running water in the household is probably no longer appropriate today since nearly all farms have running water. The general notion may be useful. Farms which appear to have lacked the capacity to provide adequate living standards in the past might realistically be questioned as economic units for the future.

The Ratio of Planned Debt Repayment to Net Cash Income--It was suggested earlier that ability to repay a loan may be the most important aspect of lending. A measure of the ability to pay back a loan should be closely related to net income. This variable was a ratio of expected debt repayment during the year following the loan to expected net cash income in the same year. The sign of the coefficient was negative as expected. This variable was important in some of the equations and gives an indication that potential income and expense estimates need serious consideration to determine whether the borrower will be able to meet other expenses and still have enough money to make scheduled debt payments.

Percent of Cropland in Corn--It was expected that farms which were capable of producing a relatively large proportion of corn in relation to total crop land would provide a less risky situation for the farmer. A positive coefficient for this variable would lend support to this hypothesis. Although the evidence was not strong the coefficient for this variable was positive.

Living Expenses--It was expected that those individuals with relatively high living expenses would have more difficulty living within their income and repaying their debts. There may be some reason to believe this since the sign of the coefficient was negative as expected. It may be that some of those with higher living expenses are living at a higher level and have goals which will require successful farm operation to attain. This may be a reason why this particular variable seemed relatively unimportant.

Farm Ownership and Experience--Variables related to these factors appeared to be quite important for the farms in the PCA sample. For this sample neither farm ownership nor experience seemed important. Data on experience was limited to the year before the loan application. Given data comparable to the PCA data concerning experience the findings may have been different.

Number of Sources of Credit--Unlike the results of the PCA analysis this factor seemed to be of little value as a risk predictor. The reason may be that the FHA supervisor provides management assistance. He may be able to discourage those borrowers who are inclined to obtain credit from many sources from doing so while he is supervising their FHA loan. He may also be providing credit to meet emergencies and unexpected expenditures which might otherwise be obtained from other lenders.

Debts and Assets--Variables related to debts and assets did not appear to contribute importantly to risk

discrimination. The expected signs were, however, obtained in each case.

Age--There were some slight indications that older farmers who obtained loans from this FHA office were poorer risks than were the younger borrowers. One explanation for this might be that older farmers would be expected to have assets which would be suitable as security for other lenders. The inability of these older borrowers to obtain loans elsewhere may indicate poor past financial management. For younger borrowers the inability to obtain loans elsewhere does not reflect on past financial management in the same way.

Choice of a Function and Borrower Classification

The function which was chosen to represent the Ingham County FHA data was function IIa. It was selected on the basis of several factors. First, R^2 was relatively high (.44) compared to many of the other functions. Those functions which had higher R^2 values also required an even larger number of variables. Deletion of variables from this function appeared to decrease R^2 more severely than in the other functions. Although function IIa has a fairly large number of variables, each of them appeared to add to the usefulness of the function; hence, they were not deleted.[3] Again it appears unlikely that naive or less

[3]It may be that under some circumstances risk prediction can be done accurately only if many variables are considered.

sophisticated models would be able to better discriminate between successful and unsuccessful borrowers.

It can be seen in Table 10 that this function, as with the others based on the Ingham County FHA data, was not as successful in classifying the borrowers as was the selected PCA function. Difficulty was experienced for both the successful and unsuccessful group. Not only were more of these borrowers misclassified, but the calculated "probability" values were further from the actual value.[4] For four of these farmers, the estimated values were above the mean estimated "probability" value for the successful class.

The successful borrowers who were misclassified had less serious errors. Three of the five had estimated "probability" values between .48 and .5 which might be interpreted as difficult classification decisions with a slight degree more risk than for the rest of the sample.

Using function IIa, the classification of the Ingham County FHA borrowers is as follows.

TABLE 10. Classification of the Ingham County FHA Sample According to Estimated "Probability" Values (Function IIa)

	Number of Borrowers				Total
Successful	0	5	7	16	28
Unsuccessful	16	6	2	4	28
"Probability of Success"	.29	.5	.69		

[4]The actual "probability" values were, of course, assumed to be 1 for the successful and 0 for the unsuccessful cases.

Since the lenders were asked to classify the borrowers into two dichotomous groups, the "actual probability" of success cannot be estimated. It is possible that these marginal cases were in fact marginal in the lenders' classification of successful borrowers. This would mean that the error was not as severe as it first appears. The error involved at the margin may, on the other hand, be considered to be important, since the marginal cases are those which present the greatest classification difficulties.

FHA Credit Rating Formula

Function IIa can be used to develop a credit rating formula just as was done with the PCA data. It should be recognized that this formula will probably be less accurate as indicated by the results obtained in classifying the sample. Other formulae could also be developed from similar functions.

The following weights were estimated for the variables in rating a particular borrower.

Family size--For each member of the family subtract .018.

Running water--For farms with running water add .501.

Life insurance--For farmers who have life insurance add .5478.

Health insurance--For families that have health insurance subtract .20.

The ratio of non-real estate debts to total debts--The largest possible value for this factor is 1. If the value recorded for an individual is 1, subtract .1. If it is less than 1, subtract a proportionally smaller value. For example, if the value is .5, subtract .05.

The ratio of planned debt repayment during the first year of the loan to expected net cash income in the same year--The largest possible value for this factor is 1. If the value recorded for an individual is 1, then .34 would be subtracted. When the value recorded is less than 1, subtract a proportionally smaller value. For example, if the value is .5, subtract .17.

Average annual gain in net worth--Add .057 for each $1,000 average annual increase in net worth after age 20.

Factor Substitution Rates

Substitution rates between the variables in the function are presented in Table 11. The form of the table is similar to that used in Table 7.

Analysis of the Eaton County FHA Data

The functions developed on the basis of the data from the Eaton County FHA sample appeared to predict fairly well yet are difficult to understand. Some of these functions are shown in Table 12. Several of the functions correctly classified 26 of the 28 cases.

TABLE 11. Rates of Substitution Between Variables for the Ingham County FHA (Function IIa)[a]

Units of Variables	Family Size	Running Water	Life Insurance	Health Insurance	An Increase in the Value of the Ratio Non-Real Estate Debt / Total Debt of .10	Debt Repayment / Net Cash Income	$1000 Average Annual Increase Net Worth
Family size	1	27.4	29.9	10.8	.53	1.86	.31
	+	−	−	+	+	−	−
Running water	.036	1	1.09	.40	.02	.067	.114
	−	+	+	−	−	−	+
Life Insurance	.033	.915	1	.362	.018	.062	.104
	−	+	+	−	−	−	+
Health Insurance	.092	2.53	2.76	1	.049	.171	.287
	+	−	−	+	+	−	−
An Increase in the Value of the Ratio Non-Real Estate Debt / Total Debt of .10	1.90	51.99	56.84	20.57	1	3.53	5.90
	+	−	−	+	+	−	−
Planned Debt Repayment / Net Cash Income	.54	14.74	16.12	5.83	.28	1	1.67
	+	−	−	+	+	−	−
$1000 Average Annual Increase Net Worth	.32	8.81	9.63	3.49	.17	.60	1
	−	+	+	−	−	−	

[a] Read across the table then up. For example, an increase in the value of the ratio non-real estate debt/total debt of .19 is approximately equivalent to an increase of one member in the family.

