DOCUMENT RESUME

ED 354 801 HE 026 220

AUTHOR Dynarski, Mark

TITLE Analysis of Factors Related to Default.
INSTITUTION Mathematica Policy Research, Princeton, N.J.; Mathtech, Inc., Arlington, VA.; Pelavin Associates, Inc., Washington, DC.
SPONS AGENCY Department of Education, Washington, DC. Office of Planning, Budget, and Evaluation.
PUB DATE Apr 91
CONTRACT 88001001; 90109001
NOTE 64p.
PUB TYPE Reports - General (140)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Academic Achievement; Comparative Analysis; Credit (Finance); Ethnic Groups; Financial Problems; Higher Education; *Individual Characteristics; *Loan Default; Loan Repayment; Multivariate Analysis; *Paying for College; Postsecondary Education; *Socioeconomic Influences; Student Loan Programs
IDENTIFIERS *Guaranteed Student Loan Program

ABSTRACT

This paper analyzes the factors associated with student loan default in the Guaranteed Student Loan (GSL) program for higher education. The paper provides an overview of the National Postsecondary Student Aid Study (NPSAS) Student Loan Recipient Survey, and, using data from the survey, presents a descriptive analysis of student loan recipients and of default rates, broken down by various demographic, socioeconomic, and educational-level groupings. Also, the paper explores the underlying factors that determine default, using multivariate statistical techniques. Three key results are reported based on the analysis. First, that a borrower's ability to pay is a powerful determinant of default. Second, that the default rates differ significantly by level of educational attainment. Finally, that default rates differ significantly by the characteristics of borrowers. In summarizing these key results, the report finds that: (1) borrowers leaving school with lower incomes, whose monthly GSL payments are higher, and who have more dependents are more likely to default; (2) borrowers who do not complete high school or their postsecondary education programs are more likely to default; and (3) Black and Hispanic borrowers are more likely to default, after controlling for income, education, and other individual characteristics. Contains seven references. (GLR)

ANALYSIS OF FACTORS RELATED TO DEFAULT

April 1991

Mark Dynarski
Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ 08543-2393

U.S. DEPARTMENT OF EDUCATION • OFFICE OF PLANNING, BUDGET & EVALUATION

This report has been funded either wholly or in part with Federal funds from the U.S. Department of Education, Postsecondary Education Division, under prime contract number 88001001 with Pelavin Associates and prime contract number 90109001 with Mathtech, Inc. The content of this publication does not necessarily reflect the views or policies of the U.S. Department of Education, nor does the mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government or by Mathematica Policy Research, Inc.

ACKNOWLEDGEMENTS

I have benefited from the comments, suggestions, and assistance of many individuals in preparing this report. In particular, I wish to thank Alan Ginsburg, Maureen McLaughlin, Dan Goldenberg, Dennis Carroll, and Jim Maxwell of the U.S. Department of Education, Laura Salganik of Pelavin Associates, and Rebecca Maynard of Mathematica Policy Research, Inc. for their helpful comments and insights. Renee Donahey and Linda Sperling provided excellent programming assistance. The report was produced by Denise Dunn, Monica Capizzi, and Debbie Jones, and was edited by Thomas Good.

Mark Dynarski

CONTENTS

Chapter		Page
	EXECUTIVE SUMMARY	ix
I	INTRODUCTION	1
II	DESCRIPTIVE ANALYSIS OF GSL BORROWERS AND DEFAULTERS	3
	A. THE NPSAS STUDENT LOAN RECIPIENT SURVEY	4
	B. THE CHARACTERISTICS OF GSL BORROWERS	5
	C. THE CHARACTERISTICS OF STUDENT LOAN DEFAULTERS	9
	1. Default Rates by Demographic, Educational, and GSL Characteristics	9
	2. High School and Postsecondary School Completion	17
	D. DEFAULT RATES AND SOURCES OF GSL PROGRAM AND REPAYMENT INFORMATION	17
	E. THE IMPORTANCE OF VARIOUS FACTORS THAT LED TO DEFAULT	21
III	A MULTIVARIATE ANALYSIS OF DEFAULT ON GUARANTEED STUDENT LOANS	23
	A. A CONCEPTUAL FRAMEWORK FOR DEFAULT	23
	B. AN EMPIRICAL MODEL OF DEFAULT	27
	C. VARIABLES FOR THE EMPIRICAL MODEL OF DEFAULT	28
	1. Independent Variables	29
	D. HYPOTHESIS TESTS IN THE MULTIVARIATE MODEL	33
	E. ESTIMATION RESULTS	34
	1. Hypothesis Tests and Discussion	37
	2. Subgroup Results	38
	F. IMPACT ESTIMATES	43
	REFERENCES	47
	APPENDIX A: ADDITIONAL ANALYSIS OF THE EFFECT OF SCHOOL TYPE AND POST-SCHOOL EARNINGS ON DEFAULT	

