



December 2006

Multivariate Analysis
of Student Loan
Defaulters at
Prairie View A&M University

Conducted by TG Research
and Analytical Services

Sandra Barone



TG™

Multivariate Analysis of Student Loan Defaulters at Prairie View A&M University

TG Research and Analytical Services
Sandra Barone

December 2006

Multivariate Analysis of Student Loan Defaulters at Prairie View A&M University

TG Research and Analytical Services
Sandra Barone

Table of Contents

Executive Summary and Highlights	3
Introduction	5
Prior Research on the Factors Relating to Student Loan Default	6
College Performance Variables	6
Demographic Variables	7
Pre-College Variables	8
Financial Aid and Loan Related Variables	8
Methodology for Multivariate Analysis of Defaulters at PVAMU	9
Variable Selection Process	10
Results of the Multivariate Analysis	11
Grade Point Average (GPA).....	14
Expected Family Contribution (EFC).....	14
Degree Indicator.....	15
Highest Academic Level Attained at PVAMU.....	15
High School Class Rank Percentile	15
Number of Hours Transferred to PVAMU	16
Model Performance	16
Distribution of probabilities.....	16
K-S Statistic	17
C Statistic	18
Receiver Operating Characteristic (ROC) Curve	18
Classification Matrix and Misclassification Rate	19
Uses of the Findings and the Model and Areas for Future Research	20
APPENDIX A: Sample Definition and Variable Descriptions	24
TABLE A1: Characteristics of Undergraduate Borrowers at PVAMU.....	28
By Default Status	
APPENDIX B: Standard Errors and Confidence Intervals	33
Bibliography	34

Multivariate Analysis of Student Loan Defaulters at Prairie View A&M University

TG Research and Analytical Services
Sandra Barone

Executive Summary and Highlights

This study examines the default behavior of 3,325 undergraduate student borrowers who attended Prairie View A&M University (PVAMU) and entered repayment on their TG-guaranteed Federal Family Education Loan Program (FFELP) loans between October 1, 2000 and September 30, 2002 (fiscal years 2001 – 2002). Using the Department of Education's official cohort default rate formula, 624 borrowers, or 18.8 percent, were in default. The study uses a statistical technique called multiple logistic regression to analyze the effects of individual student and family characteristics on the probability of default, while controlling for the effects of the other variables in the analysis.

For students at PVAMU, success in college is the key to preventing student loan defaults. The three measures of success included in the final model: 1) persisting in college beyond the freshman year, 2) performing well while at the university, and 3) obtaining a degree, are all extremely important in decreasing a student's likelihood of default. These results suggest that expanding or intensifying campus-wide efforts to increase retention rates and providing both financial aid and academic counseling to students who perform poorly academically may be the most effective default aversion strategy for PVAMU.

These college success variables provide almost all of the predictive power in the final model. However, the study finds that even after controlling for a student's success at PVAMU, a student's high school preparation and family income also have significant relationships to a student's probability of default. The students who are the least prepared academically (as measured by high school class rank percentile) and those who have the fewest financial resources (as measured by expected family contribution) are the most likely to default on their student loans (and the least likely to obtain a college degree). These students may require additional assistance in order to successfully transition to college life and will continue to need counseling and support in order to remain in school and complete their degrees.

More specifically, the key findings of this study are:

- College grade point average (GPA) is strongly related to whether or not a borrower defaults on his or her student loan after leaving college. Borrowers who leave PVAMU with a GPA of 3.0 or higher have a likelihood of default which is at least seven percentage points lower than those who exit with a GPA of 2.5 or less, holding all other borrower characteristics constant.
- Students who obtain their degree have a likelihood of defaulting which is five percentage points lower than those who leave PVAMU without a degree.

- Borrowers who leave PVAMU after their freshman year are more likely to default on their student loans than those who persist in school, even to their sophomore year. Those who remain until their senior year have an even lower probability of default, regardless of degree attainment.
- Students who are unable to contribute financially to their education (as measured by an Expected Family Contribution of zero) are more likely to default on their student loans than students with even a small contribution.
- Students who graduated in the bottom 25 percent of their high school class have a probability of default which is four percentage points higher than those who graduated in the middle of their high school class.

Multivariate Analysis of Student Loan Defaulters at Prairie View A&M University¹

TG Research and Analytical Services
Sandra Barone

Introduction

The Federal Family Education Loan Program (FFELP) makes it possible for millions of students to obtain post-secondary education each year who would otherwise be unable to afford to attend. In fiscal year 2005 alone, students borrowed over \$50 billion through this program.² Historically, however, default rates have been high under this program.³ In fiscal year 2000, defaulted student loans cost taxpayers over \$2 billion.⁴ These defaulted student loans also hurt students' credit ratings. Because of the high costs of student loan defaults, both to the student borrower and to the taxpayers, the Department of Education sanctions schools with exceptionally high percentages of defaulters. In an effort to better understand which students are likely to default, and ultimately to design programs to reduce the number of borrowers who do default, TG⁵ and Prairie View A&M University (PVAMU) have agreed to work together to perform an analysis of student loan defaulters at PVAMU. Information obtained from this study can be used to target at-risk borrowers and lower default rates; not only by TG and PVAMU, but also by other schools and guarantors, lenders, and servicers.

This study examines the behavior of 3,325 undergraduate students who attended PVAMU and entered repayment of their student loans between October 1, 2000 and September 30, 2002 (fiscal years 2001 – 2002). Following the Department of Education's official cohort default rate formula, a borrower is considered to be in default if he or she defaults during the fiscal year that the borrower entered repayment or within the following fiscal year. Using this definition, 624 borrowers, or 18.8 percent, defaulted.

PVAMU provided detailed data on this sample, including information on the students' high school performance, college coursework and performance, and demographic information. This study closely follows the methodology used in similar studies performed by TG in conjunction

¹ The author would like to acknowledge the financial support of the U.S. Department of Education and to thank Jeff Webster, Matt Steiner and Marlena Creusere for support throughout. Any remaining errors are the author's alone.

² U.S. Department of Education, *Federal Family Education Loan Program (FFELP) – Annual and Cumulative Commitments-FY66—FY2005*. <http://www.ed.gov/finaid/prof/resources/data/05ffelpga.xls>

³ The official cohort default rate reached an all time high of 22.4 percent with the 1990 cohort. U.S. Department of Education, *Briefing on National Default Rates*, September 13, 2006.

<http://ifap.ed.gov/eannouncements/attachments/0913CDRBriefingAttach.pdf>

⁴ U.S. Department of Education. *Table 49.—Federal Family Education Loan (FFEL) program annual and cumulative default dollars and collections: FY86-FY00*. <http://www.ed.gov/finaid/prof/resources/data/fslpdata97-01/table49.xls>

⁵ TG is a public, nonprofit corporation that helps create access to higher education for millions of families and students through its role as an administrator of the Federal Family Education Loan Program (FFELP). Its vision is to be the premier source of information, financing, and assistance to help all families and students realize their educational and career dreams. Additional information about TG can be found online at www.tgslc.org.

with Texas A&M University at College Station, the University of South Florida, and Texas A&M University - Kingsville. These studies use a statistical technique called multiple logistic regression to analyze the effects of individual student and family characteristics on the probability of default, while controlling for the effects of the other variables in the analysis.

Prior Research on the Factors Relating to Student Loan Default⁶

The first studies of student loan default behavior were undertaken in order to analyze the policy of holding schools responsible for borrower defaults. Therefore, many of those studies evaluated the relative importance of borrower and institutional characteristics. Several found that institutional characteristics have little or no association with loan repayment behavior and that borrower characteristics are much more important predictors of default (Knapp & Seaks, 1990; Volkwein & Szelest, 1995; Volkwein et. al., 1995; Wilms, Moore & Bolus, 1987). Because the present analysis of borrowers at PVAMU concerns the default behavior of students at one institution, prior work on the influence of institutional characteristics is of little relevance.

This early research did analyze many borrower characteristics that are relevant to the present study. Also, more recent work has focused directly on these borrower characteristics. These factors include demographic descriptors (such as ethnicity or race, gender, age and income), financial aid-related variables (like financial need and expected family contribution), college performance variables (such as college GPA, graduation status and number of courses failed) and some high school-related variables (like ACT scores and whether the borrower has a high school diploma).

College Performance Variables

The most consistent finding of past studies is that borrowers who complete school (as measured by graduating, earning a degree, or not withdrawing) have a much lower probability of defaulting on their loans than borrowers who do not complete school (Barone, Steiner & Teszler, 2005; Dynarksi, 1994; Knapp & Seaks, 1990; Meyer, 1998; Podgursky et. al., 2000; Steiner & Teszler, 2005; Steiner & Tym, 2005; Volkwein & Szelest, 1995; Volkwein et. al., 1995; Wilms, Moore & Bolus, 1987; Woo, 2002). For many of these studies, graduation (or completion) status was the single most important variable in predicting student loan default behavior.

Studies that have included a borrower's grade point average (GPA) in college have consistently found that a higher college GPA is associated with a lower probability of default (Volkwein et. al. (1995); Volkwein & Szelest, 1995; Barone, Steiner & Teszler, 2005; Steiner & Teszler, 2005; Steiner & Tym, 2005). This result holds, even after controlling for graduation status in multivariate analyses.

