[Journal of Student Financial Aid](https://ir.library.louisville.edu/jsfa)

Volume 27 | Issue 1 Article 3

2-1-1997

A Model for Predicting Perkins Loan Defaulters

Tina M. Thobe

Barbara M. DeLuca

A Model for Predicting Perkins Loan Defaulters

By Tina M. Thobe and
Barbara M. DeLuca

Tina M. Thobe was an undergraduate honors student in the Department of Economics and Finance at the University of Dayton at the time this article was written. Barbara M. DeLuca is an Associate Professor in the Department of Educational Administration in the School of Education at the University of Dayton.

The Perkins Loan Program was established in 1958 by the federal government to create loans for financially needy students. This program is one of many student loan programs that has come under heavy scrutiny in recent years because of high default rates. There are relatively few studies to help university administrators identify borrowers likely to default and almost no predictive models to assess the default probabilities of individual borrowers.

This project analyzes several characteristics of University of Dayton student borrowers to discover their relative impact on whether or not a student defaults. Further, a model is developed to help predict those students who will potentially default on their Perkins Loans. Financial aid policy recommendations aimed at reducing default costs at the University of Dayton are made based on that model.

The purpose of this research is to develop a model to predict potential Perkins Loan defaulters earlier in an effort to reduce default. Since legally no one can be denied a Perkins Loan based solely on risk of future default, any efforts aimed at reducing default are best undertaken at the time the student separates/graduates from college and begins repayment. If default on Perkins loans could be predicted at the time of separation/graduation, universities and colleges could use this information to minimize or at least manage potential losses due to default. Financial aid counselors and loan administrators would have more information with which to assess the default risk associated with student borrowers. As a result, "high risk" borrowers could be counseled and monitored in an attempt to increase their likelihood of repayment.

The losses associated with default are suffered by many. The federal government (U.S. Department of Education) incurs a loss because it initiates the fund and makes contributions (albeit, minimal) to it thereafter. Further, the school has a stake in the losses because it makes annual contributions equal to one ninth of government contributions. Additionally, the school often subsidizes the fund each year. However, the real loss associated with default on Perkins Loans is felt by future financially-needy students of the school. Because of the revolving nature of the fund and the absence of significant increased yearly funding, additional loans cannot be made if previous loans are not eventually paid off.

Federally-funded student loans, including Perkins Loans, are awarded on the basis of student need only. Unless the student clearly indicates his or her intention not to repay the loan, no student can be refused a Perkins Loan solely due to risk of future default. Thus, the default risk associated with federal student loans can potentially be greater than the default risk associated with other bank or credit union consumer loans.

Federally-funded student loan programs have come under heavy scrutiny in recent years because of default rates that began to soar during the 1980s. The cost to the federal government of paying off defaulted student loans had risen to over $2 billion annually by 1991 (Merisotis, 1991). As more students

TABLE 1
University of Dayton Perkins Loan Default Rates

Date	6/83	6/84	6/85	6/86	6/87	6/88	6/89	6/90	6/91	6/92	6/93
Default Rate (%)	7.22	6.93	7.35	6.30	7.10	5.93	5.87	4.62	3.94	3.38	2.92

Source: University of Dayton Bursar's Office

borrow to pay for rising college costs, and particularly as more lower-income, high-risk students borrow, the losses due to default also increase. Furthermore, investigations into schools with the highest average default rates have revealed waste, fraud, and abuse (Merisotis, 1991). Although wrongdoing is not prevalent at most four-year institutions, all loan programs are usually lumped together in regulators' minds and reform initiatives typically affect all programs as a matter of prevention.

Table 1 shows the default rate for the Perkins Loan Program at The University of Dayton (UD) for the last eleven years. The decline in this rate over the last six years has occurred as a direct result of the increased attention that student loan default has received from lawmakers, politicians, and college administrators. At the end of the 1992-93 academic year, the Perkins Loan default rate at UD was just under 3%. Although the Perkins Loan default rate at UD has decreased substantially over the last decade, a study of the factors commonly associated with default would help to pinpoint the origins of default, thereby allowing administrators to address the source of the problem instead of treating the symptoms of it.

