APPLICATION OF MATHEMATICAL PROGRAMMING TECHINIQUES IN CREDIT SCORING OF AGRICULTURAL LOANS

Houshmand A. Ziari, David J. Leatham, and Calum G. Turvey

Proceedings of a Seminar sponsored by
North Central Regional Project NC-207
"Regulatory, Efficiency and Management Issues Affecting Rural Financial Markets"
Hyatt-Regency Crystal City
October 3, 1994

University of Illinois
Department of Agricultural Economics
Urbana, IL 61801

April 1995

Application of Mathematical Programming Techniques in Credit Scoring of Agricultural Loans

Houshmand A. Ziari, David J. Leatham, and Calum G. Turvey[1]

Introduction

Discriminant analysis (DA) or classification methods are used to classify an individual or object, based on a set of discriminatory variables or attributes, into one of a number of mutually exclusive groups. DA has emerged as an important decision making tool in many fields. DA is extensively used in business, biology, the social sciences and other areas that require classification processes. The methods has been widely applied in business fields such as: credit scoring (Srinivasan and Kim; and Turvey), bankruptcy assessment (Mahmood and Lawrence), for prediction of various events including credit card usage and tender offer outcomes, and personal evaluation or selecting employees (Eisenbeis).

Historically, statistical DA methods have been the standard to deal with classification problems. In recent years, many researchers have expressed concern about certain features of statistical DA models. In particular, statistical DA methods require restrictive assumptions of distributional form. For example, Fisher's Linear Discriminant Analysis model which perhaps is the most widely used DA method, requires assumption of multi-variate normal populations with the same variance/covariance structure. Unfortunately, violations of these assumptions occur regularly. Esinbeis argues that deviations from the normality assumptions, at least in economics and finance, more likely are the rule rather than the exception (p. 875). For example, the financial ratios normally used in credit scoring are rarely normally distributed. In addition, most empirical data include qualitative variables that cannot be multivariate normal (Goldstein and Dillon). The performance of statistical DA models, when underlying parametric assumptions are violated, are discussed by Baladrishnan and Subrahmanian; Lachenbruch, Sneeinger, and Revo; and Press and Wilson.

The statistical DA models also assume that misclassification costs are the same for all groups (Type I and Type II errors have equal significance). For example, the cost of turning down a good loan (Type I error) and the cost of accepting a bad loan (Type II error) are assumed to be the same. Furthermore, statistical DA models are not apt to adequately handle a complex discrimination problem. In certain situations, a side constraint might be necessary which would prohibit the use of statistical DA models. The aforementioned shortcomings of statistical DA models have prompted researchers to the development of several nonparametric DA techniques such as neural network, mathematical programming, and search methods. This paper focuses exclusively on the mathematical programming (MP) DA techniques.

---

[1]Houshmand A. Ziari is a visiting assistant professor and David J. Leatham is an associate professor of agricultural economics at Texas A&M University, College Station, Texas. Calum G. Turvey is an associate professor of agricultural economics and business, University of Guelph, Ontario, Canada.

In recent years, considerable theoretical research has been devoted to the use of MP techniques to the classification problem. Hand; Freed and Glover (1981a, b) were the first to introduce the use of MP in DA. Glover, Keene and Duea argue that the MP approach to DA offers certain advantages over the statistical DA models. These include:

- MP methods are free from underlying parametric assumptions;
- Varied objectives and more complex problem formulation are easily accommodated;
- Varied misclassification costs can be easily incorporated into the model;
- Some MP methods, especially Linear Programming, lend themselves to sensitivity analysis; and
- MP are less sensitive to outliers since the model is based on the L1 metric rather than the L2 metric.

In several experiments utilizing Monte Carlo simulation data, researchers have found that some of MP techniques rival or outperform the statistical DA techniques in terms of the relative classification performance (Bajgier and Hill; Freed and Glover, 1986; Joachimthaler and Stam, 1988; and Rubin). This is specially true when the underlying assumptions of statistical DA models are not satisfied. In spite of these experimental results, there has not been an extensive study which compare the performance of alternative MP models using real-world data. Mahmood and Lawrence; and Srinivasan and Kim are the only researchers that have applied MP discriminant models to actual business data. But, the MP models they used was a rudimentary form of general MP models that have been found to perform poorly in practice (Bajgier and Hill; Markowski and Markowski, 1987). Moreover, neither author attempted to take advantage of inherent flexibility of MP models as stated above. The purpose of this paper is to compare alternative MP formulations in more detail and apply them to actual business data.