TABLES

Table		Page
1	LOAN RECIPIENT CHARACTERISTICS BY RACE/ ETHNICITY	6
2	LOAN RECIPIENT CHARACTERISTICS BY TYPE OF POST-SECONDARY INSTITUTION MOST RECENTLY ATTENDED	8
3	DEFAULT RATES BY DEMOGRAPHIC, EDUCATIONAL, AND GSL CHARACTERISTICS	10
4	DEFAULT RATES BY RACE/ETHNICITY, AND TYPE OF POSTSECONDARY INSTITUTION MOST RECENTLY ATTENDED	14
5	DEFAULT RATES BY RACE/ETHNICITY AND ANNUAL EARNINGS	15
6	DEFAULT RATES BY ANNUAL EARNINGS AND TYPE OF POSTSECONDARY INSTITUTION MOST RECENTLY ATTENDED	16
7	DEFAULT RATES BY WHETHER EDUCATIONAL PROGRAM WAS COMPLETED AND BY TYPE OF POSTSECONDARY INSTITUTION MOST RECENTLY ATTENDED	18
8	DEFAULT RATES BY TYPE OF POSTSECONDARY INSTI-TUTION MOST RECENTLY ATTENDED AND BY SOURCE OF INFORMATION ON THE GSL PROGRAM	19
9	DEFAULT RATES BY TYPE OF POSTSECONDARY INSTI-TUTION MOST RECENTLY ATTENDED AND BY SOURCE OF INFORMATION ON GSL REPAYMENT	20
10	IMPORTANCE OF FACTORS RELATED TO DEFAULT, BY TYPE OF POSTSECONDARY INSTITUTION	22
11	DESCRIPTIVE STATISTICS FOR THE MULTIVARIATE ANALYSIS	30
12	SIMPLE CORRELATION MATRIX	32
13	ESTIMATION RESULTS FOR THE ENTIRE SAMPLE	35
14	ESTIMATION RESULTS FOR RACE/ETHNICITY SUB-GROUPS	39

TABLES (continued)

Table		Page
15	ESTIMATION RESULTS FOR POSTSECONDARY SCHOOL SUBGROUPS	41
16	IMPACT ESTIMATES FOR THE LINEAR MODEL OF DEFAULT	44
A-1	ESTIMATION RESULTS FOR A REGRESSION MODEL OF POST-SCHOOL EARNINGS	50

EXECUTIVE SUMMARY

The Guaranteed Student Loan (GSL) program is the largest source of federal assistance to help needy students pay for their postsecondary education. In fiscal year (FY) 1990, 4 million students will participate in the GSL programs, borrowing approximately $11 billion.

As the popularity of the GSL program has increased, so have defaults. The costs of defaults have risen eightfold since FY 1981. By FY 1990, nearly 44 percent (over $2 billion) of program funds went to pay default costs.

As part of the Department of Education's ongoing effort to reduce defaults, it commissioned this report to analyze the factors associated with default. The report uses data from the NPSAS Student Loan Recipient Survey to analyze the characteristics of student loan recipients and to compare student loan defaulters and nondefaulters along a variety of dimensions, including their demographic profiles, their socioeconomic characteristics, and their educational attainment.

Three key results emerge from our analyses:

- **A borrower's ability to pay is a powerful determinant of default.** The likelihood of default is greater for borrowers whose incomes after leaving school are lower, whose monthly GSL payments are higher, and who have more dependents.

- **Default rates differ significantly by level of educational attainment.** Borrowers who did not complete high school and borrowers who did not complete their postsecondary education programs were more likely to default. After adjusting for other factors, including post-school earnings, borrowers who had most recently attended proprietary or two-year schools were also more likely to default. Differences in (adjusted) default rates between borrowers attending proprietary or two-year schools and borrowers attending four-year schools are greater when the effect of school type on post-school earnings is taken into account.

- **Default rates differ significantly by the characteristics of borrowers.** Black and Hispanic borrowers were more likely to default, after income, education, and other individual characteristics were controlled for.

Other researchers have reached conclusions similar to these, but their results were generally based on samples of borrowers within single states or schools (Wilms et al., 1987; and Greene, 1989) or on

samples that contained only limited information about the characteristics of borrowers (U.S. General Accounting Office, 1988). In contrast, our results are derived from a nationally representative sample of borrowers that contains rich detail about their characteristics. We view our results as both corroborating the findings of other researchers and providing new insights into the underlying determinants of student loan default.