This research project identifies variables that are significantly related to default so that potential defaulters may be distinguished from non-defaulters at the time of separation from the University. Specifically, a model is developed that predicts potential defaulters. From this model, policy recommendations are suggested to show how the model can help to reduce default costs.

Literature Review Shows Difficulty of Predicting Default

The basis for this research stems chiefly from studies done by Wilms, Moore, and Bolus (1987), and Greene (1989). Both studies develop predictive models to assess students' default potential. They also offer policy suggestions for the reduction of default costs.

Wilms et al. (1987) used cross-tabulations and discriminant function analysis to analyze a large sample of students attending California proprietary and community colleges in 1982-83. They identified six variables which they found to have a significant association with default: program completion, ethnicity, average annual income, citizenship, prior education, and program of study. Wilms et al. conclude that default rates associated with California proprietary and community colleges stem mainly from students' background characteristics as opposed to the characteristics or administrative practices of the institutions they attend. They suggest rewarding efficient program administration; taking action against schools that abuse the system by failing to make proper refunds, falsifying applications, etc.; and reducing defaults among low-income borrowers by tying student loans to student performance, or converting loans to grants for all low-income, first-year students.

In a later study, Greene (1989) used the Tobit technique to examine a small sample of students at the University of North Carolina at Greensboro, a four-year, public university. She used the Wilms et al. (1987) model as a starting point for her research by evaluating her data with their model, and she arrived at similar results. However, the main thrust of her work is to present an alternative method of analysis, the Tobit technique. She prefers the Tobit technique because it can analyze both the student's categorical status (default/no default) and the dollar amount of default for students who have defaulted, whereas discriminant function analysis can only account for the categorical default/no default decision.

Greene identified three variables which she found to be significantly related to default: program completion, scholarship/grant aid received, and grade point average. Also, she notes that two variables that Wilms et al. found to be significant, ethnicity and average annual income, were found to be insignificant when she used Tobit. Greene recommends that her model be used to analyze student characteristics that are available before a student's loans have come due for those students who are associated with high probabilities of default. Then, the Greene model can also be used by loan program administrators to predict dollar losses to defaults from a given loan portfolio of borrowers, thereby enabling administrators to prepare to meet those costs. She proposes that schools identify borrowers with potentially high probabilities of default. Then, those students can be helped to meet the demands of their programs and facilitate completion. School administrators can also counsel students on the difficulties of meeting debt burdens and students can be trained in job-seeking skills.

Most government research has only produced descriptive statistics [for instance, U.S. General Accounting Office (GAO), September 1990]. One government study identified nine characteristics most frequently cited in published studies. Only one of the studies reviewed, the Wilms et al. (1987) study, used defaulter characteristics to predict default; the rest were descriptive in nature. Generally, students likely to default on their loans were ones who attended vocational/trade schools, had low incomes, had little financial support, had minority backgrounds, lacked high school diplomas, failed to complete their educational programs, attended school for one year or less, borrowed small amounts, and were unemployed when defaulting. Nonetheless, while most defaulters have certain characteristics, the majority of defaulters with these characteristics do not default on their loans. Moreover, the government has not tried to validate the results of any study (U.S. GAO, April 1991).

Methodology

After explaining the sample source, a brief rationale for the primary method of analysis, discriminant function analysis, is offered. The variables theoretically related to default are identified. Descriptive statistics for the variables collected are calculated, using SPSS (Statistical Package for the Social Sciences, a computer software package) to offer an overview of the sample and identify potential relationships between characteristics. Next, discriminant function analysis is performed, again using SPSS, to determine which variables are significantly related to default and to develop a predictive model.