Prior studies have also tested the impact of other variables related to a borrower's success in college on the students' probability of defaulting on their student loans. Much of this research is

⁶ For a more comprehensive review of student loan default research, see TG's Student Loan Default Literature Review, McMillion (2004) available at <http://www.tgslc.org/publications/index.cfm> under Literature Reviews.

inconclusive. Several studies have included the student's major in the analysis. Volkwein et. al. (1995) found that borrowers who majored in science or technology had a significant but relatively small decrease in the probability of default. Among borrowers at Texas A&M University, Steiner and Teszler (2005) found that students who attended a college other than the Liberal Arts college were slightly less likely to default on their student loans. In contrast, Woo's (2002) study indicated that whether or not a borrower studied a business or computer curriculum did not have a significant association to default.

Past research suggests that students who transfer from one institution to another have a slightly lower probability of defaulting on their student loans. However, the impact of this variable is also inconclusive. Volkwein et al. determined that a variable signifying that the borrower was a transfer student did not have a significant relationship to default. A related study by Volkwein and Szelest (1995) had similar results. Steiner and Teszler, however, found that being a transfer student had a small but significant decrease in the probability of default. Woo (2002) established that borrowers who attended more than one school were less likely to default. She also found that attainment of a graduate or professional degree greatly reduces the chances of default. (Woo noted that the impact of attending more than one school partially reflects the fact that borrowers who go to graduate school frequently have attended more than one school.)

Past research does agree that remaining in school and performing well in school both decrease a borrower's probability of default. Steiner and Tym (2005) and Steiner and Teszler (2005) both found that borrowers who failed more than ten hours of coursework were significantly more likely to default on their loans than those who did not fail any courses. Meyer (1998) found that as the academic level attained by a borrower increases, the probability of default decreases. This result was confirmed by both Barone, Steiner & Teszler (2005) (using the same measure of persistence) and by Steiner and Tym, who measured persistence by total hours attended.

Demographic Variables

Multivariate default studies have consistently found that ethnicity/race is strongly related to default (Barone, Steiner & Teszler, 2005; Dynarksi, 1994; Knapp & Seaks, 1990; Podgursky et. al., 2000; Volkwein & Szelest, 1995; Volkwein et. al., 1995; Wilms, Moore & Bolus, 1987; Woo, 2002). In particular, being African-American greatly increases the probability of default. In four of the studies (Barone, Steiner & Teszler, 2005; Volkwein & Szelest, 1995; Volkwein et. al., 1995 and Woo, 2002), being African-American had the largest effect of all variables, and in the remainder of the cited studies, being African-American was the second most influential factor.

Past studies have also found that gender is significantly related to default behavior, with female students being less likely to default than male students (Barone, Steiner & Teszler, 2005; Podgursky et. al., 2000; Steiner & Teszler, 2005; Steiner & Tym, 2005; Volkwein et. al., 1995; Woo, 2002). However, both Knapp & Seaks (1992) and Volkwein and Szelest (1995) found that the borrower's gender was not significantly related to the borrower's likelihood of default.

Family financial resources have also been included in many models, measured in slightly different ways. Higher parental income levels, total family income, expected family

contribution, and student income levels have all been found to be associated with decreases in the probability of default (Barone, Steiner & Teszler, 2005; Dynarksi, 1994; Knapp & Seaks, 1992; Steiner & Teszler, 2005; Steiner & Tym, 2005; Volkwein et. al., 1995; Wilms, Moore & Bolus, 1987; Woo, 2002).

Podgursky et. al., Woo, Meyer, Steiner & Teszler and Steiner & Tym all examined the age of the borrower and found it to have a significant but small effect on default behavior, with increases in age related to higher probabilities of defaulting. In contrast, Knapp & Seaks and Barone, Steiner & Teszler could not detect a statistically significant relationship for the age of the borrower.

Other demographic variables that researchers have found to significantly increase a borrowers probability of default are not having two parents at home, (Knapp & Seaks, 1992), having parents who did not attend college (Volkwein et. al., 1995), having parents who did not complete high school (Steiner & Teszler, 2005; Steiner & Tym, 2005), being Hispanic (Barone, Steiner & Teszler, 2005; Dynarksi, 1994; Woo, 2002), having dependents (Dynarksi, 1994; Volkwein & Szelest, 1995; Volkwein et. al., 1995; Woo, 2002), being an unmarried borrower (Dynarksi, 1994; Volkwein & Szelest, 1995; Volkwein et. al., 1995), and having low post-college borrower income (Dynarksi, 1994; Volkwein & Szelest, 1995; Volkwein et. al., 1995; Woo, 2002).

Pre-College Variables

Some researchers have evaluated characteristics reflecting the borrower's experience before college. These data are not readily available, so it has not been analyzed as frequently as some of the variables discussed above. Several studies have found that graduation from high school reduces the likelihood of default (Dynarksi, 1994; Volkwein et. al., 1995; Wilms, Moore & Bolus, 1987 and Woo, 2002). However, Volkwein and Szelest did not detect a significant relationship between having a high school diploma and default behavior. Barone, Steiner & Teszler found that students who graduated in the top twenty percent of their high school class were less likely to default than students who graduated in the middle of their class. Podgursky et. al. found that a borrower's ACT scores have a small negative effect on the probability of default.

Financial Aid and Loan Related Variables

Financial aid practioners and researchers are greatly interested in the effect of the amount and type of financial aid received on a borrower's probability of default. The most commonly tested financial aid related variables are family income and assets, which were discussed above. Volkwein et. al. tested several other financial aid-related variables – such as whether the borrower received scholarships/grants, participated in work study or had other employment – but found none of them to be significant. Meyer, however, determined that the probability of default declined with increases in the cost of attendance, controlling for the type of institution. He further discovered that the likelihood of default increased substantially for borrowers who received more than \$1,000 from non-loan aid sources. He noted a small decrease in the chances of defaulting as the expected family contribution of borrowers increased.

Many researchers have looked at the relationship between the total amount of loans borrowed and the risk of default. Several analyses determined that there was not a statistically significant relationship between the amount of loans borrowed and default behavior (Barone, Steiner & Teszler, 2005; Knapp & Seaks, 1992; Steiner & Teszler, 2005; Steiner & Tym, 2005; Volkwein & Szelest, 1995; Volkwein et. al., 1995; Woo, 2002). In fact, at a bivariate level, Barone et al. found that students with more debt were less likely to default. In this case, the total debt amount was a proxy for length of time in school. Once the college success variables were included in the multivariate analysis, the total loan amount did not have a significant impact on the probability of default. Meyer, however, found that each \$1,000 of total debt increases the probability of default by about one percentage point. Dynarski determined that the probability of default increases with increases in the size of borrowers' monthly loan payments. Furthermore, Woo detected a small increase in the likelihood of default associated with an increase in the number of loans a borrower has. Meyer also examined the types of federal loans that borrowers received and showed that borrowers with only subsidized Stafford loans had the highest probability of default. He also found that borrowers who utilized deferments had a somewhat smaller chance of defaulting.

Methodology for Multivariate Analysis of Defaulters at PVAMU

The present study follows closely the methodology used in TG's recent multivariate analyses of borrowers at Texas A&M University – College Station (TAMU), University of South Florida (USF), and Texas A&M University – Kingsville (TAMUK).⁷ The studies use the logistic regression method of multivariate analysis. This is the appropriate method to use when analyzing an outcome which can assume one of two classes, like defaulting or not defaulting. The statistical analysis proceeds by determining the relationships between borrower characteristics and default behavior within a past population of borrowers. The known outcomes (i.e., default behaviors) of this population serve as the basis for statistical estimation. The result of the analysis is a set of coefficients or weights. The logistic regression method chooses the set of weights that would produce predictions of default that match as closely as possible to the known outcomes of default. The sign (plus or minus) of a coefficient indicates whether the presence of the characteristic increases or decreases the likelihood of default, and the size of a coefficient generally reflects the strength of the relationship between the characteristic and the occurrence of default. It is important to remember that the coefficients reported measure the effect on a borrower's likelihood of default when a specific characteristic is changed, *holding all other characteristics constant at their reference level*.

⁷ These studies are available at <http://www.tgslc.org/publications/index.cfm>.

Variable Selection Process

The multivariate model presented below measures the impact of each of the included variables on a borrower's probability of default, after controlling for the effect of the other variables in the model. This model contains only a subset of the total number of variables that were gathered for the study. Appendix A contains a complete list of the variables gathered for this study and their bivariate relationship to the probability of default. The goal of the final model is to present the subset of variables that, when taken together, best explains default behavior.

There are several reasons why a variable is excluded from the final model. The first is if the variable does not have a statistically significant relationship to default behavior. Highest degree earned is an example of such a variable. While borrowers who obtain their degree are much less likely to default than borrowers who do not complete a degree, the type of degree earned does not appear to be related to the probability of a borrower defaulting on his or her student loan.