Specifically, the objective of the paper is to evaluate alternative MP techniques in credit scoring of agricultural loans using statistical DA models, namely Fisher's Linear Discriminant Analysis (FLDA) and Logit Discriminant Analysis (LDA), as a performance benchmark. The MP and statistical DA models are compared on the basis of classification ability on in-sample and hold-out sample data set. The paper is organized into four major sections. First, a two-group discrimination problem is discussed. This is followed by a brief discussion of statistical DA models. Next, we present MP discriminant models. Finally, we compare the classification performance of statistical and MP models.

Two-Group Discrimination Problem

The two-group discriminant problem deals with discrimination between two predefined groups and is the fundamental problem in DA. Two-group discriminant problem assumes that there are two well-defined populations, G_1 and G_2 (e.g., good loans vs bad loans), and it is possible to measure j discriminatory variables or attributes for each member of either population. The focus of DA is the determination of a numerical rule or discriminant function that allows the investigator to distinguish between two populations using the j attributes. A linear discriminant function can be expressed

$$Z_i = X_0 + B_{i1} X_1 + B_{i2} X_2 + \ldots\ldots B_{ij} X_j \qquad (1)$$

where, X_0 is a constant term; X_j is the weight assigned to variable j; B_{ij} is the value of the j^{th} variable for the i^{th} individual; and Z_i is the discriminate value for the i^{th} individual. For a cutoff or boundary value of b, the classification rule then becomes: if $Z_i \geq b$ then individual i is assigned to group G_1, otherwise individual i is assigned to group G_2. The cutoff value does not have to be the same for both groups. But, for simplicity, we assume the cutoff values for both groups are the same in this paper.

The goal of any DA model is to estimate parameters X and b so as to minimize the number of misclassifications for in-sample and/or hold-out sample data set. DA models are inherently different from each other according to their choice of criterion function and/or distribution assumption(s). In all DA models, X and b, are however determined from a set of observations for which their group membership is known.

Statistical Linear Discriminant Analysis

There exists an extensive body of literatures which discuses the statistical DA models. Interested readers are referred to Altman et al. and Maddala for a detailed discussion of statistical models in classification studies. For a discussion of credit scoring models and the theoretical consideration of credit scoring in agriculture, the interested reader is referred to Betubiza and Leatham; Miller and LaDue; Turvey; and Turvey and Brown. The statistical procedures of FLDA and LDA have been discussed extensively in the literature, and their detailed formulations are not repeated here.

Fisher Linear Discriminant Analysis

FLDA procedure computes the linear discriminant function (1) by maximizing the ratio of the between-group variance to the within-group variance. The derived linear discriminant function is known to be optimal in context of minimizing the total probability of misclassifications, provided the following conditions are held: (a) the distributions of the variable are multivariate normal, and (b) the variance-covariance of the variables are the same for both population groups (Johnson and Wichern). The coefficients for FLDA model are estimated by

$$\begin{aligned} X &= [((n_1-1)S_1+(n_2-1)S_2)/(n_1+n_2-2)]^{-1}(\mu_1-\mu_2) \\ X_0 &= -X'(\mu_1+\mu_2)/2 \end{aligned} \qquad (2)$$

where, S_g and μ_g are the variance-covariance matrix and mean vectors for group g ($g=1,2$), respectively, and n_g is the number of observations in group g. The cutoff value for FLDA is calculated by $b=ln(c_1p/c_2(1-p))$. Where c_1 and c_2 represent the misclassification costs for population 1 and 2, and p is the prior probability that the individual comes from population 1. The cutoff value for an FDLA model is equal zero, if the prior probability of group membership, and misclassification costs are the same.

Logit Discriminant Model

Some of the statistical DA models, such as LDA and PROBIT, define the discriminator value Z_i as a probability. The LDA model assumes a logistic distribution function to represent the probability that an individual i belongs to group g:

$$F(Z_i) = 1/[1 + \exp(-X_0 - B_{i1} X_1 - B_{i2} X_2 - \dots - B_{ij} X_j)] \quad (3)$$

Where $F(Z_i)$ converts the value of Z_i to a probability value. The Maximum-likelihood technique is usually used to estimate the weights (Maddalla). The selection of the cutoff value for the LDA model is rather arbitrary. Typically, if the estimated probability is greater than 0.5, then the first alternative is selected (Amemiya).