TABLE 12. Selected Discriminant Functions for Eaton County FHA Borrowers[a]

III_a $P = .85490 + \underset{(.22737)}{.50337}\ X_1 + \underset{(.000006)}{.000001}\ X_2 - \underset{(.60306)}{2.41136}\ X_3 - \underset{(.22076)}{.13814}\ X_4$

$+ \underset{(.39294)}{.24897}\ X_5 - \underset{(.00006)}{.00006}\ X_6$

$R^2 = .55$ (classified 26 of 28 correctly)

III_b $P = .89964 + \underset{(.22353)}{.49809}\ X_1 + \underset{(.17192)}{.06696}\ X_7 - \underset{(.16490)}{.05870}\ X_8 - \underset{(.62057)}{2.35624}\ X_3$

$- \underset{(.19607)}{.19074}\ X_4 + \underset{(.40548)}{.23356}\ X_5 - \underset{(.00006)}{.00005}\ X_6$

$R^2 = .56$ (classified 26 of 28 correctly)

III_c $P = .9548 + \underset{(.23208)}{.46812}\ X_1 - \underset{(.16087)}{.11575}\ X_8 - \underset{(.03876)}{.03242}\ X_9 - \underset{(.20965)}{.04471}\ X_{10} + \underset{(.00081)}{.000008}\ X_2$

$- \underset{(.63098)}{2.08328}\ X_3 - \underset{(.00005)}{.00007}\ X_6$

$R^2 = .56$ (classified 26 of 28 correctly)

III_d $P = .76682 + \underset{(.22073)}{.47844}\ X_1 + \underset{(.000005)}{.000003}\ X_2 - \underset{(.59324)}{2.38523}\ X_3 + \underset{(.38666)}{.2331}\ X_5 - \underset{(.00005)}{.00007}\ X_6$

$R^2 = .54$ (classified 26 of 28 correctly)

Each variable is described below. The variability coefficient for each coefficient is noted in parenthesis. R^2 is shown for each function.

P = An estimated value for the dependent variable. The estimated "probability" that a particular individual will be successful. Individuals for whom values of P are estimated to be greater than .5 are those who are expected to be successful. Those for whom values of P are estimated to be less than .5 are expected to be unsuccessful.

TABLE 12. (Continued)....

X_1 = One for borrowers with running water in their homes and zero for those without running water.

X_2 = The dollar value of the borrower's assets.

X_3 = The percent of the total cropland in corn or to be planted to corn.

X_4 = The percent of the total debts which are non-real estate debts.

X_5 = The percent of the expected net cash income which will be used for debt repayment in the first year.

X_6 = The average annual increase in net worth after the borrower is 20 years old.

X_7 = One for borrowers with health insurance and zero for those without health insurance.

X_8 = One for borrowers who operated the same farm the previous year and zero for those who did not operate the same farm the previous year.

X_9 = The number of sources of credit which were disclosed by the prospective borrower at the time of application.

X_{10}= A one for borrowers who own their farms and zero for renters.

[a] Additional functions are shown in the appendix (Table 5).

Percent of Cropland in Corn--Although it is difficult to explain, deletion of the variable for the ratio of corn to total crop acres decreased R^2 by the greatest amount. The coefficient for this variable was also negative whereas it was expected to be positive. A possible explanation for this negative sign may be that the farmers who were raising large corn acreages were doing so at the expense of their more profitable dairy enterprise.

It is possible that they were following a short-run profit maximization program which was not suitable on the farms in the area over a period of several years. A third possibility is that they were attempting to offset some known deficiency in their businesses.

Level of Living--The variable concerned with household conveniences as represented by running water in the household again appeared to add to the ability of the equation to predict loan outcomes and was found to be quite important. Again the conclusion is that insufficient past income from the farm is indicative of future low income with the resulting low standard of living.

Family Size--In the Ingham County FHA sample, the variable which was just discussed was supported by another variable concerned with the level of living. This variable was the size of the borrower's family. For this sample the analysis showed this variable not to be important for discriminating between the successful and unsuccessful borrowers. The sign of the coefficient for family size was

the same (negative) as for the Ingham County sample. Families were slightly smaller in the Eaton County sample than for the other FHA sample (3.9 and 4.0, respectively). Too, expected net cash incomes were on the average somewhat higher. The unimportance of the variable concerned with family size may thus be due to the fact that these farm families had higher incomes per person.

The Ratio of Planned Debt Repayment To Net Cash Income--As suggested by the results of the Ingham County FHA sample, the ability to repay debts is to some extent reflected by the relationship between income and plans for debt repayment in the year following the loan. For this sample, the variable was not particularly important although its sign was negative as expected.

The Ratio of Non-Real Estate Debts To Total Debts--Large non-real estate debts in relation to total debts appeared to contribute toward loan failure as was suggested by both of the other samples.

Assets--The variable for the level of assets was expected to have a positive relationship to success. The sign of this variable lends support to the hypothesis. These results differ from the results of the PCA equations where the sign of the coefficient was negative. A possible explanation is offered in the section concerning the results of the analysis of the PCA sample.

Net Worth Increases--The functions developed for the Ingham County FHA and the PCA samples differed in the sign of the variable which reflects the ability to make gains in net worth. For this sample the sign of the coefficient agrees with the sign obtained in the PCA function. This suggests a somewhat more conservative approach to lending and the desire of the lender to avoid those who are able to make net worth gains by being risk takers, that is, by borrowing heavily.

Health Insurance--Borrowers who had health insurance did not appear to be poorer risks in this sample as was found in the other FHA sample. This is more as would be expected logically. Although the coefficient was positive, the variable was relatively unimportant.

Farm Ownership and Experience--These variables appeared not to contribute importantly to the function. This compares favorably with the result of the other FHA sample but contrasts with the PCA analysis.

Number of Sources of Credit--Again in agreement with the Ingham County FHA sample analysis, the number of sources of credit appeared to be of little importance. This, it is suggested, is due to the management aid provided by the FHA county supervisor.

Life Insurance--Ownership of a life insurance policy did not appear to be important as it was for the Ingham County FHA sample. The sign of the coefficient was negative whereas it was expected to be positive.

Development of a Credit Rating Formula

The results of the analysis of this sample were in conflict with theoretical considerations and the justification of the results seems questionable. Consequently, the data will not be used to develop a formula for prediction as was done for each of the other samples.

Part of the doubt about the advisability of developing a formula stems from the relatively small number of borrowers considered in the analysis even though all borrowers who the lender was willing to classify as unsuccessful and an equal number of successful individuals were considered.

Some General Results

Several equations were fitted by least squares to the data from the three samples. The preliminary models had many variables. Some variables, as was expected, added little to the predictive ability of the function; thus they were generally not included in subsequent equations. Some of the variables in the functions were found not to be important in risk prediction but were found to increase R^2 and to aid in prediction. Such variables were considered useful and were not deleted.

Prediction models were developed independently for each of the samples used in the analysis since there were indications that the data from individual lenders were

structurally different. Basically, the form of these functions was much alike although the importance of the different variables did change. The procedure used was dictated in part by the theoretical consideration developed in earlier sections.

In most cases, the signs of the coefficients of the independent variables were the same as would have been expected from theoretical considerations. Some variables which were expected to be rather important were found to be of little use in predicting risks. This does not necessarily mean that information about these factors is unimportant to lenders, since it may be useful as a measure of the level of security.