The second, and most common, reason that a variable is excluded from the final model is that it is measuring the same borrower characteristic as another variable. In such a case, the two variables are highly correlated. If two or more highly correlated variables are included in a multivariate model, only one will prove to be significant. This is because the multivariate model describes the impact of each variable on a borrower's probability of default *after controlling for the impact of the other variables in the model*. Many of the variables gathered for this study are highly correlated. For example, Table A1 describes the relationship of many variables which measure college success to a borrower's probability of default at PVAMU. The variables Highest Academic Level at PVAMU, Total Hours Attempted and Transferred, Number of Hours Passed, and Number of Hours Failed all appear to be strongly related to a borrower's probability of default. However, these variables are also highly correlated with each other. In other words, borrowers who reach a higher academic level also tend to complete more hours and fail fewer hours than those who don't remain in school as long. Therefore, only Highest Academic Level at PVAMU is included in the final model.

Other reasons that variables may be excluded from the multivariate model are that there is little variation in the variable, or that a large percentage of the sample is missing data for a particular variable. Ethnicity is an example of a variable with little variation at PVAMU. Since over 96 percent of the borrowers in our sample are African-American, this variable is excluded from the model, even though it has proven to be significant in past research. Marital Status is another variable which has been shown to be related to default behavior in other studies, but is excluded from the final model at PVAMU since 87 percent of the sample is missing this data.

Sometimes a variable is so important from a theoretical or practical standpoint that it is included in the final model, even if it is not found to be significant. Number of hours transferred is an example of such a variable at PVAMU. Incorporating all of these considerations, the final default model is the combined result of statistical relevance, theoretical importance, data availability and human judgment.

Results of the Multivariate Analysis

The multivariate analysis produced a default model containing the variables listed in Table 1. The table lists each variable, its reference group, the coefficient, and the change in probability of default, each of which will be explained below.

The impacts of the variables in the model are all measured in relation to a reference group for the variable. The multivariate estimation process produces a coefficient for each variable. This coefficient measures the impact of the variable on a student's likelihood of default, as compared to the reference group, when the values of all other variables are held constant at their reference group. The sign (positive or negative) of a coefficient indicates whether the presence of the variable increases or decreases the likelihood of default, and the size of a coefficient generally reflects the strength of the relationship between the variable and the occurrence of default.

The presence of an asterisk next to a coefficient indicates that the variable has a statistically significant relationship to default behavior. Statistical significance means that there is a relatively high confidence that a relationship really exists – that the size of the coefficient did not result from the peculiarities of the sample that we analyzed. The more asterisks there are, the higher the level of confidence that a true relationship exists between a variable and default behavior.

Unfortunately, due to the non-linearity of the logistic regression model, the coefficients are difficult to interpret in their raw form. In order to more easily understand their meaning, it is necessary to convert them to another form. The last column of the table represents the percentage point change in the probability of default given the presence of a characteristic, when all other characteristics are measured at their reference point. This change is only reported for variables that are statistically significant.⁸

For example, the variable Grade Point Average (GPA) = 3.01 - 4.00 has a coefficient equal to -0.989. This means that a student with a GPA in this range has a lower likelihood of defaulting on his or her student loan within the two year cohort period than a student with a GPA between 2.01 and 2.50, the reference group. The presence of three asterisks next to the coefficient indicates that there is a 99.9 percent degree of confidence that students with the higher GPA have a lower likelihood of defaulting than students with an average GPA. In other words, there is only a one tenth of a percent chance that the difference in coefficients results from a particular characteristic of this sample, rather than representing a true relationship.

Looking only at this raw coefficient, it is difficult to translate -0.989 into an effect on a student's default rate. The last column of the table assists in interpreting the coefficients. A baseline probability of default was calculated, based on the model when all variables are valued at their reference group. This baseline probability of default is 12 percent. The change in probability column gives the change in probability of default from this baseline due to moving a variable from its reference group to the indicated value. For example, a student with all variables measured at their reference group has a 12 percent probability of default. However, if this student has a GPA in the range 3.01 to 4.00, as opposed to 2.01 to 2.50, this probability of

⁸ For those who would like to calculate additional measures of significance, Appendix B includes a table containing the standard errors of the coefficients and confidence intervals for the change in probability.

default drops by seven percentage points to five percent. A discussion of the impact of each group of variables follows the table.

It is important to realize that the change in probability reported in Table 1 represents the change in probability only when all other variables are measured at the value of their reference group. To illustrate, consider a student who stayed at PVAMU through his or her senior year, but did not obtain a degree. If this student had a GPA of 2.01 to 2.50, his or her probability of default is eight percent (four percentage points lower than the baseline probability of 12 percent). However, if this student had a GPA of 3.01 to 4.00, his or her probability of default would not drop an additional seven percentage points to only one percent. It would only drop five percentage points to three percent. This is due to the non-linearity of the logistic regression model. As one approaches the tails of the default probability distribution, the impact of any one individual variable is greatly reduced. In other words, for students with either extremely high or extremely low probabilities of default, changing the value of one variable in the model will have a negligible effect on the borrower's overall probability of default.

When interpreting these results it is important to remember that there is always uncertainty in any statistical model. The results of a statistical analysis tend to best describe the sample from which they were produced. Therefore, care must be taken when generalizing the results of any particular study. Despite these general limitations, there is a great deal of information to be learned from this specific study. The results of this multivariate study of students at PVAMU are very robust,⁹ providing a great deal of confidence in their applicability to current and future students. The value of the model is not in predicting that students with GPAs above 3.0 have a probability of default that is exactly seven percentage points less than that of students with a GPA between 2.0 and 2.5. Rather, the value of the model is that it provides a high level of confidence that a student's GPA provides a significant amount of information about that student's probability of default, even after controlling for the student's other characteristics.

Following the table is a discussion of each of the variable groups included in the model. The variables are discussed in roughly the order of their strength of association to default.

⁹ A robust statistical model is one in which the results are not highly dependent on the variables included in the model. The current study is very robust. The author ran several versions of the model and found that the results were highly consistent across the various models.

TABLE 1
Results of Multivariate Analysis
Prairie View A&M University Undergraduates

Variable	Value	Reference Group	Coefficient	Change in Probability
	Intercept		-2.000 ***	
College Success Variables				
Degree Indicator	Has Degree	No Degree	-0.609 ***	-5%
Grade Point Average	0-1.00	2.01-2.50	0.536 **	7%
	1.01-2.00	2.01-2.50	0.274 *	3%
	2.51-3.00	2.01-2.50	-0.278	
	3.01-4.00	2.01-2.50	-0.989 ***	-7%
Highest Academic Level Attained at PVAMU	Freshman	Sophomore	0.371 **	4%
	Junior	Sophomore	-0.228	
	Senior or Higher	Sophomore	-0.510 *	-4%
Attendance Pattern				
Number of Hours Transferred	One or more hours	Zero	-0.186	
College Preparedness				
High School Class Rank Percentile	Less than 25%	50% - 69%	0.322 *	4%
	25% - 49%	50% - 69%	0.125	
	70% - 89%	50% - 69%	-0.126	
	90% - 100%	50% - 69%	0.068	
	Missing	50% - 69%	0.255	
Financial Aid Variables				
Expected Family Contribution (EFC)	Zero	\$1-750	0.461 **	6%
	\$751-2,000	\$1-750	-0.298	
	\$2,001-5,000	\$1-750	-0.071	
	\$5,001 and higher	\$1-750	-0.025	
	Missing	\$1-750	-0.304	

Sample Size: 3,320 Defaulters: 622 (18.7 percent)
-2 log likelihood: Intercept and covariates: 2,699
Chi-Square: 503.93 with 19 degrees of freedom (Pr > ChiSq = <.0001)
C Statistic: 77.0 percent
Baseline probability of default (intercept only): 11.9 percent
* Statistically significant at the 0.05 level
** Statistically significant at the 0.01 level
*** Statistically significant at the 0.001 level

Grade Point Average (GPA)

The higher a student's grade point average is, the less likely the student is to default on his or her student loan, after controlling for the other variables in the model. Borrowers with a GPA between 3.01 and 4.00 have a probability of default which is seven percentage points lower than borrowers with a GPA between 2.01 and 2.50. Borrowers with a GPA between 2.51 and 3.00 have the same probability of default as borrowers with an average GPA. However, borrowers with a low GPA of 0 to 1.00 have a probability of default which is seven percentage points higher than borrowers with a GPA between 2.01 and 2.50, and borrowers with a GPA of 1.01 to 2.00 have a probability of default which is three percentage points higher than borrowers with an average GPA. This result is especially useful for a school — such as PVAMU — that is interested in lowering its default rate. It is relatively simple for a financial aid office to obtain information about a student's GPA. This result suggests that by monitoring students with low GPAs and providing additional financial aid counseling to these students, PVAMU may be able to lower its cohort default rate.

These results of the multivariate analysis confirm the relationship between GPA and default as noted in the bivariate table in Appendix A. That table shows that borrowers with a GPA of 3.01 to 4.00 have a default rate of only 3.2 percent, whereas borrowers with a GPA of 0 to 1.00 have a default rate of 38.4 percent. The multivariate analysis reveals that even after controlling for factors such as obtaining a degree, persistence in college and other background variables, there is a strong relationship between a borrower's GPA and his or her probability of default.