I. INTRODUCTION

The Guaranteed Student Loan Program (GSL) grants low-interest long-term loans to eligible students who wish to attend participating postsecondary schools.[1] The federal government makes interest payments on GSL loans while students are in school, pays private lenders--who provide the capital--special allowances if the market interest rate exceeds the GSL interest rate, and guarantees the principal and interest on GSL loans indirectly by reinsuring guaranty agencies. In turn, these agencies pay lenders directly in the event of the death, disability, bankruptcy, or default of the borrower.

The GSL program has been popular since its inception in 1966. The cumulative volume of loans disbursed over the 21-year period ending in 1987 was roughly $74 billion (U.S. Department of Education, 1988). The volume of loans in 1987 alone was roughly $8.6 billion, with 3.5 million loans committed and an average loan equalling roughly $2,500. The annual volume of loans was 560 percent greater in 1987 than in 1977.

The growth of the GSL program has been paralleled by a growing rate of default on GSL loans. Default rates can be defined in a number of ways, but the general conclusion is that default rates rose significantly between 1983 and 1987. For example, the default rate on cumulative matured paper was 9.1 percent in 1983 and 13.1 percent in 1987, an increase of 44 percent over a five-year period.[2] In 1987, guaranty agencies paid lenders for 454,000 defaulted loans whose total value exceeded $1.3 billion.

---

[1]The Guaranteed Student Loan Program consists of three separate programs: the "regular" Guaranteed Student Loan program (GSL), the Parent Loans to Undergraduate Students program (PLUS), and the Supplemental Loans to Students program (SLS). In this report, we focus on the regular GSL program, which is by far the largest of the three loan programs (accounting for 88 percent of the volume of loans in 1987). The GSL program was named the Stafford loan program in July 1988. We use the older term because our data sources and documentation are based on individuals who received GSL loans prior to 1988.

[2]Cumulative matured paper is the total value of GSL loans that have entered repayment since the inception of the program.

The significant and growing volume of defaulted student loans has called for a more in-depth understanding of the factors underlying default. In response, the U.S. Department of Education commissioned the National Postsecondary Student Aid Study (NPSAS) Student Loan Recipient Survey (SLRS), which collected data on a nationally representative sample of individual borrowers in the GSL program. A significant proportion of the SLRS sample consisted of borrowers whom guaranty agency,records identified as having defaulted on at least one of their student loans. The remainder of the sample consisted of borrowers who were identified as having repaid their student loans in full, and of borrowers who were in the process of repaying their loans and had not previously defaulted. Data from over 8,000 respondents were collected through telephone and field interviews.

The outline of this report is as follows. In Chapter II, we provide an overview of the NPSAS Student Loan Recipient Survey, and, using data from the survey, we present a descriptive analysis of student loan recipients and of default rates, broken down by various demographic, socioeconomic, and educational-level groupings. In Chapter III, we explore the underlying factors that determine default, using multivariate statistical techniques. Many of the findings from the descriptive analysis are sustained in the multivariate analysis.

II. DESCRIPTIVE ANALYSIS OF GSL BORROWERS AND DEFAULTERS

In this chapter we use data from the NPSAS Student Loan Recipient Survey to provide a descriptive analysis of the characteristics of individuals who obtain guaranteed student loans, and the characteristics of individuals who default on their student loans. We first briefly discuss the Student Loan Recipient Survey, and provide an overview of the characteristics of loan recipients. We then examine patterns of default in terms of the demographic and socioeconomic characteristics of loan recipients, the types of postsecondary institutions they attended, the amount of student loan indebtedness they incurred, and their sources of information about the GSL program and the repayment process. We also examine the frequency of various factors cited by defaulters as having led to their default.

Our results can briefly be summarized as follows:

- The characteristics of GSL borrowers differ widely across racial/ethnic subgroups and across types of postsecondary institutions. Borrowers from minority groups and borrowers from proprietary schools and two-year schools are more likely to be from low-income and low-education households, and have low earnings themselves after leaving school. Larger proportions of minority group and proprietary school borrowers fail to complete their postsecondary program.

- Default rates differ widely across demographic categories, socioeconomic characteristics, and educational levels. Borrowers are more likely to default if they are black or Hispanic, if they are unmarried or have more dependents when their loans come due, if they have not completed high school, if they have most recently attended a proprietary school or two-year school, if they have not completed their postsecondary program, or if they have low earnings after leaving school.

- Default rates are higher among borrowers who learn about the GSL program only through their postsecondary institution, which among our sample occurred more frequently for proprietary school borrowers. Default rates are also higher among borrowers who receive no information about the repayment process, or who receive information about the repayment process only from their postsecondary institution at the time the loan is made. Proprietary school borrowers are more likely to receive information about the repayment process only from their postsecondary school, and are less likely to learn about repayment from multiple sources.