Conclusions drawn for any one sample do not usually appear to apply to the others. It was expected that different factors would be important at least between the PCA and FHA loans. Differences between the FHA samples, although not entirely unexpected, were greater than was anticipated.

An attempt to fit a function to data from all observations for the three samples was unsuccessful in providing a prediction equation. The observations were classified into their respective successful and unsuccessful classes rather inaccurately. About 37 percent of the borrowers were misclassified. R^2 was only .27. This suggests that important differences exist between lenders concerning the variables under study and leads to the

conclusion that serious difficulty would result from attempting to predict loan outcomes for one class of lenders on the basis of information that includes data from another class.

When the data from the two FHA samples were combined, results were found to be similar to those obtained by combining PCA and FHA data. This suggests that these two FHA county supervisors were not using the same criteria to differentiate between good and poor risks and that they may have had different objectives for their lending programs. These differences probably led them to use dissimilar factors for approving loans and for classifying individuals in the successful or unsuccessful categories.

Tables 1 and 2 show that borrowers from the Ingham County FHA sample were generally not as well off financially as were borrowers from the Eaton County FHA sample. The PCA borrowers appeared to be considerably better off financially than the borrowers from either of the FHA samples.

Not all conclusions are unlike for the different lenders. It seems that lenders generally may have overemphasized the role of the level of debts and assets as risk predictors. None of the functions showed the level of debts, assets or a ratio of these to be particularly important for prediction given the kinds of data which were available. This finding arouses suspicion that lenders use the relationship of debts to assets as an

indication of their own position in the event of default rather than as an indicator of the risk of default. It needs to be noted that the variable concerned with the number of different debts did sometimes indicate that riskiness increases as the number of debts increases.

The analysis indicates that it may be easier to predict the outcome of loans for PCA borrowers than it is for FHA borrowers. This may result in classification and prediction difficulties. There is little question but that the classification of borrowers as "successful" or "unsuccessful" as done by the FHA and PCA are not the same. Although rules were suggested for classification which dealt largely with collection problems, it is possible that FHA supervisors are more tolerant of certain collection difficulties than other lenders. Since FHA supervisors supply considerable technical and managerial assistance which could minimize these difficulties, their tolerance may be justified.

The fact that the functions developed from the loan form data generally accounted for 65 percent or less of the variation suggests that PCA and FHA loan applications do not contain sufficient information to allow these lenders to predict accurately the outcome of their loans. Lenders do have access to some additional subjective information not covered by the loan forms. Discriminant analysis however, has the advantage of being more objective and can be more precise.

If it is assumed that the lenders would not have loaned to the borrowers classified as unsuccessful had they known about them at the time of the loan, it appears that the functions did relatively better than the lenders in classifying the borrowers in the samples. The functions generally misclassified some of the borrowers. The lenders appear to have misjudged one-half of the borrowers included in the samples. It should be realized, of course, that these unsuccessful cases are actually a small percentage of the total number of loans.

There is some evidence that lenders may have recognized that the unsuccessful cases were subject to greater risk, since the loan size was generally smaller for the unsuccessful group. There is also the possibility that the small loans obtained actually failed to meet the needs of these borrowers and thus contributed to their lack of success.

Since the variability coefficients were fairly large, it appears that predictions which lenders can make from their records are subject to considerable uncertainty.

Credit rating formulae, such as the ones described in this chapter, must be combined with the judgment of the lender. They may be less useful in agriculture than for other types of credit since so many factors change from farm to farm. It should be possible, however, given satisfactory data, to develop fairly effective formulae by obtaining appropriate data from farm firms

which have several important characteristics in common. Also, since the sample on which these formulae are based does not include many very high risk individuals, it may be useful for the lender to do some preliminary screening for individuals who are easily recognized high risk cases. It is likely that most of these would also be classified as high risk cases by the formula.

An Appraisal of the Technique

The results of this study were varied. Some of the findings confirmed those expected logically. Some were in contrast to the results expected and pointed toward unsuspected factors or could be otherwise explained. Other results were unclear.

Evaluation of the technique can be divided into the choice of the source of data and the choice of a method of analysis.

Source of Data--If farm loans are to be analyzed, the most logical places to turn for data are the offices of the agricultural lenders. It might be supposed that each lender would have detailed records on his borrowers. As was noted, it has been found by others that few of the lenders maintain complete records. Based on information from other sources, it appears likely that PCA's and the FHA are probably the most satisfactory source of data available.

Given the fact that the technique was untried for agricultural loans, it was felt that efforts limited to individual lenders would serve to eliminate the effect of the lender in a particular sample. This was later found to have serious consequences in the form of reduced sample size. The organization of agricultural lenders and their experience record in the post World War II period restricted the sample form from any one lender to a relatively small number of borrowers.

The variability coefficients obtained for the coefficients of the independent variables together with the lack of success in combining samples suggest that there is little possibility for improving the functions by including more lenders as long as data are limited to those found in present loan applications.

It appears that the real data problems encountered in this study result from the fact that lenders are collecting information for different purposes than has generally been thought. They appear in fact to be much more interested in security than in prediction. If this is so, it might be theorized that the failure is circular. Lenders are interested in security because of their inability to predict and those who attempt to predict find they are unable to do so satisfactorily because the information which is collected describes security.

Method of Analysis--The statistical technique was generally found to be a useful tool. It seemed to be

applicable to this kind of analysis. Even though much information is probably lacking, the function did surprisingly well in discriminating between the good and poor risk borrowers. Additional theoretical work in developing appropriate statistical tests should make the technique even more useful.

The Need for Further Research

Like most studies the conclusions which can be drawn from this work must be stated tentatively. Confirmation of the statements made in this report requires further research. Work with successful borrowers for whom information is now available could provide a further check on the results of this analysis. Additional evidence based on new data and by other individuals may strengthen or disprove the results. It is hoped that this study suggests to others a technique which offers promise but which has previously not been used in agricultural credit research.

An area in which further work is suggested is in improving the data for future loan analysis. One way to do this might be to work with lenders in designing loan forms for prediction of risks. This is suggested since the results of this study show that present loan applications do not contain some of the kinds of information needed. Some agricultural lenders have been improving their loan forms. However, these forms still tend to emphasize security and lack needed information on the ability, character

and attitudes of borrowers. Loan applications developed for use by lenders and for further work in risk discrimination need to be based more on economic theory than has been true in the past. Also the results of this study and the experience of lenders should suggest factors which need to be closely examined.

The final evaluation of a loan should be based on successful repayment, thus future work should stress factors which are related to income and debt repayment ability. Since a farmer's ability to produce income and act in a responsible manner is closely related to his ability as a manager, it should be useful in working with lenders to examine factors which are indicators of management.

After revised loan forms have been developed and used by lenders for a period of several years loan risk analysis and risk prediction might be considerably improved by again applying discriminant analysis to these problems. This would also aid in empirically testing the new loan forms. Success in such a study should provide information which would aid in risk discrimination and should also suggest additional improvements in loan applications.

Another possible technique for improving the results of discriminant analysis as used in studying risks could be the use of survey data. Persons who were known to have obtained a loan could be interviewed without

reference to the loan or its source. At some time, perhaps five years later, the borrowers could be classified as successful or unsuccessful by some objective method and discriminant analysis could be used as a method of analysis. This method could provide much more detailed data and a more objective measure of success or failure of the borrower than is currently available from loan applications and lenders. It would have the disadvantage of requiring a very large sample since the number of unsuccessful cases is relatively small compared to the total number of loans. A small sample might lack enough unsuccessful cases to allow use of the analytic procedure.