There are many reasons why a higher GPA may lead to a lower probability of default. It is likely that this variable measures personal characteristics such as conscientiousness, persistence, motivation, intelligence and discipline which lead to success both in college and in loan repayment after college. Students who perform well in college are also more likely to complete their degrees and may earn more after college, making it easier to repay their student loans.

Expected Family Contribution (EFC)

Borrowers with an Expected Family Contribution of zero have a probability of default that is six percentage points higher than borrowers whose families are able to contribute to the cost of attending PVAMU. However, the amount the family is able to contribute does not affect a borrower's probability of default.

In general, theory suggests that higher amounts of expected family contribution are associated with higher family incomes. For independent students, the EFC represents resources that are directly available to repay student loans. For dependent students, the EFC represents the income of a student's parents. These results suggest that students whose families were able to contribute toward the cost of obtaining higher education might have more financial resources available to them in times of repayment difficulties, making it less likely that they will default on their student loans.

It is also likely that students with a zero EFC find it necessary to work while in school in order to pay for their education. The U.S. Department of Education has found that students who work full-time while in school are at a higher risk of dropping out without completing their education – also putting them at a higher risk of defaulting on any loans they may have taken out while in school.¹⁰

Degree Indicator

Students who obtain a degree from PVAMU, or arrive at PVAMU with a degree from another institution, have a likelihood of defaulting that is five percentage points lower than a borrower who does not obtain a degree, all other things being equal. Looking at the frequency table in Appendix A, we see that the impact of obtaining a degree is even stronger at the bivariate level. Only 35.6 percent of the sample of borrowers from PVAMU obtained a degree. However, of those borrowers who did obtain a degree, only 3.5 percent defaulted on their student loans. Borrowers who did not obtain a degree had a default rate of 27.2 percent. The results of the multivariate analysis show that this highly significant relationship between graduation status and default holds true, even after controlling for other student characteristics. Though financial aid officers might have little direct impact on whether or not borrowers complete their degrees, this variable might assist them in identifying at-risk borrowers (i.e., the ones who do not complete a degree program) toward whom they can direct default aversion efforts.

Highest Academic Level Attained at PVAMU

Staying in school longer reduces a borrower's probability of defaulting on a student loan, even after controlling for a student's grade point average and whether or not he or she obtained a degree. Students who remain at PVAMU through their senior year have a probability of defaulting that is four percentage points lower than students who leave after their sophomore year. By contrast, students who leave after their freshman year are four percentage points more likely to default on their loans than students who remain even one more year. As noted in Appendix A, this difference is even greater at the bivariate level. Those who left PVAMU after their freshman year had a default rate of 33.9 percent, whereas those who stayed through their senior year had a default rate of only 5.3 percent. The data show that persistence in school is directly related to student loan default behavior.

High School Class Rank Percentile

This variable is a measure of a student's preparedness for college. Theory predicts that students who are better prepared for college will have an easier time adjusting to college life and will be more successful in college. Since we know that college success is a strong factor in default aversion, it follows that high school success will help to predict default behavior. The median student at PVAMU graduated between the 50th and 69th percentile of his or her high school class.

¹⁰ See U.S. Department of Education (1996). Also, for more information on the impact of working while in school, see McMillion (2005) http://www.tgslc.org/pdf/HEA_Work_Loans.pdf.

Therefore, this group was used as the reference group. The only group of students that had a likelihood of default which is statistically different from this reference group is the group of students who graduated in the bottom 25 percent of their high school class. While this group makes up less than 13 percent of the sample of students, it is important to know that these students may require extra counseling in order to be successful – both in their academics, and in repaying their student loans.

Number of Hours Transferred to PVAMU

In this final model, the number of hours a borrower transferred to PVAMU does not have a significant effect on his or hers probability of default. However, in earlier models and in other research, this variable was found to be significant. In preliminary models which did not include the Highest Academic Level Attained variable, students who transferred hours to PVAMU were less likely to default on their student loans than students who did not transfer hours. However, when a borrower's highest academic level is added to the model, the number of hours transferred is no longer significant. This suggests that the presence of transfer hours is an indication of the borrower's persistence in school, with those borrowers who persist in school being less likely to default on their student loans than those who do not persist.

Model Performance

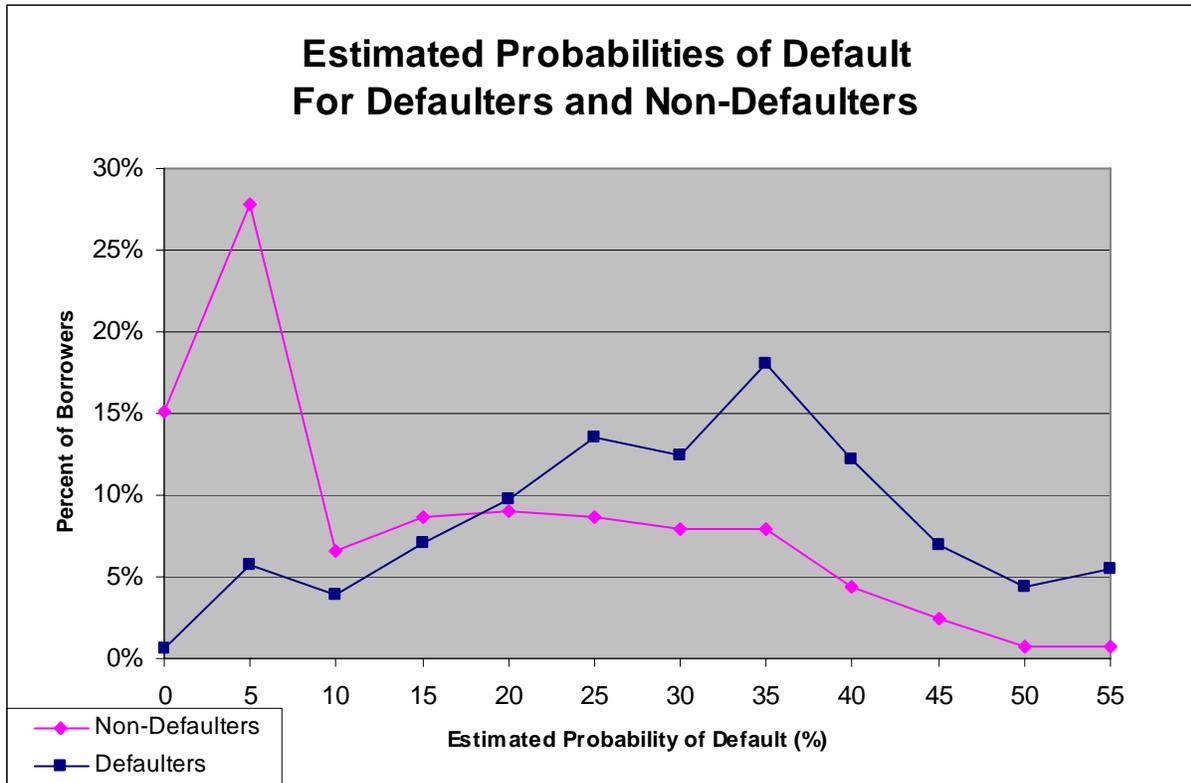
Based upon the characteristics of a borrower, it is possible to sum the coefficients for the variables in the prior section and to convert that sum to a probability that the borrower will default. The estimated probability can then be compared to the known outcome for the borrower. This comparison can be made for all borrowers in the study in order to gauge the performance of the multivariate model. In general, the performance measures in this section assess how well the statistical model correctly classifies defaulters and non-defaulters.

The performance measures indicate that this statistical model performs very well. It does a good job in assigning high probabilities of default to borrowers who actually defaulted and low likelihoods of default to borrowers who did not actually default.

Distribution of probabilities

The following chart shows the default probabilities assigned by the multivariate model to borrowers in the study. The chart provides a separate distribution of probabilities for actual defaulters and actual non-defaulters. (Each borrower's estimated probability of default was rounded to the nearest five percent.) The vertical axis shows the percentage of borrowers who were assigned each probability. Thus, whereas the model assigned estimates of a five percent (rounded) probability of default or less to 43 percent of actual non-defaulters (28 plus 15), it assigned a five percent (rounded) probability of default or less to only seven percent of actual defaulters (six plus one). In general, if the model is performing well, the curve for the non-defaulters should be higher than the curve for the defaulters on the left side of the chart.

Similarly, the curve for the defaulters should be higher than the curve for the non-defaulters on the right side of the chart. The visual impression of this chart is that the model appears to have performed well.