Mathematical Programming Discriminant Analysis Models

MP approach to discriminant problems, like statistical DA models, try to construct a discriminant function or a separating hyperplane to classify an individual or an object into a prespecified group. For a two-group problem, the objective is to determine a weighting vector X and a scalar b so that it assigns as correctly as possible the individuals of Group 1 to one side of the separation hyperplane and the individuals of Group 2 to the other side. Stating it mathematically, the objective of a MP model is to find b and nonzero X, satisfying:

$$A_1 X \geq b \qquad i \in G_1 \qquad (4)$$

$$A_2 X < b \qquad i \in G_1 \qquad (5)$$

where, A_g is an $n_g \times j$ matrix of observations and $i=1,2....N$, where N is the total number of observations ($N=n_1+n_2$).

The separating hyperplane, $AX=b$, provides the boundary between two groups. When two-group are not linearly separable, then one needs a criterion to separate the group classifications. Then, the MP formulation of a discriminant problem can be cast as:

$$\underset{X,\, b}{Optimize} \quad F(X, b) \qquad (6)$$

s.t:

$$A_1 X \geq b \qquad i \in G_1 \qquad (7)$$

$$A_2 X < b \qquad i \in G_2 \qquad (8)$$

$$X \neq 0 \qquad (9)$$

where $F(X,b)$ is the criterion function. The objective of this problem is to determine X and b that optimizes a certain criterion function. To develop the criterion function, one can incorporate deviation variables into (7) and (8).

$$\underset{X,\, b}{Optimize} \quad F(E_1, I_1, E_2, I_2) \qquad (10)$$

s.t:

$$A_1 X + E_1 - I_1 = b \qquad i \in G_1 \qquad (11)$$

$$A_2 X - E_2 + I_1 = b \qquad i \in G_2 \qquad (12)$$

$$X \neq 0 \tag{13}$$

where, E_g and I_g are deviation variables (what Glover, Keene, and Duea call external and internal deviations, respectively). A deviation is said to be external/internal if its associated observation is incorrectly/correctly classified (i.e., falls on the wrong/right side of the separating hyperplane). External/internal deviations represent the extent to which an observation is incorrectly/correctly classified. So, external deviations are undesirable while internal deviations are desirable. The above problem can be easily modified to handle multi-group classifications, as shown by Freed and Glover (1981b) and Gehrlein.

Depending on the choice of a criterion function, researchers have recently developed assorted MP models to deal with classification problems. Among the MP models are the minimize the sum of distances (MSD), the minimize the maximum distance (MMD), the mixed-integer (MIP), and the general l_p distance approaches. A variety of combinations of these basic methods have been proposed in the literature. Erenguc and Koehler (1990b) provide a comprehensive survey of various MP model formulations. As noted earlier, some of these models have proved to yield promising predictive power in studies using simulated data (Bajgier and Hill, Freed and Glover (1986); Joachimthaler and Stam 1988; and Rubin) and also using real data (Mahmood and Lawrence; and Srinivasan and Kim).

In last few years, there has been considerable research which has identified certain MP discriminant models, that, under certain data condition, exhibit some pathological problems which have not been experienced in applications of MP in other fields. Glover, Keene, and Duea classified these problems under the headings of degeneracy and stability. The solution to MP is said to be degenerate or unacceptable if $X=0$. The solution is unacceptable since all observations will be assigned to one group. The resultant discriminant functions lack any discriminatory power. The stability problem is referred to a situation where the solutions are not invariant to linear data translation and transformation. For a theoretical discussion of these problems see Koehler (pg. 19, 89b); Markowski and Markowski (1985); Freed and Glover (1986b); and Glover, Keene, and Duea.