Sample and Database

The sample consists of 392 University of Dayton student records. The University of Dayton is a four-year, comprehensive, private university with approximately 6,000 undergraduate and 4,000 graduate students. It is considered primarily a regional university located in Dayton, Ohio, serving the Midwest and East Coast. The sample used records of those students who had Perkins loans outstanding in 1993, separated from the university between the fall term of 1988 and the summer term of 1990, inclusive, and filed a financial aid application during their last year at UD as an undergraduate. Since the fall of 1988 was the first term in which an integrated database of financial aid, registration, and application data was available at UD, the sample begins at that point.

Further, in order to give a student who left the university time to begin repaying (or possibly defaulting) on his or her Perkins Loan, the sample could extend no further than students who left the university during the summer of 1990. For instance, if a student left UD in July of 1990, and assuming his or her grace period was nine months and he or she had not applied for a deferment, he or she would have begun repayment in May of 1991. This would have given the student two-and-one-half years to make appropriate payments and/or default since the Perkins Loan payment data ran until December of 1993. Other reasons for choosing the years identified above are related to limitations of the data unique to each institution, in this case the University of Dayton.

All financial aid data from the student's Financial Aid Form (FAF) was obtained from the financial aid record of the student's last year in attendance at UD as an undergraduate. The data from this year was seen as the closest indication of the financial position of the student and his or her family when the student separated from UD and was about to enter into repayment.

Discriminant Analysis

Discriminant function analysis is used as the main tool of analysis. Discriminant analysis is a statistical technique that allows the researcher to study differences between two or more groups of objects with respect to several variables simultaneously. Specifically, an optimal set of discriminating variables is determined by using a stepwise procedure that maximizes the overall separation between the two groups as measured by Rao's V. The stepwise procedure is used so that poor discriminators or discriminators that share the same discriminating characteristics are not included in the final equation.

In the framework of this project, discriminant analysis will be used for both interpretation and classification. Interpretation is used when studying the ways in which groups differ. In other words, one looks at the ability of a set of characteristics to discriminate between the groups, the extent to which they can discriminate, and which characteristics are the most powerful discriminators. The other application is to derive one or more mathematical equations for the purpose of classification. These equations, or discriminant functions, combine the group of characteristics in a way that will allow one to identify the particular group which a case most closely resembles (Klecka, 1980).

Variables

Criterion Variable. The criterion variable used to sort the sample is DEFAULT. A Perkins Loan is considered "in default" when the borrower has not made a payment on the loan for over 120 days. (This definition of default was chosen for this research study only; it is not that used by the U.S. Department of Education.)

Technically, a borrower can float into and out of default throughout the life of the loan. The borrower may fail to make payments for a certain period of time, but then may begin to make regular payments again. Each time a payment has not been made for more than 120 days, it is noted on the borrower's record. Thus, even though a person may have been in default several times over the life of the loan, he or she may eventually pay off the entire loan, although over a period of time longer than originally planned.

On the other hand, some loans that enter into default remain in default for extended periods of time. Such loans are determined to be uncollectible by the institution and are assigned to the U.S. Department of Education. At that point, the loan is written off as a loss to the Perkins Loan fund at the institution even though the U.S. Department of Education may eventually collect on it. Loans for which a payment has not been made for 120 days impose considerably more administrative costs on the Perkins Loan Program. Thus, for this study, a loan is labeled as being in default if the borrower did not make a payment on the loan for 120 days at least one time throughout the life of the loan.

"If default on Perkins Loans could be predicted at the time as separation/ graduation, schools could use this information to minimize, or at least manage, potential default losses."

The sample is divided into two groups: those students who have been 120 days late on at least one payment (defaulters) and those who have not (non-defaulters). Thus, DEFAULT, a dichotomous (two-category) variable, is the dependent variable of the proposed model. DEFAULT equals 1 if the borrower made at least one 120-day or later payment on the Perkins Loan and 0 if the borrower never made a payment that was 120 days late. The results of the model identify significant characteristics that can be used to predict whether or not a student will default on a Perkins Loan. First, several variables are hypothesized as being possible determinants of Perkins Loan default. These independent variables consist of student traits, student academic characteristics, student financial data, and parental information.