- According to our sample of defaulters, being unemployed or having a low income after leaving school is an important factor that leads to default. Proprietary and two-year school borrowers are more likely to cite unemployment as a factor that leads to default.

A. THE NPSAS STUDENT LOAN RECIPIENT SURVEY

In this section we provide an overview of the NPSAS Student Loan Recipient Survey (SLRS). A more complete discussion of the methodology used in the Student Loan Recipient Survey is found in Knight et al. (1988).

The universe of potential respondents for the SLRS was defined as all individuals who had received a regular guaranteed student loan (GSL) or a federally insured student loan (FISL), and who had left a postsecondary institution between 1976 and 1985. According to GSL program data, this sample frame contained roughly 8 million loan recipients.[1] Recipients were classified as "in default" if they had ever defaulted on any GSL, as "paid in full" if they had completed repayment on any GSL and had not defaulted on a GSL, and as "in repayment" if they had not completed repayment on a GSL and had not defaulted on a GSL.

The sample frame was divided into eighteen strata, according to whether respondents had left a HEGIS or non-HEGIS institution,[2] whether respondents were classified as in default, in repayment, or paid in full, and whether respondents were 1 to 2 years out of school, 3 to 5 years out of school, or 6 to 10 years out of school. Disproportionate sample sizes were allocated to the various strata to guarantee a sufficient number of observations for stratum-level analysis.

The sample was selected using a two-stage process. In the first stage of sampling, schools were selected with probabilities proportional to their total number of loans in the 1976-1985 period. In the second stage of sampling, GSL recipients from sampled schools were selected with probabilities that depended on the strata into which recipients were classified.

---

[1]Recipients who had more than one loan were counted only once in the sample frame.

[2]The Higher Education General Information System (HEGIS) maintains data for most two-year and four-year postsecondary institutions. Non-HEGIS institutions consist primarily of proprietary schools.

The initial sample design called for a target sample size of 14,025 respondents. Due to difficulties in locating some respondents, the target sample size was later reduced to 11,847 respondents. The final sample consisted of 8,223 respondents, for an overall completion rate of 69.4 percent.[3]

Stratum-level response rates ranged from a low of 56 percent (for the in-repayment, non-HEGIS, 1976-1980 stratum) to a high of 82 percent (for the paid-in-full, HEGIS, 1984-1985 stratum). The response rates by loan status were 68 percent for respondents classified as in repayment, 71 percent for respondents classified as paid in full, and 72 percent for respondents classified as defaulters.[4]

B. THE CHARACTERISTICS OF GSL BORROWERS

An important feature of the SLRS data is the nationally representative nature of the sample. Thus, sample characteristics from the SLRS are estimates of the characteristics of the population of GSL borrowers who left school during the 1976-1985 period.

Table 1 shows weighted sample means for a variety of demographic, educational, and GSL-related characteristics. As shown in column 1 of Table 1, GSL borrowers as a whole were relatively young, were likely to be unmarried when their loans came due, and were likely to have most recently attended a four-year school. The average GSL amount borrowed was $4,629, and the average monthly GSL payment was $73.50. Within the first two years after their loans came due, GSL borrowers reported average annual gross earnings of $13,400, or about $1,100 monthly. The average GSL monthly payment thus represented 6.6 percent of gross monthly earnings.

Sharp differences in the characteristics of the population of GSL borrowers are evident when the sample is separated into racial/ethnic subgroups and postsecondary institution subgroups. As

---

[3]The final sample included 177 respondents who said that they had not received a GSL or FISL loan, and 50 respondents who were out-of-scope (e.g., deceased), which reduced the possible sample size to 11,797 respondents. Classifying the 177 respondents who said that they had not received a loan as out-of-scope reduces the overall response rate to 69.2 percent.

[4]The sample weights included in the data file incorporate the differential response rates by stratum.