Early MP models constrained b to be a constant to avoid the unacceptable solutions. It was tacitly assumed that choice of b would just scale the solutions but further research in this area found that it is not the case and still leads to $X=0$ for certain data configurations (Glover). Recently, several normalization alternatives have been suggested to overcome with these anomalies. Details regarding alternative normalizations can be found in Koehler (1990). Since it is possible that normalization eliminates a feasible region with potential optimal solutions, a user should be cautious when employing normalization. To this end, Rubin recommends that

> ... The precise impact of a particular normalization constraint is generally difficult to assess, and so the selection of normalization constraints tends to be arbitrary. Since trivial solutions generally do not occur when the training sample (estimation sample) are separable, perhaps users should initially omit such normalization constraints and incorporate them only after obtaining a trivial solution for a particular data set" [explanation in parentheses added] (Rubin, p16).

In this paper, normalization was incorporated into the MP models, if it was deemed to be necessary. In the remaining section, we will present four variants of MP discriminant models. These models are chosen among alternative MP models based on their competitive classification power and also their appropriateness dealing with credit scoring problem.

The first MP model, hereafter referred to as a MSD model, can be summarized as follows:

$$\underset{X,\, b}{\textit{Minimize}} \quad e_1 E_1 + e_2 E_2 \tag{14}$$

s.t:

$$A_1 X + E_1 \geq b \qquad i \in G_1 \tag{15}$$

$$A_2 X - E_2 < b \qquad i \in G_2 \tag{16}$$

$$e_3 X + b = 1 \tag{17}$$

$$X, b \quad u.r.s. \tag{18}$$

$$E_1, E_2 \geq 0 \tag{19}$$

where e_1, e_2, e_3 are $1 \times n_1$, $1 \times n_2$, $1 \times j$ matrices of ones, respectively, and E_g has dimension $n_g \times 1$. X and b are unrestricted in signs.The MSD model minimizes the sum of exterior deviations from the hyperplane. Equation (17), a normalization constraint suggested by Freed and Glover (1986a), is included to overcome the difficulties associated with unacceptable solutions. The normalization constraint requires the sum of all coefficients to be equal to some arbitrary (positive) constant (1 is used here). The constant term is only a scaling constant and does not affect the classification rates. The MSD model, without normalization constraint (17), was originally published by Freed and Glover (1981b).

The second MP model used in this study, called the optimize sum of distances (OSD) by Bajgier and Hill, has the following form:

$$\underset{X}{\textit{Minimize}} \quad e_1 E_1 + e_2 E_2 \tag{20}$$

s.t:

$$A_1 X + E_1 \geq 1 \qquad i \in G_1 \tag{21}$$

$$A_2 X - E_2 < 1 \qquad i \in G_2 \tag{22}$$

$$X \quad u.r.s. \tag{23}$$

$$E_1, E_2 \geq 0 \tag{24}$$

MSD model is similar to OSD. Both model attempt to minimize the sum of external deviations from the hyperplane. But, in the OSD model, the cutoff value is preassigned to be equal to an arbitrary number (1 is used here) which precludes the need for normalization constraint.

The third MP model, hereafter referred to as the HB, seeks to:

$$\underset{X,\, b}{Minimize} \quad h_1 E_1 + h_2 E_2 - m_1 I_1 - m_2 I_2 \qquad (25)$$

$s.t:$

$$A_1 X + E_1 - I_1 = b \qquad i \in G_1 \qquad (26)$$

$$A_2 X - E_2 + I_2 = b \qquad i \in G_2 \qquad (27)$$

$$e_j X + b = 1 \qquad (28)$$

$$X,\ b \quad u.r.s \qquad (29)$$

$$E_1,\ E_2,\ I_1,\ I_2 \geq 0 \qquad (30)$$

$$h_1 \geq m_1,\ h_2 \geq m_2 \qquad (31)$$

where, h_g and m_g are $1 \times n_g$ matrix of nonnegative objective coefficients. The objective function of the HB model maximizes the weighted sum of interior deviations and minimizes the weighted sum of exterior deviations. Constraint (28) is included in the model formulation to prevent potential unbounded solutions. In practice, the h_g and m_g may reflect the relative importance of incorrect/correct classification to a particular group or individuals in the group. By modifying these weights and parameters, usually by LP post-optimization techniques as proposed by Glover, the solution might be tailored to meet a decision maker's specific goals. In other words, it might be possible to find a set of weights that achieves balancing of errors for a decision maker's particular set of data. (Markowski).