Student Traits. The ethnicity, nature of entry into the university, citizenship, marital status while an undergraduate, and gender of the student were the student traits that were entered into the initial analysis to see if they significantly help to discriminate between defaulters and non-defaulters.

- The ethnicity of the student is denoted by a binary (two-category) variable (ETHNIC) that equals 1 if minority and 0 if Caucasian.
- The nature of entry is symbolized by a binary variable (ENTRY) that equals 1 if the student entered as a transfer student and 0 if the student entered as a first-year student.
- Citizenship is represented by a binary variable (CITIZEN) that equals 1 if the student is an eligible noncitizen and 0 if the student is a U.S. citizen.
- The marital status of the student while an undergraduate is indicated by a binary variable (SMARSTAT) that equals 1 if married and 0 if single.
- The gender of the student is denoted by a binary variable (GENDER) that equals 1 if female and 0 if male.

Student Academic Characteristics. The number of years registered, nature of separation from the university, and grade point average upon separation were the academic characteristics that were used in the initial analysis.

- The number of years registered at UD (REGYEARS) gives an indication of how long the student attended before departing due to graduation or simply leaving the university.
- The nature of separation from UD is symbolized by a dichotomous variable (GRAD) that equals 1 if the student did not graduate and 0 if the student graduated.
- The UDGPA variable indicates the cumulative grade point average that the student achieved as a undergraduate while at UD.

Student Financial Data. The income, assets, dependent status, grants, scholarships, on-campus earnings, total loans, cumulative Perkins Loan amount, and minimum Perkins Loan payment of the student are the financial data of the student.

"The losses associated with default are suffered by many."

- SINC signifies the adjusted gross income of the student during the year preceding the student's last year as an undergraduate.
- SASSET shows the assets (cash, value of business, value of home, and other revenue) that were declared by the student during the year preceding the student's last undergraduate year.
- The dependent status of the student is represented by a binary variable (DEPSTAT) that equals 1 if dependent and 0 if independent. DEPSTAT indicates whether the parents of the student claimed that child for income tax purposes and provided a significant amount of financial support for the child during the last year he or she was an undergraduate.
- GRANTS equals the amount of grant aid (non-repayable) received by the student during the last undergraduate year.
- SCHOLAR is equivalent to the amount of merit-based scholarship aid (non-repayable) received by the student during the last undergraduate year.
- EARNED equals the amount of money earned by the student through on-campus jobs during his or her last undergraduate year.
- LOANACC equals the total dollar value of all loans received by the student during his or her last undergraduate year.
- LOANAMT is equivalent to the total dollar value of all Perkins Loans accumulated by the student while at UD.
- MPAY equals the minimum monthly Perkins Loan payment that the student is required to pay.

Parental Information. The marital status, number of dependents, income and assets of the parent(s) of the student constitute the parental information that was entered into the initial analysis to see if it significantly helps to discriminate between defaulters and non-defaulters.

- The marital status of the student's parents is denoted by a dichotomous variable (FAMTYPE) that equals 1 if the undergraduate student came from a single-parent home and 0 if the student came from a two-parent family.
- DEP represents the number of dependents in the undergraduate student's family (including the student).
- PARINC signifies the adjusted gross income of the student's parents during the year preceding the last year the student was an undergraduate.