This model has been restricted to the analysis of loan cases based on information available at the time of the loan. Since some loans continue over several years, it may be more realistic, particularly for long term loans, to continue to collect information and make revised predictions at different times after the loan was approved. The type of investigation suggested might result in valuable indicators of how loan cases progress toward "success" or what events precede failure. Lenders currently make revised estimates of their various loans from time to time. Their decisions are in part based on this "new" information as it is obtained.

More work needs to be done to further isolate the factors which affect loan outcomes under different circumstances. This should include consideration of real estate loans. After such work appears to be fruitful,

attention should be focused on the costs of "lending mistakes." Determinations need to be made of the relative importance of the two types of errors in misclassification. That is, classifying an individual as a potentially successful borrower when, in fact, he is unsuccessful or in classifying an individual as potentially unsuccessful when he is successful.

Since some questions arise concerning the classification of borrowers into successful and unsuccessful classes, further analysis could limit unsuccessful cases to those which actually failed to repay their debts or some other more objective classification system. This may require aggregation of lenders from larger groups. Aggregation may also cause difficulties due to heterogeneous elements in the lending and classifying techniques of lenders even though they are from the same group or credit agency.

There are many reasons to think that soil productivity is an important factor in agricultural loan risk analysis. Inadequate attention has been given to soils in this study. It is suggested that real gains in loan risk prediction may result from a study designed particularly to analyze the effects of soil productivity. Farms should be selected so that the latest soil maps can be used. These are available now for a limited number of counties. This needs to be a joint project with the Department of Soils. Soil maps, farm boundaries, land use capability

and current land use could be obtained from the Soil Conservation Service.

Another area of suggested research is in studying the relationship between time preference and risk. If present consumption is extremely important to an individual or family when compared to future consumption they may be unwilling to make capital investments needed for efficient operation of their farm. This could result in decreased income and more risky loans. Too, they might use credit for home improvements and immediate consumption which seemed desirable but which was in excess of what their income could provide in the long run. Important relationships may exist particularly for low income and high risk borrowers. These are of course of real concern to lenders. Attempts to get ideas about time preference will probably need to depend on survey data or on cooperation with a group of lenders.

Summary

This chapter has been primarily concerned with presenting the results of the analysis. The findings are discussed in terms of each of the samples with the effect of the particular variables being described for each. Two credit rating formulae were developed.

Final sections in this chapter were also used to evaluate the technique and to suggest further areas for research.

CHAPTER VI

SUMMARY AND IMPLICATIONS

This study was conducted to accomplish the following objectives: (1) to evaluate the importance of various borrower characteristics in discriminating "successful" from "unsuccessful" loan applicants, (2) to develop a model which can aid in discriminating "successful" from "unsuccessful" loan applicants based on information available at the time the loan is under consideration and (3) to evaluate the effectiveness of present loan applications as sources of data for predicting the outcome of loans.

Changes within agriculture, expanded use of short and intermediate term credit and a need for improved loan arrangements suggest greater emphasis on risk forecasting and the use of more objective risk prediction techniques. Historically little has been done to improve risk prediction by the use of objective methods.

In initiating this study, factors which appeared to be logical indicators of risk were reviewed. These factors were classified as those arising due to non-firm influences and those originating within the firm. For this study the risk factors associated with the firm were given most of the emphasis.

Three offices of agricultural lenders provided data. The lenders were asked to select a dichotomous sample of "successful" and "unsuccessful" borrowers. The Farmers'

Home Administration and a Production Credit Association were used as data sources since these lenders have more complete information on their borrowers than other agricultural lenders.

Discriminant analysis was chosen as the method for analyzing the data. The function was of the form $P = a + b_1x_1 + b_2x_2 + ... + b_nx_n + \alpha$ where the dependent variable is assigned a value of one or zero according to whether the borrower was classified by the lender as "successful" or "unsuccessful." The function was solved by least squares methods. The estimated values for the dependent variable gave an indication of the "probability" of successful repayment for each borrower. Borrowers for whom the estimated value of the dependent variable was near one were expected to succeed. Those near zero were predicted to be unsuccessful. The actual discriminating value was set at one-half.

The equations which were selected as being most useful were those which could accurately predict loan outcomes with relatively few variables and had rather high values for R^2. The importance of the variables was indicated by their coefficients. Substitution rates between the variables in the equations were given in tabular form to make comparisons easier. Loan risk formulae were developed to aid in applying the results of the study to actual loans.

It was found that the PCA chattel loan candidates were more likely to be successful if they owned their farms and had lived on the particular farm for several years at the time of their loan application.[1] Farmers with larger operations were also found to be more likely to succeed. The relationship between success and the level of debts and assets did not seem particularly useful as risk predictors but the number of sources of credit was related to the outcome of loans. Although it was less pronounced than some other factors, the relationship between non-real estate debts and total debts appeared to be associated with risk. Relatively large non-real estate debts indicated greater risk.

Surprisingly PCA borrowers who were able to make the greatest average annual gains in net worth were less likely to be later classified as successful by the lender. This may be because these individuals were able to make their net worth gains by taking risks and the lender tended to discriminate against risk takers.

Analysis of the Ingham County FHA sample produced evidence that the relationship between the firm and the household needs to be given more consideration for these borrowers. Indicators of level of living seemed important

---

[1]Ownership as defined does not preclude the existence of real estate debts.

probably because of the relatively low incomes of these farmers in relation to family size. There was also an indication that attitudes toward insurance, risk and planning for the future may be important. It may be that FHA supervisors could understand the level of risk in their loans better if more attention were given to the goals of their borrowers.

FHA borrowers in this sample appeared to be somewhat better risks when their non-real estate debts were relatively low in relation to their total debts. Borrowers who were committed to make large debt repayments in relation to their income were poorer risks. The ability to make annual increases in net worth prior to the loan seemed to be an indicator that the borrower would be more likely to succeed. Again factors such as assets and debts appeared to be of little value in risk prediction.

The Eaton County FHA sample again showed that the past level of living was an indicator of the potential future capacity of the farm to produce needed income. Again plans for large debt repayment in relation to income in the early years following the loan seemed to indicate higher risk. Also large non-real estate debts in relationship to total debts suggests greater risk. Factors related to insurance did not appear as important in the case of this sample as with the other FHA sample. Rather highly intensified cropping programs were found to be related to higher risks for this sample of dairy farmers.

Again variables related to the level of assets and debts did not seem particularly important in predicting risks.

The study showed that it is possible to develop a function which will aid in discriminating between "good" and "poor" risks on the basis of loan application data. It was successful in indicating factors which are evidently important in predicting risk. It also furnished evidence that present loan applications do not contain all of the information needed for accurate risk prediction.

There are indications that lenders would be more successful in predicting loan outcomes if they found factors which show the ability of the borrower to repay debts. If the factors which are indicators of debt repayment ability and character can be isolated and used to predict loan outcomes, then lenders can place less emphasis on security.

A serious problem encountered in this study was a lack of the kind of data which are useful for prediction. Loan forms are still based largely on security and much of this information is not of much value for risk prediction. Better data will probably need to come from improved loan applications which are oriented more toward risk prediction and less toward security.