The HB model was first presented by Glover, Keene and Duea; they dubbed this model as a Hybrid model since it can encompass several variations of MP models by setting the corresponding weights equal to either $+\infty$ or $-\infty$. The HB model presented here is, however, different from the model presented by Glover, Keene and Duea. For simplicity, the maximum exterior deviation and the minimum interior deviation were deleted from the model formulation.

The final MP variant used in this study is a mixed integer programming model (MIP). The MIP model has the form:

$$\underset{X,\, b}{Minimize} \quad h_1 Y_1 + h_2 Y_2 \qquad (32)$$

$s.t:$

$$A_1 X + q Y_1 \geq b \qquad i \in G_2 \qquad (33)$$

$$A_2 X - q Y_2 \leq b \qquad i \in G_2 \qquad (34)$$

$$e_j X + b = 1 \qquad (35)$$

$X, b \quad u.r.s.$ (36)

$Y_1, Y_2 \quad \epsilon (0, 1)$ (37)

where h_g denotes the misclassification costs associated with group g; Y_i is binary variable that equal one if individual i is misclassified and zero otherwise; and q is a large positive number. The objective function of MIP model minimizes total misclassification costs. The interesting feature of the MIP model, as noted by Bajgier and Hill, is that it is the only model that directly attacks the goal of minimizing the number of misclassifications. Whereas, all other DA models (including parametric and nonparameteric models) use a surrogate criterion function to achieve the goal. If misclassification costs for both groups are the same, then, MIP directly minimizes the number of misclassification. Whereas, all other models minimize the amount or extent of misclassification from the hyperplane which might not be intuitively appealing to the users. Another interesting feature of MIP is that a constraint can be easily incorporated into the model to balance the number of misclassification for each group. In spite of its potential, the MIP model has not been widely utilized by researchers and practitioners because of a large computational cost and lack of efficient software. Koehler and Erenguc (1990a) recently developed a special purpose mixed integer algorithm which takes advantage of the problem's structure. Moreover, because of the recent decrease in computing cost and increase in computing power, some of general purpose mixed integer program packages can now be conveniently applied to solve larger MIP problems.

The discussion in the last two sections emphasized the estimation of a linear discriminant function for a two-group classification problem using alternative econometrics and MP credit scoring models. We have, however, omitted the theoretical aspects of credit scoring problem. As noted by Turvey "...the credit scoring models themselves will not be successful in assessing the success of a particular loan". There are several other factors that one has to consider in order to establish a successful credit scoring practice. The choice of discriminant variables, the level of subjectivity, the institutional constraints and several other factors are important consideration in any credit scoring study, but in this paper we only concentrate on providing a comparison between the classification performance of MP and statistical credit scoring models. Chhikara; Betubiza and Leatham; and Miller and LaDue provide a detailed discussion of credit scoring issues.

Data and Variable Selection

To perform a comparative analysis, the above models are applied to estimate the corresponding discriminant functions using a sample of credit application data. The classification power of these models are then tested based on their performance using in-sample and hold-out sample data. The credit application data used in this paper were collected by Canada's Farm Credit Corporation. The data are from actual 1981, 1982, and 1983 loan applications for which loans were made in the Saskatchewan Province. The applicants in Group 1 consist of individuals with recent histories of delinquent credit payments (noncurrent loans) and applicants in Group II consists of those individuals without recent delinquent credit payments (current loans) based on the status of the loan as of March 1990. The sample consisted of 754 current loan applications (38%) and 1,245 of noncurrent loan applications (62%). The sample data was divided into two subsamples- an in-sample and a hold-out samples. The in-sample data set was used for model developments and the resulting models were then validated using the in-sample and hold-out sample sets. In this study, 60% (1,199 loans) of total sample was used for model estimation.

The usual procedure in credit scoring studies is to select a large group of explanatory variables and reduce that to a smaller number of statistically significant variables. The above data set was recently used by Turvey in a study in which he compared alternative statistical credit scoring models. Our investigation only included the explanatory variables used in his study to avoid potential overfitting biases. Definition of the explanatory variables are presented in Table 1. Turvey and Brown; and Turvey provide a more formal definitions and explanations of the explanatory variables.