TABLE 1

LOAN RECIPIENT CHARACTERISTICS BY RACE/ETHNICITY

Characteristic	(1) Total Sample	Race/Ethnicity (2) White	(3) Black	(4) Hispanic
Age (Years)[a]	30.2	30.2	30.8	29.8
Sex (% Male)	49.7	51.2	38.3	47.6
Married[b] (%)	24.5	25.3	18.0	29.0
Number of Dependents[b]	0.56	0.50	1.00	0.76
Parental Income[c]				
<$17,000 (%)	27.6	22.4	61.2	52.0
$17,000-$30,000 (%)	31.3	32.4	24.4	24.7
>$30,000 (%)	41.0	45.0	14.3	23.3
Parental Education[c]				
<High School (%)	10.7	7.8	26.2	30.5
High School (%)	56.8	56.7	60.4	48.3
>High School (%)	32.4	35.5	13.3	21.2
Type of Postsecondary Institution Most Recently Attended				
Proprietary School (%)	19.2	15.2	42.0	34.5
Two-Year School (%)	15.3	15.4	15.1	15.4
Four-Year School (%)	65.5	69.5	43.0	50.1
Completed Postsecondary Program (%)	75.3	77.8	58.9	67.3
Total GSL Amount Borrowed	$4,629	$4,859	$3,100	$4,068
Total GSL Monthly Payment	$73.50	$75.00	$63.60	$68.00
Average Annual Earnings[b]	$13,423	$14,200	$8,631	$10,983
Proportion of Months with No Reported Earnings[b] (%)	24.9	22.7	38.9	31.8
Payment as a Percent of Monthly Earnings	6.6	6.3	8.8	7.4
Sample Size[d]	7,613	6,027	1,199	387

NOTE: Estimates are weighted to offset differential sampling rates and nonresponse.

[a]As of January 1, 1988.
[b]These variables refer to the two-year period beginning in the month that the loan repayment was scheduled to begin.
[c]This category pertains to the time that the borrowers first entered postsecondary education.
[d]The sample sizes shown refer to the total number of cases in each subgroup. The sample sizes for individual characteristics differ due to nonresponse. In addition, the total sample includes 19 respondents who refused to report their race/ethnicity.

shown in Table 1, white GSL borrowers are more likely to have attended a four-year school, and to have higher parental income, higher parental education, and higher earnings after leaving school. Black and Hispanic borrowers are more likely to have attended proprietary or two-year schools, and to have lower parental income, lower parental education, and lower earnings after leaving school. For example, as shown in column 3 of Table 1, 61 percent of black borrowers had parental income below $17,000 compared with 22 percent for white borrowers. The parents of 26 percent of black borrowers and 31 percent of Hispanic borrowers were high school dropouts, compared with 8 percent of the parents of white borrowers.[5] A smaller proportion of black borrowers were married when their GSL loan came due, and the average number of dependents was larger. Only 59 percent of black borrowers completed the postsecondary program that they had attended most recently, versus 67 percent for Hispanics and 78 percent for whites. Average annual earnings for black borrowers was $8,631, which was 61 percent of annual earnings of $14,200 for white borrowers. The average monthly GSL payment of $63.60 for black borrowers represented 8.8 percent of monthly earnings.

Table 2 reports weighted averages for subgroups defined by the type of postsecondary institution most recently attended. Three types are used here: proprietary schools, two-year schools, and four-year schools. In general, the characteristics of borrowers who had attended two- year schools were similar to the characteristics of borrowers who had attended proprietary schools. Borrowers who had attended proprietary schools and two-year schools were more likely to be black or Hispanic and to have low parental income and education compared with borrowers who had attended four-year schools. Borrowers from four-year schools had greater GSL indebtedness and greater monthly GSL payments, but they also had greater average annual earnings after leaving school ($15,800 for four-year-school borrowers, versus $10,100 for two-year-school borrowers and $8,261 for proprietary-school borrowers). The average monthly GSL payment of $81.80 for four-year-school borrowers thus repre-

[5]The parental education variable is the highest level of education reported for the mother and the father, or the educational level that is reported if the level is known only for one parent.

TABLE 2

LOAN RECIPIENT CHARACTERISTICS BY TYPE OF POSTSECONDARY INSTITUTION MOST RECENTLY ATTENDED

Characteristic	(1) Total Sample	Postsecondary Institution (2) Proprietary	(3) Two-Year	(4) Four-Year
Age (Years)[a]	30.2	29.1	29.9	30.6
Sex (% Male)	49.7	42.3	41.9	53.6
Married[c] (%)	24.5	22.7	24.4	25.0
Number of Dependents[b]	0.56	0.77	0.77	0.45
Race/Ethnicity				
White (%)	84.2	66.6	84.5	89.4
Black (%)	11.6	25.8	11.4	7.4
Hispanic (%)	4.6	8.2	4.6	3.5
Parental Income[c]				
<$17,000 (%)	27.6	41.8	40.4	21.6
$17,000-$30,000 (%)	31.3	30.7	33.0	31.1
>$30,000 (%)	41.0	27.4	26.6	47.3
Parental Education[c]				
<High School (%)	10.7	20.0	15.4	7.0
High School (%)	56.8	66.9	64.7	52.1
>High School (%)	32.4	12.9	19.9	40.9
Completed Postsecondary Program (%)	75.3	71.1	56.6	80.9
Total GSL Amount Borrowed	$4,629	$2,975	$2,717	$5,568
Total GSL Monthly Payment	$73.50	$59.00	$55.40	$81.80
Average Annual Earnings[b]	$13,423	$8,261	$10,110	$15,800
Proportion of Months with No Reported Earnings (%)	24.9	35.9	28.1	20.9
Payment as Percent of Monthly Earnings	6.6	8.6	6.6	6.2
Sample Size[d]	7,397	1,407	1,371	4,618

NOTE: Estimates are weighted to offset differential sampling rates and nonresponse.