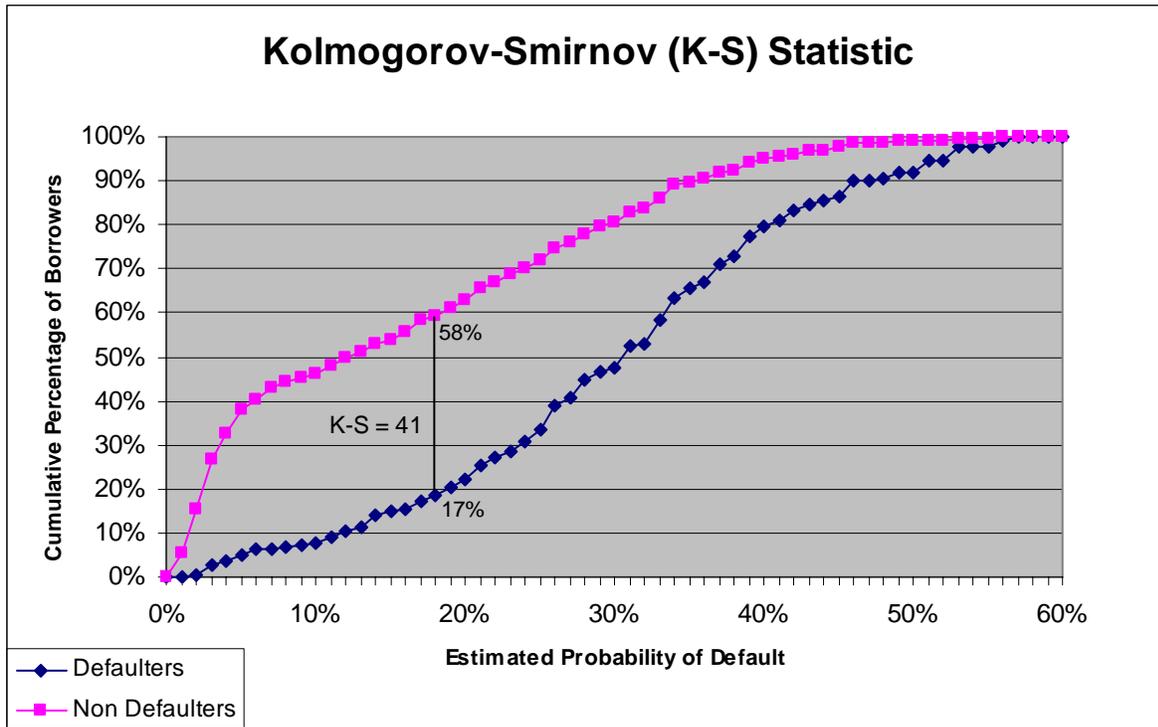


K-S Statistic

The previous distributions can be transformed into a set of cumulative distributions. Cumulative distributions give the percentage of borrowers who have an estimated probability that is equal to, or less than, a given point along the horizontal axis. For example, the chart below shows that 58 percent of actual non-defaulters have an estimated probability of default that is less than or equal to 17 percent and that only 17 percent of actual defaulters have an estimated probability of default in that range. As it turns out, at 17 percent (along the horizontal axis), the curves for defaulters and non-defaulters are separated by the greatest distance. This distance is known as the Kolmogorov-Smirnov (K-S) statistic. For the present model, the K-S statistic is 41 percent (58%-17%). Models with large K-S statistics are said to have done a good job of distinguishing between defaulters and non-defaulters. Forty-one percent (41%) is a high K-S statistic and indicates that the model does well in separating defaulters and non-defaulters.

A high K-S means that a model will predict default outcomes for a much higher percentage of actual defaulters than non-defaulters. Suppose this model predicted default for borrowers to whom the model assigned a default probability greater than 17 percent. The K-S of 41 percent indicates that using 17 percent as the prediction cutoff means that this model will predict default 41 percent more frequently for defaulters than for non-defaulters. At 17 percent, the model

would predict 83 percent of actual defaulters to default (that is, one minus the 17 percent with probabilities less than or equal to 17 percent), but it would only predict 42 percent of actual non-defaulters to default (one minus the 58 percent with probabilities less than or equal to 17 percent). The K-S statistic illustrates that this particular model does an excellent job of predicting default, while its ability to predict who will not default is weaker.



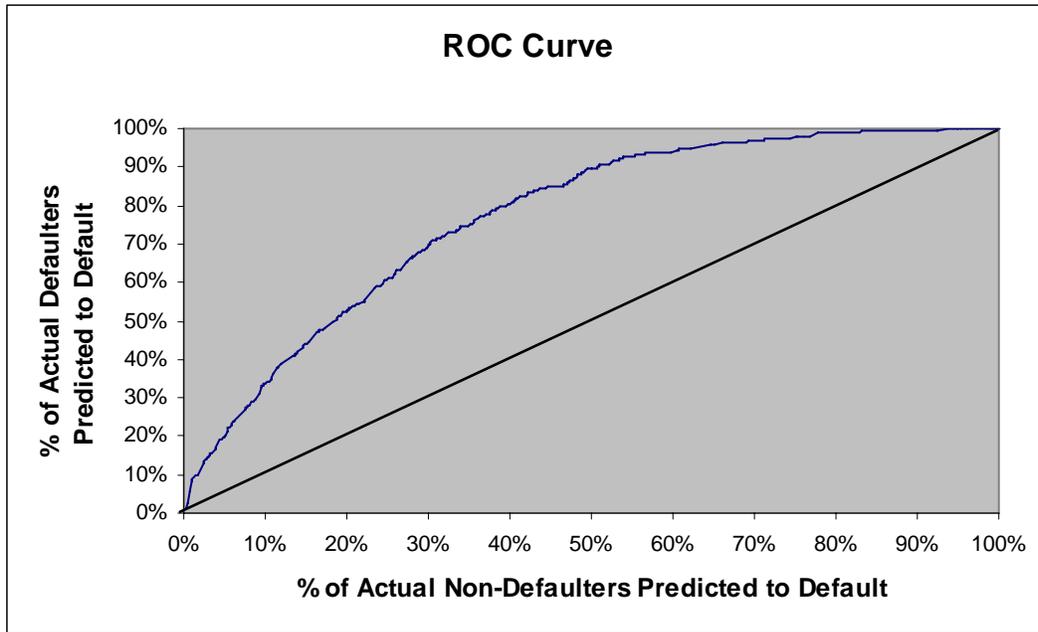
C Statistic

The c statistic measures how consistently a model assigns higher probabilities to actual defaulters than it does to actual non-defaulters. It compares each defaulter with each non-defaulter. In the present analysis, there are 1,685,424 pairings (624 defaulters multiplied by 2,701 non-defaulters). The c statistic indicates the proportion of these cases for which the model assigns a higher probability of defaulting to the defaulter than it assigns to the non-defaulter. For the present model, the c statistic is 77.0 percent – a high value for this statistic.

Receiver Operating Characteristic (ROC) Curve

The c statistic is represented graphically in the chart below. The area under the curve – called a Receiver Operating Characteristic (ROC) curve – is the c statistic: 77.0 percent of the chart is below the curve. A statistical model that assigned the same probabilities to defaulters and non-defaulters – a model that does no better than chance – would have an ROC curve that formed a diagonal running from the lower left corner of the chart to the upper right corner. To the extent that an ROC curve bows above the diagonal, the performance of the model increases. A model

that perfectly separates defaulters and non-defaulters would have an ROC curve that hugged the left-hand side and top of the chart. The ROC curve for this model ranges well above a diagonal and indicates a high level of performance.



Classification Matrix and Misclassification Rate

Constructing a classification matrix provides an easy way to assess how well the statistical model classifies defaulters and non-defaulters. In the following example, the matrix employs a classification rule: if the model assigns a probability of default of 17 percent or more, the borrower is classified as a defaulter; a borrower with less than a 17 percent probability of default is predicted to be a non-defaulter. The matrix shows the numbers of actual defaulters that the classification rule predicts to be defaulters and non-defaulters. It also provides the same information for actual non-defaulters.

		Predicted Outcome	
		Default	Non Default
N=3,325			
Actual Outcome	Default	516	108
	Non Default	1,132	1,569

It is possible to derive a misclassification rate from the classification matrix. When the predicted outcome does not align with the actual outcome, the classification rule resulted in a misclassification. The total number of misclassifications (1,240) is the sum of the defaulters who the model predicted to be non-defaulters (108) and the non-defaulters who the model predicted to be defaulters (1,132). The misclassification rate is 37 percent ($=1,240/3,325$).

Whether or not this misclassification rate is good depends upon the frame of reference. If the school's alternative to using the model is to treat all borrowers as if they are potential defaulters, then a misclassification rate of 37 percent is very good. Treating all borrowers as potential defaulters will misclassify all 2,701 non-defaulters and result in a misclassification rate of 81 percent. In this comparison, using the model results in a 46 percent reduction in the misclassification rate.

If the school's alternative to using the model is to provide counseling to borrowers who have a GPA of 2.5 or lower, the misclassification rate will be about 45 percent, since 86 defaulters in the study have GPAs greater than 2.5, and 1,395 non-defaulters have GPAs of 2.5 or less. Relative to this alternative, the model still provides a modest reduction in the misclassification rate.

This misclassification rate is comprised of two measurements. A low misclassification rate indicates that a method of prediction is successful at predicting both defaulters and non-defaulters. By successfully predicting both, a school can most effectively target its resources to the predicted defaulters. Using the model, with a cutoff of 17 percent to predict defaulters, PVAMU would correctly identify 83 percent of the defaulters and 58 percent of the non-defaulters. As can be seen in the above chart, this means that they would needlessly counsel 1,132 non-defaulters and fail to counsel 108 defaulters.