TABLE 2
Percentage of Variable Occurances*

Variable	Overall (n = 392)		Non-defaulter (n = 315)		Defaulter (n = 77)	
	0	1	0	1	0	1
Whether or not Default Occurred 0 = no default, 1 = default	80.4%	19.4%	100.0%	0.0%	0.0%	100.0%
Status at Time of Entrance to UD 0 = freshman, 1 = transfer	83.2	16.8	82.2	17.8	87.0	13.0
Ethnic Group 0 = white, 1 = minority	89.2	10.8	92.7	7.3	75.3	24.7
Gender 0 = male, 1 = female	53.1	46.9	51.1	48.9	61.0	39.0
Eligibility Status Based on Citizenship 0 = U.S. citizen, 1 = eligible noncitizen	99.2	0.8	99.4	0.6	98.7	1.3
Student Martial Status 0 = single, 1 = married	95.2	4.8	94.6	5.4	97.4	2.6
Status at Time of Separation 0 = graduated, 1 = not graduated	68.6	31.4	73.3	26.7	49.4	50.6
Dependent Status of Student 0 = dependent, 1 = independent	87.5	12.5	88.3	11.7	84.4	15.6
Types of Family 0 = two-parent, 1 = single-parent	66.3	33.7	68.3	31.7	58.4	41.6

*Variables with only two categories, 0 or 1.
The values in the body of the table represent the percent of the sample in each respective category.

- PASSET shows the assets (cash, value of business, value of home, and other revenue) that were declared by the student's parents during the year preceding the last year he or she was an undergraduate.

Findings

Descriptive Statistics. Table 2 identifies the frequency for each binary variable that is present in each of three samples: the overall sample (n = 392), non-defaulter subsample (n = 315), and defaulter subsample (n = 77). Of the 392 students in the overall sample, 19.4% "defaulted" on their Perkins Loan, or in other words, made at least one 120-day or later payment. (Recall that this is not the U.S. Department of Education definition of default.)

The proportion of transfer students in the non-defaulter subsample is higher than the proportion in the defaulter subsample; nearly 18% of the non-defaulters were transfer students while only 13% of the defaulters were. Nearly one-fourth of the defaulter subsample is minority while only about 7.3% of the non-defaulters were minorities. Interestingly, there is a higher proportion of males than females in the defaulter subsample, 61% and 39% respectively. A considerable portion of the defaulter subsample did not graduate from UD (50.6%), while only 26.7% of the non-defaulters did not graduate. Relatively more independent students defaulted (15.6%) as compared to the portion in the non-defaulter subsample (11.7%). Lastly, 41.6% of the defaulter subsample came from a single-parent family, whereas only 31.7% of the non-defaulter subsample came from a single-parent family.

Table 3 identifies the mean and standard deviation for each continuous variable for the overall sample (n = 392), the non-defaulter subsample (n = 315), and the defaulter subsample (n = 77). The minimum and maximum value for each variable are also given.

Several differences between the two subsamples are apparent. The mean GPA achieved by a defaulter (2.25, C+) is considerably less than the mean GPA achieved by a non-defaulter (2.84, B-). Both defaulters and non-defaulters averaged approximately the same number of years at UD. As might be expected, the student's income and assets and the parent's income and assets were substantially lower on average for defaulters. Also, the range of these financial measures for defaulters was either equal to or significantly less than those for non-defaulters. In addition, the typical defaulter received less grant and scholarship aid and earned less money through on-campus employment than the typical non-defaulter. Still, the mean amount of loan aid accepted by the student did not vary across defaulters and non-defaulters. Further, the total amount of Perkins Loans accumulated by the student also did not differ between defaulters and non-defaulters on average.

It should be noted that a few variables were omitted from the analysis. The cumulative high school grade point average (another indicator of a student's academic performance) and the occupation and salary that a student attained after separation from the university could not be obtained. Occupation and salary could have significant explanatory value, but were unavailable for analysis. Additionally, SMARSTAT, REGYEARS, and MPAY had to be eliminated from further analysis as they are highly correlated (r^2 0.75) with other independent variables.

Discriminant Analysis Results

Using stepwise discriminant analysis to predict default, only nine of the original twenty-one variables help the equation accurately sort defaulters from nondefaulters (UDGPA, PASSET, DEPSTAT, ETHNIC, SINC, GRANTS, ENTRY, EARNED, and SCHOLAR), and five of these variables are statistically significant (UDGPA, PASSET, ETHNIC, SINC, and GRANTS). (Complete explanations for each acronym used are given earlier in this paper.)