[a] As of January 1, 1988.
[b] These variables refer to the two-year period beginning in the month that the loan repayment was scheduled to begin.
[c] This category pertains to the time that the borrowers first entered postsecondary education.
[d] The sample sizes shown refer to the total number of cases in each subgroup. The sample sizes for individual characteristics differ due to nonresponse. In addition, the total sample includes 571 respondents whose most recent type of postsecondary school was unspecified.

sented only 6.2 percent of their average gross monthly earnings, whereas the average monthly payment was 8.6 percent of average monthly earnings for proprietary-school borrowers.

C. THE CHARACTERISTICS OF STUDENT LOAN DEFAULTERS

This section focuses on the characteristics of defaulters in terms of their demographic and socioeconomic background, their educational level, their earnings, the GSL amount borrowed, and their sources of information about the GSL program and the repayment process.

For purposes of the descriptive analysis, respondents were classified as defaulters if they responded that they had ever defaulted on a GSL payment to one or two of the lenders from which they had received their loan(s).[6] (For purposes of the survey, the maximum number of lenders from which default could have occurred was limited to two.) We defined the *default rate* as the number of respondents who reported that they had defaulted, divided by the total number of respondents. Responses were weighted to adjust for differential sampling probabilities and nonresponse. The accompanying tables also report the (unweighted) total number of sample respondents in various categories.

1. Default Rates by Demographic, Educational, and GSL Characteristics

Table 3 displays default rates and sample sizes for various demographic, educational, and GSL characteristics. The estimated overall default rate from the NPSAS Student Loan Recipient Survey is 17.0 percent.[7] Six important results emerge from Table 3:

---

[6]Some respondents who were sampled as defaulters responded that they had paid their loans in full or were in the process of repaying their loans. On the other hand, some respondents who were sampled as being in repayment or having paid their loans in full reported that they had defaulted. Movements across sample statuses were roughly offsetting. Seventy percent of those sampled as defaulters reported that they had defaulted. The remaining 30 percent of those sampled as defaulters reported that they had not defaulted, but a roughly equal number who were not classified as defaulters reported that they had defaulted (Knight et al., 1988). Because borrowers changed their loan status over time, we did not use loan status at the time of sampling as an indicator of default.

[7]The default rate was 14.2 percent for loan recipients who left postsecondary institutions in the 1976-1980 period, 17.3 percent for loan recipients who left in the 1981-1983 period, and 17.1 percent for loan recipients who left in 1984 or 1985.

TABLE 3

DEFAULT RATES BY DEMOGRAPHIC, EDUCATIONAL, AND GSL CHARACTERISTICS

Variable	Default Rate	Sample Size
All	17.0	7,954
Age		
21-30	15.5	4,952
31-40	17.8	2,412
>41	23.6	531
Gender		
Male	15.4	3,846
Female	18.0	4,108
Race/Ethnicity		
White	11.3	6,027
Black	48.0	1,199
Hispanic	25.4	387
Other[a]	19.6	250
Marital Status at Time of First Payment		
Married	14.8	1,804
Not Married	17.6	6,037
Number of Dependents at Time of First Payment		
No Dependents	13.1	5,466
1-2	24.1	1,673
More than 2	31.5	677
High School Education		
Did Not Complete	56.2	207
GED	44.2	465
Diploma	14.4	7,158
Type of Postsecondary Institution Most Recently Attended		
Proprietary	31.5	1,407
Public, Two-Year	26.0	1,222
Private, Two-Year	23.1	149
Public, Four-Year	11.1	2,779
Private, Four-Year	9.2	1,839

TABLE 3 (continued)

Variable	Default Rate	Sample Size
Highest College Degree Earned		
No Degree	32.6	1,965
Certificate (Two-Year or Less)	28.5	1,299
Associate's	13.3	681
Bachelor's	8.2	2,610
Master's	6.1	668
Doctor's	2.6	119
Professional	5.8	319
GSL Amount Borrowed		
<$5,000	21.4	4,906
$5,000-$9,999	10.4	1,824
≥$10,000	6.8	794
Unspecified	19.0	443
Monthly GSL Payment		
≤$50	20.8	2,918
$51-$100	13.1	4,106
>$100	11.0	1,560
Unspecified	18.9	1,299
Average Annual Earnings[b]		
<$10,000	36.5	1,332
$10,000-$14,999	22.6	397
$15,000-$24,999	16.5	969
≥$25,000	9.6	1,962

NOTE: The default rate is defined as the weighted proportion of respondents who reported they had defaulted on their loan(s). Weights were used to offset differential sampling rates and nonresponse.