However, if a school can counsel additional borrowers at a very low cost, it may choose to use a more aggressive method in order to capture a greater percentage of defaulters. For example, if PVAMU used a student's GPA as the predictor of default and provided additional counseling to all students with a GPA of 2.5 or lower, they would correctly identify 86 percent of the defaulters (538 of 624) but only 48 percent of the non-defaulters (1,306 of 2,701). In this scenario, PVAMU would provide unneeded counseling to 1,395 non-defaulters but fail to counsel only 86 defaulters.

Which method a school uses to identify potential defaulters will ultimately depend on the costs of implementing the prediction versus the costs of needlessly counseling borrowers who would not otherwise default. Given the simplicity of using GPA as an indicator of the need for default counseling, combined with the added benefits of counseling students who may be at a higher risk to leave school without completing their degree, it may make the most sense for PVAMU to use this model simply as support for implementing a simpler predictor of default, such as GPA.

Uses of the Findings and the Model and Areas for Future Research

There are many ways in which the results of this study can be used to assist PVAMU in preventing defaults by their student borrowers. These possibilities range from simple solutions requiring only minor changes to existing policies and procedures, to slightly more involved solutions that coordinate financial aid goals with the efforts of other campus functions such as academic advising, career counseling, and instruction. There are also more sophisticated ways in which PVAMU could use this statistical model to identify at-risk borrowers. However, given the

strong predictive power of a student's GPA (which is easily obtainable), these more complicated methods may not be the most cost-effective for this campus.¹¹

The findings of this study are largely consistent with the findings of other studies of default behavior. However, this study does highlight the strong significance of a relatively few number of variables at PVAMU. The importance of the college success variables suggest that targeting increased intervention to certain at-risk students could have a significant impact on PVAMU's default rate. Other research has suggested that by involving the entire campus community in assisting the financial aid office with default aversion policies, an institution can not only lower its default rate, but have the added benefits of increasing its completion rate and promoting student success on campus.¹² The results of this study suggest that early intervention with students who are not performing well academically could benefit both PVAMU's completion rate and default rate.

- A student's grade point average is a very powerful predictor of student loan default. The model found that a student with a GPA of 3.01 to 4.00 has a probability of default which is seven percentage points lower than a student with a GPA of 2.01 to 2.5, even after controlling for background characteristics and degree attainment. On a bivariate level, the relationship is even stronger. Students who left PVAMU with a GPA of 2.5 or less had a default rate of 27.8 percent, whereas those who left PVAMU with a GPA above 2.5 had a default rate of only 6.2 percent. By requiring students whose GPAs drop below 2.5 to receive both academic and financial aid counseling, PVAMU may reach students before they decide to leave school.
- The model illustrates the effect of obtaining a degree on reducing a borrower's likelihood of default. After controlling for other characteristics, students who graduate have a probability of default which is five percentage points lower than students who do not graduate. Once again, this effect is particularly dramatic on a bivariate level. Students who left PVAMU without a degree have a default rate of 27.2 percent, whereas those who obtain their degree have a default rate of only 3.5 percent.
- Persistence in school has a significant effect on a borrower's probability of default, even after controlling for degree obtainment and other borrower characteristics. Students who leave PVAMU after their freshman year have a probability of default that is four percentage points higher than students who remain on campus even one year longer. Students who stay through their senior year see another four percentage-point reduction in the probability of default. This means that students who remain through their senior year see an eight percentage point reduction in their probability of default as compared to those students who leave during their freshman year. These results highlight the importance of working with the entire campus community to increase retention and graduation rates. Increases in these rates are good for both the students and the campus as a whole, and as a bi-product, default rates decrease.

¹¹ The conclusions drawn in this section represent the informed opinions of the author. They are in no way intended to be exhaustive or exclusive of other conclusions. These ideas are shared in the spirit of starting discussions, not ending them.

¹² See "A Clear and Present Danger to Institutional and Student Success." TG and the Council for the Management of Educational Finance, http://www.tgslc.org/pdf/default_model.pdf for more details.

This model highlights the fact that successful default aversion strategies must be a part of a larger, campus-wide effort to retain students and improve their academic performance. Prairie View A&M University has implemented, or is in the process of implementing, several policies based on this principle.

- *More frequent contact between the registrar and the financial aid office.* In the past, the financial aid office (FAO) was not notified of a student's withdrawal from classes until the semester was over. These offices are now working together to provide class attendance sheets to the FAO early in the semester. This will make it easier for financial aid officials to contact students and counsel them regarding their loan repayment obligations.
- *Working with lenders to conduct default management seminars.* PVAMU has begun to require its largest lenders to provide default management seminars on campus. The lenders conduct several seminars each semester in which they provide students not only with information about repaying their student loans, but also with basic financial management tools which will help them throughout their lives.
- *Using the University College to provide additional counseling to freshmen.* PVAMU already provides its freshman students with significant academic and social support through its University College program. The FAO currently works within this system to provide additional financial aid counseling to these young students. This partnership provides an excellent opportunity for PVAMU to lower its default rate, given the much higher default rate among students who leave school during their freshman year.

As noted in the Model Performance section of the paper, the current model does an excellent job of predicting defaulters, but is not as accurate in predicting non-defaulters. In other words, the model too often predicts that borrowers will default in cases in which the borrowers do not actually default. In order to more accurately predict default, further research will need to be conducted. This research could consist of improving the current model by including additional variables which will help to more accurately predict default, or it could involve other types of research.

This model could be improved with the addition of two variables which were unavailable for this study and have been significant in other research. One is very simple: gender. Although this variable is not generally used in developing campus policies, it has proven to be a significant control variable in almost all studies in which it was included. Past research has found that women are less likely to default on their student loans than men. It may be that more of the students that this model predicts to default — and who do *not* default — are women than men.

The second variable which has the potential to strengthen the current model is an indicator of what type (if any) of exit counseling the borrower received. This is a variable which is under the control of campus financial aid offices. If exit counseling is effective, the data should reflect that many of the students that the model predicts to default and in fact do not default, received exit counseling; whereas, many of those who did in fact default left school without receiving the required exit counseling.

Currently, financial aid officials attempt to conduct in-person exit counseling with all students who leave Prairie View A&M University. However, it is often difficult to reach students who withdraw before completing their degree program. All students who do not complete in-person exit counseling are contacted and provided with the opportunity to complete on-line exit counseling. If a student fails to complete either form of counseling, a hold is put on his or her transcript, meaning that no academic records will be released until the required counseling has been completed. PVAMU also hopes that the default management seminars which are being given by lenders to students in their freshman year will help students to more fully understand their student loan repayment obligations.

Other methods of research may also be helpful in learning more about why some borrowers default while others, with the same measurable characteristics, do not. A survey of students who left PVAMU without completing a degree could be very helpful, although it may be difficult to get a satisfactory response rate among this population. A more feasible alternative is to attempt to hold focus groups with students who leave after their freshman year to determine why they left, what their financial situation is, and whether or not they intend to repay their student loans. By comparing these answers to actual default behavior, researchers may gain valuable insight into why students default.

APPENDIX A

Sample Definition and Variable Descriptions

Sample Definition

Using TG's loan data, students were identified who attended PVAMU and who entered repayment on their TG-guaranteed Federal Family Education Loan Program (FFELP) loans between October 1, 2000 and September 30, 2002 (fiscal years 2001 – 2002). This resulted in a sample of 3,692 borrowers. Staff at PVAMU then retrieved detailed data for each of these borrowers from the school's internal databases. As described in the text, only a subset of these variables was used in the final analysis. However, we include a description of each of the variables in this section. Table A1, following the variable descriptions, contains basic descriptive statistics for each variable. Although data were collected for all 3,692 borrowers, this report analyzes the data for only the 3,325 borrowers who were undergraduate students at PVAMU. Therefore, the descriptive statistics are provided for only those undergraduate students.¹³

Variable Descriptions

Default or Not

This study uses the Department of Education's official cohort default rate formula to determine if a borrower is in default. Under this definition, a borrower is considered to be in default if a default claim is paid on his or her behalf during the fiscal year in which the borrower enters repayment, or the following fiscal year.¹⁴ Out of the 3,325 borrowers in the sample of undergraduates, 624 (or 18.8 percent) defaulted under this definition. This variable is the focus of this study.

The subsequent variables all describe the characteristics of the borrowers. These variables help determine if default behavior among borrowers at PVAMU varies among groups of borrowers with different characteristics.

College Success Variables

These variables were provided by PVAMU and measure the borrowers' performance while in attendance at PVAMU. These variables measure both the quantity of education received at PVAMU and the quality of the students' performance in college. It is expected that students who stay in school longer, obtain a degree, and receive higher grades are more likely to repay their student loans. Students who remain in school and receive high grades are more likely to obtain jobs and have higher earnings after college, enabling them to repay their student loans. Students who fail fewer hours are more likely to repay their student loans. This variable, as well as variables such as 'GPA' and 'Has Degree From Any School,' are correlated with personal qualities, such as persistence and discipline, that make a borrower more likely to repay his or her

¹³ Statistics for the full sample, including graduate students are available from TG upon request.