The standardized coefficients displayed in Table 4 show the relative importance of each variable. The standardized coefficients are used to ascertain the relative contribution of the associated variable to the function by looking at the magnitude of these coefficients (ignoring the sign). The higher the magnitude, the greater that variable's contribution. In addition, the sign indicates whether the variable is making a positive or negative contribution. In this case, any variable with a negative coefficient contributes to default. Alternatively, any variable with a positive coefficient contributes to repayment.

As indicated in Table 4, UDGPA makes the greatest contribution in helping to determine who will default. Further, as UDGPA increases, the likelihood of default falls. PASSET and SINC also contribute to repayment. On the other hand, ETHNIC (1 = minority) and GRANTS increase the likelihood of default; these coefficients have negative signs. Most importantly, each of these coefficients is statistically significant at least at the 0.05 level.

Justification for these results could be posited. Individuals who achieve a higher GPA in college will probably be more likely to obtain a good job with

TABLE 3
Sample Description: Means and Standard Deviations For Continuous Variables*

Variable	Overall (n = 392) Mean (s.d.)	Non-defaulter (n = 315) Mean (s.d.)	Defaulter (n = 77) Mean (s.d.)	Range	
				Minimum	Maximum
University of Dayton (U.D.) Grade Point Average	2.72 (0.65)	2.84 (0.57)	2.25 (0.73)	0.000	4.000
No. of Years Registered at U.D.	3.158 (1.445)	3.184 (1.339)	3.052 (1.820)	0.000	12.000
Adjusted Gross Income of Students**	$3,656.19 ($5,142.51)	$3,922.69 ($5,537.31)	$2,565.97 ($2,806.14)	$0.00	$45,160.00
Student Assets**	$291.62 ($280.46)	$319.47 ($289.16)	$177.67 ($206.98)	-$406.00	$1,837.00
Grant Aid***	$2,643.98 ($1,650.29)	$2,539.37 ($1,602.06)	$3,071.95 ($1,782.28)	$0.00	$6,490.00
Merit-based Scholarship Aid***	$238.83 ($714.33)	$263.47 ($772.34)	$137.99 ($384.52)	$0.00	$9,000.00
Money Earned On Campus***	$713.87 ($921.22)	$745.96 ($905.24)	$582.59 ($979.13)	$0.00	$7,252.00
Value of All Loans Received for Students***	$3,147.63 ($1,552.85)	$3,134.56 ($1,583.95)	$3,201.10 ($1,427.03)	$0.00	$10,145.00
Value of All Perkins	$3,521.89	$3,551.73	$3,399.81	$140.00	$7,800.00
Loans While at U.D.	($1,792.28)	($1,779.11)	($1,852.05)		
Monthly Minimum Payment on Perkins Loans	$41.93 ($15.27)	$42.06 ($15.11)	$41.42 ($16.00)	$1.92	$120.00
Adjusted Gross Income of Student's Parents**	$27,614.22 ($18,823.98)	$28,689.07 ($18,617.43)	$23,217.10 ($19,146.26)	-$30,000.00	$92,585.00
Assets of Parents**	$29,161.63 ($28,045.58)	$31,946.93 ($28,915.76)	$17,767.22 ($20,698.32)	-$40,600.00	$183,700.00
Number of Dependents in Family	2.952 (1.733)	2.978 (1.694)	2.844 (1.892)	-1.000	10.000

*Variables with the possibility for consecutive values.

**During the year preceding the student's last year as an undergraduate.

***During the last year student was an undergraduate.

s.d. = Standard Deviation

a higher salary, which will aid them in making loan payments. Further, students with parents who have accumulated a significant amount of assets (cash, personal business, home, and other revenue) will more likely have a safety net in which to fall if they run into financial difficulty. Likewise, the more income a student earns (the majority usually earned over the summer months), the less money the student will need to borrow to pay for college expenses. Conversely, since minorities more often come from less affluent backgrounds than whites, being a minority can reasonably be seen as contributing to default. Further, if GRANTS are seen as an indicator of need such that only the neediest students receive grants, these needy students are also the poorest students and thus would be more likely to default.