This study suggests a need for further work in risk prediction. Since character and the ability to repay a loan are closely related to the farmer's management ability, future research may be more fruitful if criteria can be

found which indicate good or poor management. These criteria might then be related to good or poor credit risks. Loan forms could first be improved; then additional research could be based on these improved data.

Other areas associated with risk prediction are suggested as potential research topics. Survey data may be useful in some cases. This could include studies of the relation of time preference and borrower motivation to risk. More objective methods of classifying the borrowers in the sample might be devised. Also risk analysis could be considered a continuing analysis of a lending situation rather than a one time estimate of risk. The study of the relationship between soils and risk may also provide an opportunity for valuable contributions.

Suggested Loan Form Improvements

Several improvements in loan forms have been suggested. Further emphasis on the nature of these improvements may be helpful.

Loan application data can serve to both predict risks and indicate security. Present loan applications seem to be well designed to do the latter. Despite improvements which have been made in loan forms, these forms appear to need further improvement to allow accurate risk prediction.

Loan forms need to give information on both volumn and efficiency of the farm business. Size of business can

be obtained by determining factors such as gross income, the number of tillable acres, acres of the various crops and the number of animals. Some important aspects of efficiency can be shown by gross income per man, gross income per acre and acres per man. Other measures which indicate the kind of job being done are yield per acre, fertilizer expense per acre, milk produced per cow, pigs weaned per litter, pounds of feed per pound of gain and income per $100 of cash expense. Lenders should find it helpful to consider these and similar measures of efficiency.

It should be useful to ask the borrower about his family size, level of living and expected family living expenses. If the borrower cannot provide enough income to meet family living, farm expenses and planned debts repayment, the loan should not be granted.

In general, the lender needs to find out: (1) Is the size of business adequate for a good living and for supporting the proposed debt? (2) Is the operator capable enough to get the kind of return on his labor, capital, land and animals to make the farm a paying proposition?

Appendix Table 1 - Discriminant Functions for PCA Borrowers[1]

$$I_f\ P = -.12923 + \underset{(.00767)}{.01771}\ X_1 - \underset{(.03099)}{.06061}\ X_2 + \underset{(.14728)}{.41088}\ X_3 + \underset{(.000005)}{.000001}\ X_9 + \underset{(.00057)}{.00131}\ X_5$$

$R^2 = .55$ (classified 34 of 40 correctly).

$$I_g\ P = .22246 + \underset{(.00770)}{.01590}\ X_1 - \underset{(.02934)}{.04430}\ X_2 + \underset{(.15094)}{.44513}\ X_3 + \underset{(.00052)}{.00203}\ X_5 - \underset{(.000007)}{.000007}\ X_{11} - \underset{(.15213)}{.03852}\ X_{12} - \underset{(.00001)}{.00002}\ X_7$$

$R^2 = .60$ (classified 35 of 40 correctly).

$$I_h\ P = .13994 + \underset{(.00738)}{.01750}\ X_1 - \underset{(.03164)}{.06002}\ X_2 + \underset{(.14942)}{.41081}\ X_3 + \underset{(.000005)}{.000001}\ X_9 - \underset{(.00001)}{.000002}\ X_{10} + \underset{(.00074)}{.00138}\ X_5$$

$R^2 = .55$ (classified 34 of 40 correctly)

$$I_i\ P = -.25608 + \underset{(.00932)}{.02914}\ X_1 - \underset{(.03107)}{.07612}\ X_2 + \underset{(.14225)}{.41745}\ X_3 - \underset{(.000003)}{.000007}\ X_4 + \underset{(.000007)}{.00001}\ X_9 + \underset{(.00001)}{.00002}\ X_{10} + \underset{(.00075)}{.00193}\ X_5$$

$R^2 = .61$ (classified 34 of 40 correctly).

$$I_j\ P = .11219 - \underset{(.00484)}{.00596}\ X_{16} + \underset{(.01120)}{.03670}\ X_1 - \underset{(.03401)}{.07839}\ X_2 + \underset{(.00001)}{.000002}\ X_{15} + \underset{(.17919)}{.45966}\ X_3 - \underset{(.000004)}{.000009}\ X_4 + \underset{(.00001)}{.00002}\ X_9 + \underset{(.00002)}{.00003}\ X_{10} + \underset{(.00078)}{.00178}\ X_5 - \underset{(.49656)}{.54826}\ X_{17} - \underset{(.31778)}{.11395}\ X_{18}$$

$R^2 = .64$ (classified 35 of 40 correctly)

[1]Several selected functions for PCA borrowers were shown on page 59 and are not repeated here.

$$I_k\ P = .25331 + \underset{(.00088)}{.00149}\ X_5 + \underset{(.30408)}{.42496}\ X_3 + \underset{(.01008)}{.01383}\ X_1 -$$

$$\underset{(.03363)}{.04693}\ X_2 - \underset{(.00002)}{.00002}\ X_7 - \underset{(.34766)}{.23193}\ X_8 - \underset{(.000004)}{.000002}\ X_4 +$$

$$\underset{(.00001)}{.00001}\ X_9 - \underset{(.68392)}{.98071}\ X_6 - \underset{(.00001)}{.000001}\ X_{15} + \underset{(.01066)}{.00160}\ X_{13} -$$

$$\underset{(.04250)}{.00806}\ X_{14}$$

$R^2 = .65$ (Classified 37 of 40 correctly)

$$I_l\ P = .19369 + \underset{(.00882)}{.00809}\ X_{13} - \underset{(.04211)}{.02012}\ X_{14} - \underset{(.00494)}{.00765}\ X_{16} +$$

$$\underset{(.01173)}{.03284}\ X_1 - \underset{(.03270)}{.07891}\ X_2 + \underset{(.000009)}{.000006}\ X_{20} + \underset{(.000009)}{.000008}\ X_{15} +$$

$$\underset{(.18280)}{.55425}\ X_3 - \underset{(.000005)}{.00001}\ X_4 + \underset{(.00001)}{.00002}\ X_9 + \underset{(.00002)}{.00003}\ X_{10} -$$

$$\underset{(.13305)}{.17623}\ X_{21} + \underset{(.00079)}{.00131}\ X_5 - \underset{(.54007)}{.23540}\ X_{19} - \underset{(.67918)}{1.43150}\ X_6$$

$R^2 = .71$ (Classified 37 of 40 correctly)

$$I_m\ P = .12999 + \underset{(.00966)}{.02265}\ X_{13} - \underset{(.05036)}{.02492}\ X_{14} + \underset{(.000007)}{.000008}\ X_{15} -$$

$$\underset{(.48240)}{.39503}\ X_6$$

$R^2 = .18$ (Classified 29 of 40 correctly)

Appendix Table 2 - Variables Used in the Discriminant Functions for the PCA Borrowers.

P = An estimated value for the dependent variable. The "probability" that a particular individual will be successful. Individuals for whom values of P are estimated to be greater than .5 are those who are expected to be successful. Those for whom values of P are estimated to be less than .5 are expected to be unsuccessful.

X_1= The number of years that the farmer has operated the present farm.

X_2= The number of sources of credit which were disclosed by the prospective borrower.

X_3= A one for borrowers who own their farms and zero for renters.

X_4= The dollar value of the borrowers assets.

X_5= The size of the farm as measured by total acres of land.