Table 1. Explanatory Variables for Credit Scoring Application

Variable	Variable Definition
Liquidity Ratio	Current asset / current liabilities
The Rate of Return on Assets	(Net farm income + interest expense) /total assets
Debt-to-Asset Ratio	Debt / assets
Loan-to-Security Ratio	(Loan + other prior changes + FCC prior mortgages + statutory charges) / total security
Contribution Margin	(Total Revenue - variable cost)/total revenue
Repayment Ratio	(Net farm income + depreciation + off-farm income-living costs + interest on term loans) /(principal + interest on term loans)
Refinancing Status	1 if loan is required for refinancing, 0 otherwise.

The HB and MIP models require the parameters h_g and m_g to be specified. As was discussed earlier, in practice, these parameters could be solicited from the user. Since the actual benefits and costs of external and internal deviations from the hyperplane for current and noncurrent loan applicants were not available for this study, a set of arbitrary values were selected for these parameters. Subsequently, four variants of HB model (denoted by HB-1, HB-2, HB-3, and HB-4) and three variants of MIP model (denoted by MIP-1, MIP-2, MIP-3) were tested. Table 2 presents the objective coefficients associated with variants of HB and MIP models.

The HB-1 model maximizes the total interior distances from the hyperplane and minimizes the total exterior distances from the hyperplanes. The HB-2, HB-3, and HB-4 maximize the weighted sum of interior distances and minimizes the weighted sum of exterior distances from the hyperplane. The objective is to provide a better balancing of errors between current and noncurrent loans by varying the objective coefficients assigned to interior and/or exterior deviations. As was discussed earlier, in contrast to other MP models, the objective function of MIP model has a direct and meaningful interpretation. For example, the MIP-1 model assumes that the misclassification costs for current and noncurrent loans are the same, hence, the MIP-1

Table 2. The Objective Coefficients for Varients of HB and MIP Models

	m_1	h_1	m_2	h_2
HB-1	1	1	1	1
HB-2	1	1	1	0.5
HB-3	1	1	1	0.25
HB-4	1.25	1	1	0.25
MIP-1	n.a.	1	n.a.	1
MIP-2	n.a.	0.645	n.a.	0.364
MIP-3	n.a.	2	n.a.	1

n.a.- Not Applicable

model directly minimizes the number of misclassifications for both groups. The MIP-2 model is similar to MIP-1 model, but the weights are proportionally weighted based on sample size in each group. The MIP-3 model, however, assumes that the misclassification cost for a noncurrent loan is twice as much as a current loan.

Overall, we tested eleven models, two parametric and nine nonparametrics models. The next section discusses the classification results of these models.

Classification Results

The classification performance on the calibration and hold-out samples of alternative models are presented in Tables 3 and 4. Table 3 presents the classification performance in terms of number of loans in calibration and hold-out samples, while Table 4 presents the same results in percentage. The results in Table 4 and 5 show that the classification performance for parametric models are not significantly different from each other. Both LDA and FLDA models, however, predict better than a pure naive model (i.e. predict better than the proportional prior probabilities for current loan, 36.4% and noncurrent loan, 63.6%). But, both MSD and OSD models perform significantly worse than LDA, FLDA and the naive model for both calibration and hold-out samples. Among four HB models tested, the classification performance of the HB-4 model in hold-out sample is worse than the other three HB models. The results in Table 3 and 4 show the HB-2, HB-3, and HB-4 models perform as well as statistical models in hold-out samples. However, all three MIP models perform marginally better than LDA and FLDA models for calibration and hold-out samples. The LDA classified correctly 601 of noncurrent loans in the calibration sample for a correct classification rate of 65%. While, the overall correct classification rate for LDA is 66% for the calibration sample (Table 4). The MIP-1, MIP-2 and MIP-3 models classified correctly 622, 609, and 609 of noncurrent loans in the calibration sample for a correct classification rate of 66%, 68%, and 68%, respectively. The overall correct classification rate for MIP-1, MIP-2, and MIP-3 are 67%, 67%, and 68% for the calibration sample, respectively, which marginally better than LDA and FLDA models. The results in Table 3 and 4 show that the LDA and FLDA models, however, provide a more balanced discriminant solutions than MP models. None of the MP models tested here show a higher correct classification for current loans.