"Entrance and exit interviews should stress more rigorously academic concerns as well as financial ones."

One final note is necessary about the coefficients. The negative coefficient on SCHOLAR is puzzling. The University of Dayton only gives out merit-based scholarships and anyone receiving one is expected to have a lower debt load throughout college and thus be in a better financial position to pay any debt. Nonetheless, this coefficient is not significantly different from zero and therefore does not pose a considerable theoretical challenge to the model.

Moreover, to judge the importance of the discriminant equation as a whole, one should look at a statistic called the square of the canonical correlation as well as the significance of another statistic called the Wilk's lambda. The square of the canonical correlation coefficient identifies the percent of the variation in the discriminant equation that is explained by the two groups. In this case, canonical correlation squared equals 0.1968458. This means that the discriminant equation, which includes only the nine variables in Table 4, accounts for only 19.7% of the variation in the groups, that is, the difference between defaulters and nondefaulters. In other words, 80.3% of the distinction between these two groups has not been determined by this model. In addition, Wilk's lambda is a measure of the how far apart the groups are based on the variables used. Values range from zero to one; the closer the value is to zero, the greater the distinction between the two groups (meaning that the group means are separate and distinct compared to the dispersion within each group) (Klecka, 1980). The Wilk's lambda is 0.8031542 which is fairly close to one. This confirms that the group means are somewhat close together; there is not a large difference between the groups. The overall significance of the equation is 84.066 (chi-square) with 9 degrees of freedom, which is significant at the 0.00 level. In other words, there is little possibility that the final results could have occurred by chance alone.

Finally, the activity of classification provides one last check of the overall significance of the discriminant function. Classification is performed on a case by case basis such that the discriminating variables are used to predict the group to which a case most likely belongs. A case is classified into the group that has the smallest squared distance from the case to the group mean. Further, these generalized distance measures can be converted into probabilities of group membership. A case is then assigned to the group for which it has the highest probability of membership (Klecka, 1980). By classifying the cases used to derive the function in the first place and comparing actual group membership with predicted group membership, one can empirically measure the success in discrimination by observing the percent of correct classifications. Table 5 indicates that the discriminant function did a good job of identifying group

TABLE 4
Discriminant Function Results

Variables	Standardized Coefficients
UDGPA	0.75752 *
PASSET	0.28225 *
DEPSTAT	-0.27297
ETHNIC	-0.26680 *
SINC	0.24777 ***
GRANTS	-0.22685 **
ENTRY	0.21067
EARNED	0.15948
SCHOLAR	-0.12286

*, **, and *** indicate significance at the 0.0001, 0.01, and 0.05 level, respectively.

These were the final nine variables in the discriminant equation used to distinguish between defaulters and nondefaulters.

TABLE 5
Final Prediction of Default Probability*

		Predicted Group Membership	
Actual Group	**No. of Cases**	**No Default**	**Default**
No Default	313	240 76.7%	73 23.3%
Default	77	23 29.9%	54 70.1%

Percent of "Grouped" Cases Correctly Classified: 75.38%.
*Example interpretation: based on the model used (the variables included in the equation), 76.7% of all non defaulters were classified correctly.

membership using only the values of the discriminating variables. It correctly classified 75.38% of the cases. For further interpretation of the table, one should note that of the 77 actual defaulters, 23 (29.9%) were incorrectly classified as non-defaulters, but 54 (70.1%) were correctly classified as defaulters.