X_6= The ratio of debts to assets.

X_7= The average annual increase in net worth after the borrower is 20 years old.

X_8= The percent of the total debts which are non-real estate debts.

X_9= The dollar value of the borrower's debts.

X_{10}= The amount of the loan.

X_{11}= The amount of non-real estate debts.

X_{12}= A one for borrowers with life insurance and zero for those without life insurance.

X_{13}= The age of the borrower.

X_{14}= The size of the borrowers family.

X_{15}= The amount of life insurance held by the borrower.

X_{16}= The years of residence on the farm.

X_{17}= The ratio of total debt including the new loan to assets.

X_{18} = The ratio of total debt before the new loan was approved to the total debts after the new loan was approved.

X_{19} = The ratio of loan size to the value of the borrowers assets.

X_{20} = The value of the borrower's cattle.

X_{21} = A one for new debts and a zero for loans intended primarily for refinancing.

Appendix Table 3 - Discriminant Functions for Ingham County FHA Borrowers[1]

$$II_f\ P = .19264 - \underset{(.03370)}{.05343}\ X_1 + \underset{(.18679)}{.27206}\ X_2 + \underset{(.13268)}{.55649}\ X_3 - \underset{(.00043)}{.00057}\ X_7 + \underset{(.55670)}{.49346}\ X_{12} - \underset{(.04382)}{.04859}\ X_{10} + \underset{(.58154)}{1.0093}\ X_9 - \underset{(.13447)}{.23326}\ X_4$$

$R^2 = .40$ (Classified 46 of 56 correctly)

$$II_g\ P = .82192 - \underset{(.00639)}{.00605}\ X_{13} - \underset{(.03488)}{.04780}\ X_1 + \underset{(.14155)}{.49082}\ X_3 - \underset{(.00042)}{.00073}\ X_7 - \underset{(.04457)}{.03393}\ X_{10}$$

$R^2 = .29$ (Classified 44 of 56 correctly)

$$II_h\ P = .33934 - \underset{(.03471)}{.04624}\ X_1 + \underset{(.18953)}{.30788}\ X_2 + \underset{(.13524)}{.50939}\ X_3 - \underset{(.00008)}{.00004}\ X_{11} + \underset{(.56442)}{.53059}\ X_9 - \underset{(.16157)}{.16977}\ X_5 - \underset{(.21195)}{.46887}\ X_6$$

$R^2 = .35$ (Classified 45 of 56 correctly)

$$II_i\ P = .27362 - \underset{(.03405)}{.04906}\ X_1 + \underset{(.18477)}{.28833}\ X_2 + \underset{(.13423)}{.51052}\ X_3 + \underset{(.55817)}{.50443}\ X_9 - \underset{(.15760)}{.15377}\ X_5 - \underset{(.20922)}{.45690}\ X_6$$

$R^2 = .35$ (Classified 40 of 56 correctly).

$$II_j\ P = .45351 - \underset{(.03418)}{.05041}\ X_1 + \underset{(.17809)}{.31524}\ X_2 + \underset{(.13301)}{.45822}\ X_3 - \underset{(.15907)}{.18941}\ X_5 - \underset{(.21098)}{.54085}\ X_6 + \underset{(.00003)}{.00006}\ X_{10} - \underset{(.27300)}{.02411}\ X_{14}$$

$R^2 = .38$ (Classified 46 of 56 correctly)

$$II_k\ P = .54763 - \underset{(.03318)}{.06253}\ X_1 + \underset{(.17155)}{.28450}\ X_2 + \underset{(.13148)}{.45626}\ X_3 - \underset{(.13245)}{.23145}\ X_4 - \underset{(.15796)}{.27885}\ X_5 - \underset{(.21087)}{.40657}\ X_6 + \underset{(.00003)}{.00006}\ X_8 -$$

$$\underset{(.26346)}{.09855}\ X_{14} + \underset{(.00073)}{.00073}\ X_{15}$$

$R^2 = .46$ (Classified 46 of 56 correctly)

$$II_L\ P = .21867 - \underset{(.03359)}{.06343}\ X_1 + \underset{(.18344)}{.26302}\ X_2 + \underset{(.12942)}{.57298}\ X_3 -$$

$$\underset{(.13030)}{.28021}\ X_4 - \underset{(.00008)}{.00003}\ X_{11} + \underset{(.54779)}{.45949}\ X_{12} - \underset{(.04357)}{.06190}\ X_{10}$$

$$\underset{(.56365)}{1.05124}\ X_9 - \underset{(.16179)}{.31670}\ X_5 - \underset{(.21172)}{.33261}\ X_6$$

$R^2 = .46$ (Correctly classified 47 of 56.)

$$II_M\ P = .47266 - \underset{(.03207)}{.06378}\ X_1 + \underset{(.17216)}{.24744}\ X_2 + \underset{(.12749)}{.51380}\ X_3 -$$

$$\underset{(.12376)}{.31952}\ X_4 + \underset{(.54001)}{.73965}\ X_9 - \underset{(.15268)}{.28780}\ X_5 - \underset{(.20842)}{.37221}\ X_6 +$$

$$\underset{(.00003)}{.00006}\ X_8$$

$R^2 = .46$ (Classified 44 of 56 correctly)

$$II_n\ P = .68053 - \underset{(.03318)}{.06247}\ X_1 + \underset{(.17080)}{.30040}\ X_2 + \underset{(.12805)}{.48601}\ X_3 -$$

$$\underset{(.12263)}{.28128}\ X_4 - \underset{(.15733)}{.28355}\ X_5 - \underset{(.21084)}{.40338}\ X_6 + \underset{(.00003)}{.00006}\ X_{10} -$$

$$\underset{(.26341)}{.09377}\ X_{14}$$

$R^2 = .44$ (Classified 46 of 56 correctly.)

$$II_o\ P = .75749 - \underset{(.03539)}{.05658}\ X_1 - \underset{(.18242)}{.00789}\ X_{16} - \underset{(.15258)}{.08349}\ X_{17} -$$

$$\underset{(.02744)}{.01756}\ X_{18} + \underset{(.00003)}{.00006}\ X_8 - \underset{(.21245)}{.39001}\ X_5 + \underset{(.00003)}{.00001}\ X_{19}$$

$$\underset{(.00001)}{.00001}\ X_{20} - \underset{(.46088)}{.28569}\ X_{21} + \underset{(.14918)}{.50764}\ X_3 + \underset{(.21270)}{.20024}\ X_2 -$$

$$\underset{(.13836)}{.33479}\ X_4 + \underset{(.00004)}{.00002}\ X_{22} - \underset{(.25458)}{.31272}\ X_6 + \underset{(.61116)}{.64056}\ X_9$$

$R^2 = .50$ (Classified 46 of 56 correctly)

$$II_p\ P = .47883 - \underset{(.00754)}{.00987}\ X_{13} + \underset{(.14092)}{.00317}\ X_{17} + \underset{(.16536)}{.06244}\ X_{16} +$$

$$\underset{(.14433)}{.45131}\ X_3$$

$R^2 = .23$ (Classified 40 of 56 correctly)

$$II_Q\ P = .73204 - \underset{(.00661)}{.00450}\ X_{13} - \underset{(.03525)}{.03926}\ X_1 + \underset{(.13869)}{.48598}\ X_3 -$$

$$\underset{(.00043)}{.00057}\ X_7 - \underset{(.02474)}{.02810}\ X_{18}$$

$R^2 = .30$ (Classified 44 of 56 correctly)

$$II_r\ P = .36736 - \underset{(.03971)}{.04689}\ X_1 - \underset{(.02375)}{.04743}\ X_{18} + \underset{(.13472)}{.49394}\ X_3 +$$

$$\underset{(.42393)}{.29086}\ X_{12} - \underset{(.06324)}{.06680}\ X_{23}$$

$R^2 = .30$ (Classified 41 of 56 correctly)

[1] Several selected functions for Ingham County FHA borrowers were shown on page 69 and are not repeated here.