¹⁴ A fiscal year is defined as October 1 – September 30. For example, fiscal year 2002 covers the period October 1, 2001 – September 30, 2002.

loans. Fewer hours failed also means that the student enrolled in fewer courses in order to obtain the necessary credits for his or her degree, making it likely that students who failed fewer courses took on less debt while at PVAMU.

Has Degree From Any School

GPA – Grade Point Average earned at PVAMU.

Highest Academic Level Attained at PVAMU

Total Hours Taken at PVAMU or Transferred from Another School

Number of Hours Passed

Number of Hours Failed

Withdrawals – Number of hours in which student received a grade of ‘W’ (dropped without penalty).

Degree Received – Highest degree received from any institution.

Attendance Pattern

The Attendance Pattern variables describe the length and intensity of a borrower’s attendance at PVAMU. Some of the variables also indicate whether there were interruptions in the borrower’s course of study and whether the borrower was a transfer student. As a group, the variables are intended to signify the borrower’s commitment to the education he or she is pursuing. The study’s author anticipates that borrowers who finish their programs of study, finish sooner rather than later, and finish with few interruptions will default with less frequency than other groups of borrowers. It is also possible that variables such as the student’s major or college of admittance are correlated with the probability of default. Students in certain majors may earn more than those in others, making them more likely to repay their loans successfully.

Type of Admission - Indicates if a student was admitted as a freshman, a transfer student, or readmitted.

Lowest Number of Semester Hours

Number of Withdrawals

Number of Times Student was Readmitted to PVAMU

Number of Summer Semesters Attended

Number of Hours Transferred

Number of Major Changes

Major Field of Study

College of Attendance – College (within PVAMU) in which student is enrolled.

College Preparedness

This set of variables describes students' achievements in high school, before entering college. Past research has found that success in college is related to lower default rates on student loans. It is expected that success in high school is a good predictor of success in college. Therefore, we expect students who were successful in high school to have lower default rates than those who were not successful.

High School GPA
High School Class Rank Percentile
Total SAT Score
High School Graduation Date

Demographics

A borrower's demographic characteristics describe the cultural and family background which a student brings to college. These variables are often included in multivariate analyses in order to control for some of these otherwise unmeasurable differences. These variables affect a student's probability of repaying his or her student loans in several different ways. A borrower's age may reflect the borrower's level of responsibility and experience. A borrower's marital status may proxy for the amount of resources a student has available to repay his or her student loan, with married students having more competing demands on their resources, making them more likely to default. It may be that students who left their home state to attend college are more committed to their education, making them more likely to repay their student loans.

Age of Borrower - Borrower's age at the time of entering repayment.
Ethnicity of Borrower
Marital Status of Borrower
State of Permanent Address of Borrower

Financial Aid Variables

The financial aid variables are from TG's database. They capture the financial resources available to student borrowers when they enter school. It is expected that students who come from families with lower incomes will have a higher probability of default than students from higher income families. However, it may be that these students are more committed to their education and appreciate the chance they have, making them less likely to default.

Amount of Need
Expected Family Contribution

Loan-related Variables

The loan-related variables are from TG's databases. They represent basic measures and indicators of the borrower's student loan experience that might have a relationship to default behavior. It seems that the 'Number of Loans' and the 'Total Loan Amount' should indicate the repayment burden that a borrower faces: the higher the burden, the greater the likelihood of default. However, past research has shown that these variables are usually a proxy for how long the borrower went to school: generally speaking, the higher the loan amount, the more education the borrower received and, therefore, the less likely the borrower is to default. The highest grade at which a borrower received a TG loan is a direct proxy for persistence in school. The longer a borrower remains in school, the less likely he or she is to default.

Number of Loans

Total TG Loan Amount

Highest Grade At Which Borrower Received a TG Loan

TABLE A1
Characteristics of Undergraduate Borrowers at Prairie View A&M University
By Default Status

		Total		Defaulters	
		N	% of cases	N	% of row
Total Sample		3,325	100.0%	624	18.8%
College Success Variables					
	Has Degree From Any School				
	Has Degree	1,183	35.6%	41	3.5%
	Does Not Have Degree	2,142	64.4%	583	27.2%
	Undergraduate GPA				
	0-1.00	393	11.8%	151	38.4%
	1.01-2.00	864	26.0%	270	31.3%
	2.01-2.50	676	20.3%	117	17.3%
	2.51-3.00	858	25.8%	69	8.0%
	3.01-4.00	534	16.1%	17	3.2%
	Highest Academic Level at PVAMU				
	Freshman	1,043	31.4%	354	33.9%
	Sophomore	585	17.6%	138	23.6%
	Junior	335	10.1%	60	17.9%
	Senior or Higher	1,362	41.0%	72	5.3%
	Total Hours Attempted and Transferred				
	0 - 29	554	16.7%	206	37.2%
	30 - 89	1,107	33.3%	299	27.0%
	90 - 149	722	21.7%	70	9.7%
	150 or more	942	28.3%	49	5.2%
	Number of Hours Passed				
	0 - 29	952	28.6%	334	35.1%
	30 - 89	941	28.3%	214	22.7%
	90 - 149	949	28.5%	55	5.8%
	150 or more	483	14.5%	21	4.3%
	Number of Hours Failed				
	Zero	347	10.4%	26	7.5%
	1 - 3	238	7.2%	30	12.6%
	4 - 6	240	7.2%	47	19.6%
	7 - 12	609	18.3%	133	21.8%
	13 or more	1,891	56.9%	388	20.5%
	Number of Hours Withdrawn				
	Zero	1,417	42.6%	319	22.5%
	1 - 3	666	20.0%	141	21.2%
	4 - 6	407	12.2%	68	16.7%
	7 - 9	231	6.9%	21	9.1%
	10 or more	604	18.2%	75	12.4%
	Highest Degree				
	No Degree Earned	2,142	64.4%	583	27.2%
	Certificate or License	52	1.6%	3	5.8%
	Bachelor of Arts	141	4.2%	7	5.0%
	Bachelor of Business Administration	126	3.8%	3	2.4%
	Bachelor of Science	533	16.0%	20	3.8%
	Master's or Higher	331	10.0%	8	2.4%

TABLE A1 – cont'd

		Total		Defaulters	
		N	% of cases	N	% of row
Total Sample		3,325	100.0%	624	18.8%
Attendance Pattern					
Type of Admission					
	Freshman	2,518	75.7%	521	20.7%
	Transfer	615	18.5%	86	14.0%
	Other	192	5.8%	17	8.9%
Lowest Semester Hours					
	Zero	9	0.3%	4	44.4%
	4 - 6	215	6.5%	19	8.8%
	7 - 12	1,750	52.6%	212	12.1%
	13 or more	1,351	40.6%	389	28.8%
Number of Withdrawals					
	Zero	2,898	87.2%	576	19.9%
	One	383	11.5%	44	11.5%
	Two or more	44	1.3%	4	9.1%
Number of Readmittances					
	Zero	2,742	82.5%	540	19.7%
	One	463	13.9%	62	13.4%
	Two or more	120	3.6%	22	18.3%
Number of Summer Semesters					
	Zero	1,372	41.3%	396	28.9%
	One	776	23.3%	119	15.3%
	Two	482	14.5%	64	13.3%
	Three or more	695	20.9%	45	6.5%
Hours Transferred					
	Zero	2,401	72.2%	525	21.9%
	1 - 15	560	16.8%	66	11.8%
	16 or more	364	10.9%	33	9.1%
Number of Major Changes					
	Zero	8	0.2%	3	37.5%
	One	1,431	43.0%	382	26.7%
	Two	1,019	30.6%	143	14.0%
	Three or more	867	26.1%	96	11.1%
Admission Major					
	Declared Major	2,340	70.4%	328	14.0%
	Provisional/Undecided	985	29.6%	296	30.1%
College of Admittance					
	Undeclared	317	9.5%	104	32.8%
	Agriculture	111	3.3%	27	24.3%
	Arts and Sciences	693	20.8%	118	17.0%
	Business	465	14.0%	83	17.8%
	Education	348	10.5%	72	20.7%
	Engineering and Architecture	725	21.8%	145	20.0%
	Graduate School	119	3.6%	4	3.4%
	Honors College	5	0.2%	0	0.0%
	Juvenile Justice and Psychology	102	3.1%	14	13.7%
	Nursing	440	13.2%	57	13.0%

TABLE A1 – cont'd

			Total		Defaulters	
			N	% of cases	N	% of row
Total Sample			3,325	100.0%	624	18.8%
College Preparedness						
	High School GPA					
		Missing	777	23.4%	112	14.4%
		2.00 or less	45	1.4%	8	17.8%
		2.01 – 2.50	547	16.5%	145	26.5%
		2.51 – 3.00	1,244	37.4%	274	22.0%
		3.01 – 4.00	712	21.4%	85	11.9%
	High School Class Rank Percentile					
		Less than 25%	421	12.7%	128	30.4%
		25% - 49%	759	22.8%	182	24.0%
		50% - 69%	552	16.6%	98	17.8%
		70% - 89%	461	13.9%	57	12.4%
		90% - 100%	251	7.5%	25	10.0%
		Missing	881	26.5%	134	15.2%
	Total SAT Score					
		Less than 800	855	25.7%	195	22.8%
		800 to 999	769	23.1%	104	13.5%
		1000 or more	244	7.3%	41	16.8%
		Missing	1,457	43.8%	284	19.5%
	High School Graduation Year					
		Before 1990	17	0.5%	0	0.0%
		1990 - 1994	333	10.0%	47	14.1%
		1995	303	9.1%	40	13.2%
		1996	316	9.5%	46	14.6%
		1997	404	12.2%	83	20.5%
		1998	385	11.6%	102	26.5%
		1999	363	10.9%	102	28.1%
		2000 - 2001	364	10.9%	95	26.1%
		Missing	840	25.3%	109	13.0%