Limitations and Conclusion

Several limitations of this study must be noted. The sample size for this study was constrained by data availability and could not be extended into more recent years because of the lag existing between the time when a student accepts a Perkins Loan and the time when a student begins repaying that loan. The introduction of new policies and the changing attitudes toward the severity of student loan default could change some of the aspects of the current model. Even though the model can be applied to new student borrowers, it must be used with caution.

Also, a few assumptions of discriminant function analysis were violated. See the Appendix for further discussion.

Overall, even though the model explained a small proportion of the variation in the groups (19.7%), the proportion that it explained was significantly different from zero. Most importantly, the model did a very good job of classifying cases; 75.38% of the cases were correctly classified by the model.

Although federal loan programs have gone through many policy changes recently, this study does emphasize one key issue. Since GPA was found to be the most significant contributor to repayment, entrance and exit interviews should stress more rigorously academic concerns as well as financial ones.

Much more can be said about default and the Perkins Loan Program. Since new requirements are being enacted yearly, it would be interesting to see if some of the more prominent changes truly have a significant effect on whether or not a student defaults. Alternatively, increased attention to the problem of default in general made universities more active loan collectors and this may be the primary reason for any reduction in default rates. In addition, the study of student loan defaults at UD could be expanded to other loan programs, such as the Stafford Loan Program. The data would certainly be more difficult to collect because it would be scattered throughout many financial institutions. The effort would be worth it, however, because the Stafford Loan Program is so much larger that there is greater potential for saving money by better managing defaults.

Appendix

The violation of the assumptions of a multivariate normal distribution on the discriminating variables and equal group covariance matrices can be tolerated. Several authors (in particular Lachenbruch, 1975) have shown that discriminant analysis is a robust technique which can sustain some deviation from these assumptions. Further, not all of the aspects of discriminant analysis require these assumptions. Tests of significance can be interpreted "conservatively" when the data violates the assumptions. Most importantly, if the main interest is in a mathematical model which can predict well or serve as a reasonable description of the real world, as it is in this study, the percentage of correct classifications serves as the best indicator of performance.

However, the fact that the discriminating variables must at least be measured at the ratio or interval level has yet to be addressed. This study included several binary variables. To see if these variables significantly altered the results, the analysis was also performed using continuous variables only (PARINC, PASSET, SINC, SASSET, DEP, GRANTS, SCHOLAR, LOANACC, EARNED, UDGPA, and LOANAMT). The results of the alternate analysis are similar to the analysis presented above. UDGPA was the largest significant contributor to repayment. Also, SINC made the same significant contribution as it did before. Therefore, the use of binary variables did not seem to hinder the basic analysis.

References

Greene, L. L. (1989). An Economic Analysis of Student Loan Default. *Educational Evaluation and Policy Analysis*, 11, 61-68.

Klecka, William R. (1980). *Discriminant Analysis*. Beverly Hills: Sage.

Lachenbruch, P.A. (1975). *Discriminant Analysis*. New York: Hafner.

Merisotis, Jamie P. (Ed.). (1991). *The Changing Dimensions of Student Aid.* San Francisco: Jossey-Bass.

U.S. Department of Education. (1993). *The Student Guide: Financial Aid from the U.S. Department of Education, 1994-1995.* Washington, DC.

U.S. General Accounting Office (GAO). (1991, April). *Student Loans: Characteristics of Defaulted Borrowers in the Stafford Student Loan Program, Briefing Report to the Chairman, Subcommittee on Education, Arts and the Humanities; Committee on Labor and Human Resources, U.S. Senate.* Washington, DC: Author.

U.S. General Accounting Office (GAO). (1990, September). *Defaulted Student Loans: Analysis of Defaulted Borrowers at Schools Accredited by Seven Agencies, Fact Sheet for Congressional Requesters.* Washington, DC.

Wilms, W. W., Moore, R. M. & Bolus, R. E. (1987). Whose Fault Is Default? A Study of the Impact of Student Characteristics and Institutional Practices on Guaranteed Student Loan Default Rates in California. *Educational Evaluation and Policy Analysis*, 9, 41-54.