[a]Includes American Indians, Alaskan natives, Asians, and Pacific Islanders.

[b]Average annual earnings refer to the two-year period beginning in the month that the loan repayment was scheduled to begin.

- Black and Hispanic borrowers had higher default rates.
- Borrowers who were unmarried or who had more dependents at the time that GSL repayment was scheduled to begin had higher default rates.
- High school dropouts and high school graduates who received a GED had higher default rates.
- Borrowers who had most recently attended proprietary schools or two-year schools had higher default rates.[8]
- Borrowers who received no degree or certificate had high default rates; borrowers who received bachelor's degrees or more advanced degrees had low default rates.
- Borrowers who had low earnings after leaving their postsecondary institutions had higher default rates.

It should be noted that these results are descriptive only and do not imply a direct relationship between default and the characteristics in question. For example, the finding that married borrowers had lower default rates is likely due to the fact that they had a higher household income level. When we examined reported 1986 household income, we found that household income for married borrowers was $14,000 higher than household income for unmarried borrowers, after adjusting for the borrowers' educational level and years out of school. Similarly, the finding that the number of dependents is associated with default may reflect the lower levels of discretionary income available to make loan payments for borrowers with larger families. In a sense, borrowers with larger families are less able to pay than borrowers whose families are smaller and whose incomes are equivalent.

Several other findings merit further investigation. For example, we noted in the previous section that black and Hispanic borrowers were more likely to attend proprietary and two-year schools, and that they generally had lower earnings. Each of these factors is associated with higher default rates. It is not possible to say more about the relationships among race/ethnicity, school type, earnings, and

[8]It is possible for students to have defaulted on loans that they used to attend institutions that differed from the most recently attended institutions. For example, a student may have left a two-year school and then attended a four-year school. If the student defaulted on a loan used to attend the two-year school, the construct employed here would count this default in the four-year- school category.

default on the basis of the simple averages presented in Table 3. To provide more insight on these relationships, Tables 4 to 6 show default rates cross-tabulated by race/ethnicity and type of postsecondary institution most recently attended, by race/ethnicity and earnings, and by type of school and earnings.

Table 4 shows default rates by race/ethnicity and type of school. Though the overall default rate for proprietary school borrowers was 31 percent, the default rate was 56 percent among blacks attending proprietary schools, versus 37 percent for Hispanics and 20 percent for whites. The overall default rate for public four-year-school borrowers was 11 percent, but the default rate was 38 percent among blacks attending public four-year schools, versus 16 percent for Hispanics and 9 percent for whites. The ranking of default rates by race/ethnicity is the same in each of the school categories.

Table 5 shows default rates by race/ethnicity and earnings. The ranking of default rates by race/ethnicity is the same in each of the earnings categories, with blacks having the highest default rates, followed by Hispanics and whites. In particular, black borrowers who reported earnings of more than $25,000 had a default rate of 36 percent, which was five times larger than the default rate for whites in that earnings category and almost four times larger than the default rate for Hispanics in that earnings category.

Table 6 shows default rates by earnings and type of school. Within school types, default rates fell as earnings rose. For example, the default rate for proprietary-school borrowers was 48 percent when their earnings were below $10,000, 22 percent when their earnings were between $15,000 and $25,000, and 16 percent when their earnings exceeded $25,000. Within earnings categories, borrowers in proprietary and public two-year schools consistently had the highest default rates. Even when earnings exceeded $25,000, borrowers in proprietary and public two-year schools had default rates of approximately 16 percent. These rates compare with a default rate of 7 percent for high-earning borrowers in four-year schools, and are slightly below the default rate found for borrowers in four-year schools with earnings of less than $10,000.

TABLE 4

DEFAULT RATES BY RACE/ETHNICITY, AND TYPE OF POSTSECONDARY INSTITUTION MOST RECENTLY ATTENDED

	Race/Ethnicity							
	(1) White		(2) Black		(3) Hispanic		(4) Total	
School Type	Default Rate	Sample Size	Default Rate	Sample Size	Default Rate	Sample Size	Default Rate	Sample Size
Proprietary	20.2	841	55.6	438	37.3	115	31.2	1,394
Public Two-Year	20.8	960	53.5	184	40.1	68	26.0	1,212
Private Two-Year	14.8	108	61.4	33	15.9	7	23.1	148
Public Four-Year	8.6	2,400	38.0	264	16.4	90	11.2	2,754
Private Four-Year	7.0	1,565	33.9	173	11.2	85	9.2	1,823
Total	11.9	5,874	47.7	1,092	26.2	365	17.9	7,331

NOTE: The default rate is defined as the weighted proportion of respondents who said that they had defaulted on their loan(s). Weights were used to offset differential sampling rates and nonresponse.