TABLE A1 – cont'd

		Total		Defaulters	
		N	% of cases	N	% of row
Total Sample		3,325	100.0%	624	18.8%
Demographics					
Borrower's Age When Enter Repayment					
	17-20	711	21.4%	193	27.1%
	21-22	902	27.1%	213	23.6%
	23-24	859	25.8%	118	13.7%
	25-26	401	12.1%	56	14.0%
	27-30	260	7.8%	24	9.2%
	31-34	70	2.1%	3	4.3%
	35+	117	3.5%	15	12.8%
Borrower's Ethnicity					
	African-American	3,202	96.3%	607	19.0%
	Caucasian	69	2.1%	8	11.6%
	Hispanic	32	1.0%	5	15.6%
	Other	22	0.7%	4	18.2%
Borrower's Marital Status					
	Single	403	12.1%	52	12.9%
	Married	27	0.8%	2	7.4%
	Divorced and Other	2	0.1%	0	0.0%
	Missing	2,893	87.0%	570	19.7%
State of Permanent Address					
	Texas	3,108	93.5%	590	19.0%
	Non-Texas	217	6.5%	34	15.7%
Financial Aid Variables					
Amount of Need					
	Zero	140	4.2%	15	10.7%
	\$1-2,500	147	4.4%	21	14.3%
	\$2,501-5,000	396	11.9%	75	18.9%
	\$5,001-7,500	421	12.7%	70	16.6%
	\$7,501-10,000	713	21.4%	140	19.6%
	\$10,001 and higher	875	26.3%	159	18.2%
	Missing	633	19.0%	144	22.7%
Expected Family Contribution					
	Zero	1,025	30.8%	234	22.8%
	\$1-750	381	11.5%	68	17.8%
	\$751-2,000	402	12.1%	52	12.9%
	\$2,001-5,000	474	14.3%	68	14.3%
	\$5,001 and higher	410	12.3%	58	14.1%
	Missing	633	19.0%	144	22.7%

TABLE A1 – cont’d

			Total		Defaulters	
			N	% of cases	N	% of row
Total Sample			3,325	100.0%	624	18.8%
Loan-related Variables						
Number of TG Loans						
		1	991	29.8%	281	28.4%
		2 to 4	1,509	45.4%	281	18.6%
		5 to 6	440	13.2%	44	10.0%
		7 to 9	257	7.7%	13	5.1%
		10 or more	128	3.8%	5	3.9%
Total TG Loan Amount						
		\$1 to 3,000	887	26.7%	276	31.1%
		\$3,001 to 5,000	259	7.8%	61	23.6%
		\$5,001 to 7,000	566	17.0%	131	23.1%
		\$7,001 to 10,000	325	9.8%	53	16.3%
		\$10,001 to 15,000	465	14.0%	57	12.3%
		\$15,001 or more	823	24.8%	46	5.6%
Highest Grade At Which Borrower Received A TG Loan						
		Freshman	1,388	41.7%	427	30.8%
		Sophomore	470	14.1%	96	20.4%
		Junior	305	9.2%	40	13.1%
		Senior	969	29.1%	50	5.2%
		Graduate	192	5.8%	11	5.7%

APPENDIX B
Standard Errors and Confidence Intervals

Variable Group	Variable	Reference Group	Coefficient	Standard Error	95% Confidence Interval of Change in Probability
	Intercept		-2.000	0.253	
College Success Variables					
Degree Indicator	Has Degree	No Degree	-0.609	0.122	-6% to -3%
Grade Point Average	0-1.00	2.01-2.50	0.536	0.172	2% to 13%
	1.01-2.00	2.01-2.50	0.274	0.139	0% to 7%
	2.51-3.00	2.01-2.50	-0.278	0.177	-5% to 1%
	3.01-4.00	2.01-2.50	-0.989	0.287	-9% to -4%
Highest Academic Level Attained at PVAMU	Freshman	Sophomore	0.371	0.135	1% to 8%
	Junior	Sophomore	-0.228	0.183	-5% to 1%
	Senior or Higher	Sophomore	-0.510	0.214	-7% to -1%
Attendance Pattern					
Hours Transferred	One or more hours	Zero	-0.186	0.142	-4% to 1%
College Preparedness					
High School Class Rank Percentile	Less than 25%	50% - 69%	0.322	0.163	0% to 9%
	25% - 49%	50% - 69%	0.125	0.149	-2% to 5%
	70% - 89%	50% - 69%	-0.126	0.193	-4% to 3%
	90% - 100%	50% - 69%	0.068	0.261	-4% to 8%
	Missing	50% - 69%	0.255	0.167	-1% to 8%
Financial Aid Variables					
Expected Family Contribution (EFC)	Zero	\$1-750	0.461	0.166	1% to 11%
	\$751-2,000	\$1-750	-0.298	0.213	-6% to 1%
	\$2,001-5,000	\$1-750	-0.071	0.203	-4% to 4%
	\$5,001 and higher	\$1-750	-0.025	0.211	-4% to 5%
	Missing	\$1-750	-0.304	0.178	-5% to 0%

Sample Size: 3,320

Defaulters: 622 (18.7 percent)

-2 log likelihood: Intercept and covariates: 2,699

Chi-Square: 503.93 with 19 degrees of freedom (Pr > ChiSq = <.0001)

C Statistic: 77.0 percent

Baseline probability of default (intercept only): 11.9 percent

* Statistically significant at the 0.05 level

** Statistically significant at the 0.01 level

*** Statistically significant at the 0.001 level

Bibliography

- Barone, S., Steiner, M. and Teszler, N. (2005) Multivariate Analysis of Student Loan Defaulters at Texas A&M University - Kingsville. Austin, TX. May.
- Dynarski, M. Who defaults on student loans? Findings from the National Postsecondary Student Aid Study. *Economics of Education Review*, 13(1), 55-68.
- Knapp, L.G., and Seaks, T.G. (1992). An analysis of the probability of default on federally guaranteed student loans. *Review of Economics and Statistics*, 74(3), 404-411.
- McMillion, Robin (2005). The Role of Work and Loans in Paying for an Undergraduate Education: Observations from the 2003-2004 National Postsecondary Student Aid Study (NPSAS). Austin, TX. November.
- Meyer, D. (1998). Predicting which borrowers are most likely to default. In: Webster, J., Meyer, D. & Arnold, A. Student loan defaults in Texas: Yesterday, today, and tomorrow. Austin, TX: Texas Guaranteed Student Loan Corporation.
- Monteverde, K. (2000). Managing student loan default risk: Evidence from a privately-guaranteed portfolio. *Research in Higher Education*, 41(3), 331-352.
- Podgursky, M., Ehlert, M., Monroe, R. and Watson, D. (2000). Student loan defaults and enrollment persistence. Columbia, MO. October.
- Steiner, M. and Tym, C (2005). Multivariate Analysis of Student Loan Defaulters at The University of South Florida. Austin, TX. March.
- Steiner, M. and Teszler, N (2003). The Characteristics Associated with Student Loan Default at Texas A&M University. Austin, TX. January.
- Steiner, M. and Teszler, N (2005). Multivariate Analysis of Student Loan Defaulters at Texas A&M University. Austin, TX. January.
- TG and the Council for the Management of Educational Finance (2003). *A Clear and Present Danger to Institutional and Student Success: A training model for embedding student loan default aversion within strategic enrollment management*. Austin, TX.
- U.S. Department of Education (2001). *The Cohort Default Rate Guide*. Washington D.C. <http://www.ifap.ed.gov/drmaterials/finalcdrg.html#mulifiles> .
- U.S. Department of Education (1996), National Center for Education Statistics. *Nontraditional Undergraduates: Trends in Enrollment from 1986 to 1992 and Persistence and Attainment Among 1989-90 Beginning Postsecondary Students*, NCES 97-578, by Laura J. Horn. Project Officer: C. Dennis Carroll, Washington, DC: 1996.

Volkwein, J. F., Cabrera, A. F., Szelest, B. P., and Napierski, M. (1995). *Characteristics of student loan defaulters among different racial and ethnic groups*. Boston, MA: AIR 1995 Annual Forum Paper. (ERIC Document Reproduction Service No. ED 386 972).

Volkwein, J. F., and Szelest, B. P. (1995). Individual and campus characteristics associated with student loan default. *Research in Higher Education*, 36(1), 41-72.

Wilms, W. W., Moore, R. W., and 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(1), 41-54.

Woo, J. H. (2002). Factors affecting the probability of default: Student loans in California. *Journal of Student Financial Aid*, 32(2), 5-23.