Appendix Table 4 - Variables Used in the Discriminant Functions for the Ingham County FHA Borrowers

P = An estimated value for the dependent variable. The "probability" that a particular individual will be successful. Individuals for whom values of P are estimated to be greater than .5 are those who are expected to be successful. Those for whom values of P are estimated to be less than .5 are expected to be unsuccessful.

X_1 = The size of the borrower's family.

X_2 = One for borrowers with running water in their homes and zero for those without running water.

X_3 = One for individuals who have life insurance and zero for those without such insurance.

X_4 = One for borrowers with health insurance and zero for those without health insurance.

X_5 = The percent of the total debts which are non-real estate debts.

X_6 = The percent of the expected net cash income which will be used for debt repayment in the first year.

X_7 = The percent of the expected gross cash income which will be used for debt repayment in the first year.

X_8 = The average annual increase in net worth per year of productive work.

X_9 = The percent of the total cropland in corn or to be planted to corn.

X_{10} = Planned fertilizer expenditure per crop acre in dollars.

X_{11} = The borrower's expected living expenses for the year following the loan.

X_{12} = The percent of the total acres in crops.

X_{13} = The borrowers age.

X_{14} = The ratio of expected living expenses to expected cash income in the year following the loan.

X_{15} = Total acres of land.

X_{16} = A one for borrowers who own their farms and zero for renters.

X_{17} = A one for borrowers who operated the same farm the previous year and zero for those who did not.

X_{18} = The number of sources of credit which were disclosed by the borrowers.

X_{19} = The dollar value of the borrower's debts.

X_{20} = The dollar value of the borrower's assets.

X_{21} = The ratio of total debts to total assets.

X_{22} = The borrower's expected net cash income for the year following the loan.

X_{23} = The ratio of loan size to total assets.

Appendix Table 5 - Discriminant Functions for Eaton County FHA Borrowers[a]

$$III_e\ P = 1.27795 + \underset{(.23838)}{.41443}\ X_1 + \underset{(.19493)}{.01380}\ X_7 + \underset{(.04665)}{.01729}\ X_9 -$$

$$\underset{(.42266)}{.45338}\ X_{10} + \underset{(.000008)}{.000004}\ X_2 - \underset{(.71420)}{2.29671}\ X_3 - \underset{(.47106)}{.45893}\ X_4 -$$

$$\underset{(.32961)}{.52182}\ X_5 - \underset{(.00007)}{.00011}\ X_6 - \underset{(.32507)}{.08225}\ X_{11} + \underset{(.58187)}{.94866}\ X_{12}$$

$R^2 = .67$ (classified 25 of 28 correctly)

$$III_f\ P = 1.46932 + \underset{(.22606)}{.49965}\ X_1 - \underset{(.52212)}{.60141}\ X_{10} - \underset{(.63066)}{2.58566}\ X_3 -$$

$$\underset{(.69333)}{.77192}\ X_4 + \underset{(.40247)}{.13886}\ X_5 - \underset{(.00005)}{.00004}\ X_6 + \underset{(.61038)}{.47250}\ X_{12}$$

$R^2 = .59$ (classified 26 of 28 correctly)

$$III_g\ P = .49737 - \underset{(.05258)}{.08579}\ X_9 - \underset{(.33787)}{.29797}\ X_{10} + \underset{(.00039)}{.00031}\ X_{13} +$$

$$\underset{(.00001)}{.000007}\ X_2 - \underset{(.06282)}{.00445}\ X_{14} - \underset{(.41039)}{.44418}\ X_{11} + \underset{(.68604)}{.27137}\ X_{12}$$

$R^2 = .21$ (classified 21 of 28 correctly)

$$III_h\ P = .54310 + \underset{(.23853)}{.34182}\ X_7 + \underset{(.00029)}{.00006}\ X_{13} - \underset{(.06130)}{.00069}\ X_{14} -$$

$$\underset{(.32886)}{.12089}\ X_4 - \underset{(.32505)}{.42220}\ X_5 - \underset{(.38097)}{.36345}\ X_{11} + \underset{(.66241)}{.33577}\ X_{12}$$

$R^2 = .23$ (classified 19 of 28 correctly)

$$III_i\ P = .93810 + \underset{(.22285)}{.47726}\ X_1 - \underset{(.15701)}{.11731}\ X_8 - \underset{(.03528)}{.02942}\ X_9 +$$

$$\underset{(.000006)}{.000007}\ X_2 - \underset{(.60555)}{2.10851}\ X_3 - \underset{(.00005)}{.00007}\ X_6$$

$R^2 = .56$ (classified 25 of 28 correctly)

$$III_j\ P = .83606 + \underset{(.22630)}{.50116}\ X_1 + \underset{(.16425)}{.08031}\ X_7 - \underset{(.60563)}{2.37368}\ X_3 -$$

$$\underset{(.19069)}{.18275}\ X_4 + \underset{(.38514)}{.26852}\ X_5 - \underset{(.00005)}{.00005}\ X_6$$

$R^2 = .55$ (classified 25 of 28 correctly)

$$III_k \quad P = .87409 + \underset{(.21237)}{.51738}\ X_1 - \underset{(.58560)}{2.4263}\ X_3 - \underset{(.18290)}{.16252}\ X_4 + \underset{(.37828)}{.26337}\ X_5 - \underset{(.00005)}{.00006}\ X_6$$

$R^2 = .55$ (classified 26 of 28 correctly)

$$III_L \quad P = .93737 + \underset{(.20435)}{.55136}\ X_1 - \underset{(.56138)}{2.32649}\ X_3 - \underset{(.18073)}{.16698}\ X_4 - \underset{(.00005)}{.00004}\ X_6$$

$R^2 = .54$ (classified 26 of 28 correctly)

$$III_m \quad P = .84347 + \underset{(.20371)}{.54870}\ X_1 - \underset{(.55959)}{2.31815}\ X_3 - \underset{(.00005)}{.00005}\ X_6$$

$R^2 = .52$ (classified 26 of 28 correctly)

$$III_n \quad P = 1.25890 - \underset{(.61567)}{2.38167}\ X_3$$

$R^2 = .37$ (classified 23 of 28 correctly)

$$III_o \quad P = 1.51356 + \underset{(.09361)}{.06959}\ X_{15} - \underset{(.92397)}{1.04231}\ X_{10} - \underset{(.27874)}{.09780}\ X_8 + \underset{(.05126)}{.01127}\ X_9 - \underset{(.00008)}{.00005}\ X_6 - \underset{(1.09697)}{1.03064}\ X_4 - \underset{(.00007)}{.000005}\ X_{16} + \underset{(.00003)}{.000008}\ X_2 + \underset{(1.42823)}{.77051}\ X_{12} + \underset{(.47298)}{.08277}\ X_{17} + \underset{(.28686)}{.39858}\ X_1 - \underset{(.29143)}{.04320}\ X_7 + \underset{(.61169)}{.03707}\ X_5 - \underset{(.87819)}{2.43774}\ X_3$$

$R^2 = .64$ (classified 26 of 28 correctly)

[a] Several selected functions for Eaton County FHA borrowers were shown on page 82 and are not repeated here.