Conclusions

The purpose of this study has been to compare the alternative statistical and MP credit scoring models in an empirical setting using the actual credit data. The results indicate that there are only a small differences in the classifying accuracy of statistical and MP approaches. The results of this study re-enforce the findings of the experimental studies which claim that the MP models are as competitive as statistical DA models. As was shown here, the MIP models even outperform the statistical models. We recommend the use of MP models in an applied environment when the incorporation of a side condition becomes necessary, or a small sample size is available, or a large number of explanatory variables is present, or the data set is heavily contaminated. In these situations, the MP models have the potential to perform better than the statistical DA models. Since there is no optimal DA model which fits all data sets in all situations, it may be a good practice to apply the data to alternative parametric and nonparametric DA models and choose the best model. In many credit scoring applications, even a moderate improvement in the ability to correctly classify may represent a significant increase in financial contributions.

Table 3. Classification Results Reported as a Percentage of Loans Classified Correctly and Incorrectly.

		Number of Loans Classified Correctly and Incorrectly					
		Current		Noncurrent		Classification[a]	
Data Sample	Models	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect
Calibration Sample							
	LDA	177	259	601	162	778	421
	FLDA	183	253	597	166	780	419
	MSD	143	293	493	270	636	563
	OSD	133	303	528	235	661	538
	HB1	123	313	622	141	745	454
	HB2	146	290	609	154	755	444
	HB3	159	277	594	169	753	446
	HB4	171	265	575	188	746	453
	MIP-1	107	329	687	76	794	405
	MIP-2	147	289	666	97	813	386
	MIP-3	132	304	680	83	812	387
Hold-out Sample							
	LDA	184	134	344	138	528	272
	FLDA	184	134	341	141	525	275
	MSD	109	209	316	166	425	375
	OSD	111	207	338	144	449	351
	HB1	130	188	379	103	509	291
	HB2	143	175	373	109	516	284
	HB3	148	170	365	117	513	287
	HB4	161	157	346	136	507	293
	MIP-1	116	202	420	62	536	264
	MIP-2	142	176	392	90	534	266
	MIP-3	134	184	407	75	541	259

[a] Current and noncurrent loans.

Table 4. Classification Results Reported as a Percentage of Loans Classified Correctly and Incorrectly

		Percent of Loans Classified Correctly and Incorrectly					
		Current		Noncurrent		Classification[a]	
Data Sample	Models	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect
Calibration Sample							
	LDA	41	59	79	21	65	35
	FLDA	42	58	78	22	65	35
	MSD	33	67	65	35	53	47
	OSD	31	69	69	31	55	45
	HB1	28	72	82	18	62	38
	HB2	33	67	80	20	63	37
	HB3	36	64	78	22	63	37
	HB4	39	61	75	25	62	38
	MIP-1	25	75	90	10	66	34
	MIP-2	34	66	87	13	68	32
	MIP-3	30	70	89	11	68	32
Hold-out Sample							
	LDA	58	42	71	29	66	34
	FLDA	58	42	71	29	66	34
	MSD	34	66	66	34	53	47
	OSD	35	65	70	30	56	44
	HB1	41	59	79	21	64	36
	HB2	45	55	77	23	65	36
	HB3	47	53	76	24	64	36
	HB4	51	49	72	28	63	37
	MIP-1	36	64	87	13	67	33
	MIP-2	45	55	81	19	67	33
	MIP-3	42	58	84	16	68	32

[a] Current and noncurrent loans

References

Altman, E. I., R. B. Avery, R. A. Eisenbeis, and J. F. Sinky. Application of Classification Techniques in Business, Banking and Finance, JAI Press Inc., 1981.

Amemiya, T. "Qualitative Response Models: A Survey." *J. of Econ. Lit.* 19(1981):1483-1536.

Bajgier, S. M., and A. V. Hill "An Experimental Comparison of Statistical and Linear Programming Approaches to the Discriminant Problem." *Decision Sciences* 13(1982): 604-18.

Betubiza, E. and D. J. Leatham. A Review of Agricultural Assessment Research and Annotated Bibliography, B-1688, Department of Agricultural Economics, Texas A&M University, College Station, TX, 1990.

Balakrishnan, N. and K. Subrahmaniam. "Robustness to Nonnormality of the Linear Discriminant Function: Mixtures of Normal Distributions." *Communications in Statistics- Theory and Methods* 9(1980):13-25

Chhikara, K. R. "The State of the Art in Credit Evaluation." *Amer. J. Agr. Econ.* 71(1989):1138-44.