Appendix Table 6 - Variables used in the Discriminant Functions for the Eaton County FHA borrowers.

P = An estimated value for the dependent variable. The estimated "probability" that a particular individual will be successful. Individuals for whom values of P are estimated to be greater than .5 are those who are expected to be successful. Those for whom values of P are estimated to be less than .5 are expected to be unsuccessful.

X_1= One for borrowers with running water in their homes and zero for those without running water.

X_2= The dollar value of the borrower's assets.

X_3= The percent of the total cropland in corn or to be planted to corn.

X_4= The percent of the total debts which are non-real estate debts.

X_5= The percent of the expected net cash income which will be used for debt repayment in the first year.

X_6= The average annual increase in net worth after the borrower is 20 years old.

X_7= One for borrowers with health insurance and zero for those without health insurance.

X_8= One for borrowers who operated the same farm the previous year and zero for those who did not operate the same farm the previous year.

X_9= The number of sources of credit which were disclosed by the prospective borrower at the time of application.

X_{10}= A one for borrowers who own their farms and zero for renters.

X_{11}= The ratio of the size of the loan to the farmer's assets.

X_{12}= The ratio of debts to assets.

X_{13}= Expected living expense in the year following the loan.

X_{14}= Planned fertilizer expenditure per crop acre in dollars.

X_{15}= The size of the borrowers family.

X_{16}= The dollar value of the borrower's debts.

X_{17}= One for borrowers with life insurance and zero for those without life insurance.

Appendix Table 7 - A Comparison of Mean Value for Characteristics of the PCA and FHA Samples

Characteristic	Lender		
	PCA	FHA (Ingham)	FHA (Eaton)
Age	35.5	33.8	31.3
Family size	4.5	4.0	3.9
% Owners	75	36	54
% Holding life ins.	82	73	82
Amount of life ins.	7,604	---	---
% holding health ins.	---	54	39
Years operated farm	8.4	---	---
% operated farm previous yr.	---	64	57
% households with running water	---	87	86
Expected living expense	---	1,841	1,697
Size of original loan	7,274	5,266	4,570
Net worth	47,882	8,903	12,607
Assets	66,920	14,737	22,280
Debts	19,038	5,834	9,673
Average annual increase in net worth after age 20	4,030	1,063	1,652
Non-real estate debt	6,437	2,702	3,656
Number of debts	3.5	4.3	5.2
Expected gross income	---	6,900	7,822
Expected cash operating exp.	---	2,817	3,581
Planned debt repayment first yr.	---	2,025	2,535
Debt-asset ratio	.28	.33	.38
% Debts that are non-real estate	48	77	58
% of net cash income going to debt repayment	---	49	56
Total acres land	279	195	204
Total crop land	---	151	155
% Total crop landin corn	---	28	32

BIBLIOGRAPHY

Ackerman, Joseph and Norton, L.J. Factors Affecting Success of Farm Loans, University of Illinois, Agricultural Experiment Station, Bulletin 486, August 1940.

Anderson, T. W. An Introduction to Multivariate Statistical Analysis, Wiley, 1958, pp. 108-09 and pp. 126-42.

Barnard, M. M. "The Secular Variations of Skull Characteristics in Four Series of Egyptian Skulls," Annals of Eugenics, 1935, p. 352.

Baum, E. L., Diesslin, Howard G., and Heady, Earl O. (ed), Capital and Credit Needs in a Changing Agriculture, The Iowa State University Press, Ames, 1961.

Blood, Dwight M. "Discriminant Analysis and Farm Management Research," Management Strategies in Great Plains Farming, Great Plains Council Publication Number 19, University of Nebraska, Lincoln, August 1961.

Blood, Dwight M. and Baker, C. B. "Some Problems of Linear Discrimination," Journal of Farm Economics, Vol. XL, August 1958.

Bostwick, Don, Esray, James, and Rodewald, Gordon. Attitudinal Research Relating to Farmers' Use of Short-Term Credit, ERS-25, Farm Economics Division, United States Department of Agriculture, 1961.

Brake, J. R., and Holm, C. R. "The Influence of Household Size and Income on Farm Family Expenditures in Michigan, 1960," Quarterly Bulletin, Michigan State University, Agricultural Experiment Station, East Lansing, Vol. 44, February 1962, p. 541.

Chapman, John M., and associates. Commercial Banks and Consumer Installment Credit, Financial Research Program, National Bureau of Economic Research, New York, 1940, p. 134.

Diesslin, Howard G. "Agricultural Equipment Financing," National Bureau of Economic Research, Occasional Paper number 50, 1955, Chapter 6.

_____. "A Re-examination of the Credit Needs of Agriculture," Journal of Farm Economics, Proceedings Issue, Vol. XL, December 1958, p. 1200.

Durand, David. Risk Elements in Consumer Installment Financing, Financial Research Program, Studies in Consumer Installment Financing 8, National Bureau of Economic Research, New York, 1941, p. 125.

Eckert, Phil S., and Maughan, Orlo H. Farm Mortgage Loan Experience in Central Montana, Montana State College Agriculture Experiment Station, Bulletin 372, June 1939.

Economics Research Service, United States Department of Agriculture, The Balance Sheet of Agriculture, 1962, Agriculture Information Bulletin No. 270, pp. 18-22.

Fisher, R. A. "The Use of Multiple Measurements in Taxonomic Problems," Annals of Eugenics, Vol. 7, 1936, p. 179.

_____. "Statistical Utilization of Multiple Measurements," Annals of Eugenics, Vol. 7, 1938, p. 376.

Hill, F. F. An Analysis of the Loaning Operations of the Federal Land Bank of Springfield from its Organization in March 1917 to May 31, 1929, Cornell University Agricultural Experiment Station, Bulletin 549, December 1932.

Jamba, Nicholas A. "Discussion: Intermediate-Term Credit in Agriculture," Journal of Farm Economics, Proceedings Issue, Vol. XL, December 1958, p. 1140.

Johnson, Palmer O. Statistical Methods in Research, Prentice-Hall, New York, 1949, p. 347.

Jones, Lawrence A., and Durand, David. Mortgage Lending Experience in Agriculture, National Bureau of Economic Research, New York, 1954.

Merchant, Charles H. Farm Credit in Aroostook County, Maine, University of Maine Agriculture Experiment Station Bulletin 418, April 1943.

Morrow, Stanley A. "Intermediate-Term Credit in Agriculture," Journal of Farm Economics, Proceedings Issue, Vol. XL, December 1958, p. 1131-1140.

Straus, Murray A. "Matching Farms and Families in the Columbia Basin Project." Bulletin 588, Washington Agricultural Experiment Stations, State College of Washington. June, 1958.

Warren, Stanley W. Result of Farm Mortgage Financing in Eleven Counties in New York State, Cornell University Agriculture Experiment Station, Bulletin 726, December 1939.

Williams, E. J. Regression Analysis, Wiley, New York, 1959, p. 175.