Eisenbeis, R. A. "Pitfalls in the Application of Discriminant Analysis in Business, Finance, and Economics." *J. of Finan.* 32(1979):875-900.

Freed N., and F. Glover. "A Linear Programming Approach to the Discriminant Problem." *Decision Sciences* 12(1981a):68-74.

________________ "Simple but Powerful Goal Programming Models for Discriminant Problems." *European J. of Operational Research* 7(1981b):44-60.

________________ "Resolving Certain Difficulties and Improving the Classification Power of the LP Discriminant Analysis Procedure." *Decision Sciences* 17(1986a):589-95.

________________ "Evaluating Alternative Linear Programming Models to Solve the Two-Group Discriminant Problem." *Decision Sciences* 17 (1986b):151-62.

Glover, F. "Improved Linear Programming Models for Discriminant Analysis." *Decision Sciences* 21(1990):771-85.

Gehrlein, W. V. "General Mathematical Programming Formuations for the Statistical Classification Problem." *Operations Research Letters* 5 (1986):299-304.

Glover, F., Keene, S., and Duea, B. "A New Class of Models For the Discriminant Problem." *Decision Sciences* 19(1988):269-80.

Goldstein, M., and W. R. Dillon. Discrete Discriminant Analysis. New York: Wiley, 1978.

Hand, D. J. Discrimination and Classification. New York: Wiley, 1981.

Joachimsthaler E. A. and A. Stam. "Four Approaches to the Classification Problem in Discriminant Analysis: An Experimental Study." *Decision Sciences* 19(1988):322-33.

Johnson, R.A. and D. W. Wichern. Applied Multivariate Statistical Analysis, Prentice-Hall, Inc., N.J., 1982.

Koehler, G. J. "Considerations for Mathematical Programming Models in Discriminant Analysis." *Managerial and Decision Economics* 11 (1990):322-33.

Koehler, G. J., and S. S. Erenguc. "Minimizing Misclassifications in Linear Discriminant Analysis." *Decision Sciences* 21(1990a):63-85.

______________________________ " Survey of Mathematical Programming Models and Experimental Results for Linear Discriminant Analysis." *Managerial and Decision Econ.* 11(1990b): 215-25.

Lachenbruch, P.A., C. Sneeringer, and L.T. Revo. "Robustness of the Linear and Quadratic Discriminant Functions to Certain Types of Non-Normality." *Communications in Statistics* 1(1973): 39-56

Maddala, G. S. Limited Dependent and Qualitative Variables in Econometrics. Cambridge: Cambridge University Press, 1987.

Markowski C. "On the Balancing of Error Rates for LP Discriminant Methods." *Managerial and Decision Economics* 11(1990):235-41.

Markowski, E. P., and C. A. Markowski. "Some Difficulties and Improvements in Applying Linear Programming Formulations to the Discriminant Problem." *Decision Sciences* 16(1985):237-47.

______________________________ "An Experimental Comparison of Several Approaches to the Discriminant Problem with Both Qualitative and Quantitative Variables." *European J. of Operation Research* 28(1987):74-78.

Mahmood, M., and E. Lawrence. "A Performance Analysis of Parametric and Nonparametric Discriminant Approaches to Business Decision Making." *Decision Sciences* 18 (1987):308-26.

Miller, L. H., and E. L. La Due. "Credit Assessment Models for Farm Borrowers: A Logit Analysis." *Agr. Finan. Rev.* 49(1989):22-36.

Press, S. J., and S. Wilson. "Choosing between logistic Regression and Discriminant Analysis." *J. of the Amer. Statist. Assoc.* 73 (1978):699-705.

Rubin, P. A. "A Comparison of Linear Programming and Parametric Approaches to the Two-Group Discriminant Problem." *Decision Sciences* 21(1990):374-86.

Srinivasan, V., and Y. H. Kim. "Credit Granting: A Comparative Analysis of Classification Procedures." *J. of Finan.* 42(1987):665-81.

Turvey, C. G., and R. Brown. "Credit Scoring for a Federal Lending Institution: The Case of Canada's Farm Credit Corporation." *Agr. Finan. Rev.* 50(1990):47-57.

Turvey, C. G. "Credit Scoring for Agricultural Loans: A Review with Applications." *Agr. Fin. Rev.* 50(1991):43-54.