TABLE 5

DEFAULT RATES BY RACE/ETHNICITY AND ANNUAL EARNINGS

	Race/Ethnicity							
	(1) White		(2) Black		(3) Hispanic		(4) Total	
Annual Earnings[a]	Default Rate	Sample Size	Default Rate	Sample Size	Default Rate	Sample Size	Default Rate	Sample Size
<$10,000	23.3	864	65.4	367	44.4	101	36.5	1,332
$10,000-$14,999	14.7	316	55.6	62	46.7	19	22.6	397
$15,000-$24,999	11.4	804	48.5	123	20.9	42	16.5	969
≥ $25,000	7.3	1,739	36.2	151	9.8	72	9.6	1,962
Total	12.5	3,723	55.3	703	29.7	234	19.8	4,660

NOTE: The default rate is defined as the weighted proportion of respondents who said that they had defaulted on their loan(s). Weights were used to offset differential sampling rates and nonresponse.

[a]Average annual earnings refer to the two-year period beginning in the month that the loan repayment was scheduled to begin.

TABLE 6

DEFAULT RATES BY ANNUAL EARNINGS AND TYPE OF POSTSECONDARY INSTITUTION MOST RECENTLY ATTENDED

Type of Postsecondary Institution	Annual Earnings Level[a]									
	(1) <$10,000		(2) $10,000-$14,999		(3) $15,000-$24,999		(4) >$25,000		(5) Total	
	Default Rate	Sample Size	Default Rate	Sample Size	Default Rate	Sample Size	Default Rate	Sample Size	Default Rate	Sample Size
Proprietary	48.0	473	26.3	107	22.2	234	16.3	215	33.3	1,024
Public Two-Year	43.7	209	35.8	80	18.7	146	16.7	219	28.1	654
Private Two-Year	40.7	34	21.0	9	2.0	29	8.4	36	17.9	108
Public Four-Year	19.6	282	14.0	117	13.6	317	7.3	820	11.4	1,536
Private Four-Year	19.0	212	7.4	54	12.2	188	6.9	611	10.3	1,065
Total	35.4	1,210	21.5	367	16.0	914	9.3	1,901	18.9	4,392

NOTE: The default rate is defined as the weighted proportion of respondents who said that they had defaulted on their loan(s). Weights were used to offset differential sampling rates and nonresponse.

[a] Annual earnings refer to the two-year period beginning in the month that the loan repayment was scheduled to begin.

2. High School and Postsecondary School Completion

As was shown in Table 3, the default rate for high school dropouts was 56 percent, and the default rate for borrowers who did not receive a college degree or certificate was 33 percent. Further evidence about the relationship between program completion and default is shown in Table 7, in which default rates are cross-tabulated by school type and by whether the program was completed. As shown in Table 7, default rates are considerably higher among non-completers, regardless of school type. For example, among proprietary school borrowers, the default rate was 24 percent for completers and 47 percent for non-completers. Among public 4-year school borrowers, the default rate was 8 percent for completers and 17 percent for non-completers. As with earnings, default rates within completion categories were much higher in proprietary and two-year schools. This finding is particularly true of non-completers at proprietary schools.

D. DEFAULT RATES AND SOURCES OF GSL PROGRAM AND REPAYMENT INFORMATION

Tables 8 and 9 display default rates by sources of information about the GSL program and about the GSL repayment process, cross-tabulated by type of school. As shown in Table 8, 57 percent of proprietary-school borrowers reported learning about the GSL program from a postsecondary institution only, and the default rate for these borrowers was 36 percent. An examination of the distribution columns shows that proprietary- school borrowers were more likely than other borrowers to have received GSL program information only from a postsecondary institution. On the other hand, four-year school borrowers were more likely to have received information on the GSL program from multiple sources. Borrowers who received GSL program information only from a lender had the lowest default rates across all school types.

As shown in Table 9, similar patterns emerge for GSL default rates by sources of information on repayments. Proprietary-school borrowers were more likely to have received GSL repayment information only from a postsecondary institution at the time the loan was made, and they were less likely to have received repayment information from a lender. Across all school types, default rates

TABLE 7

DEFAULT RATES BY WHETHER EDUCATIONAL PROGRAM WAS COMPLETED AND BY TYPE OF POSTSECONDARY INSTITUTION MOST RECENTLY ATTENDED

Type of Postsecondary Institution	Completed Educational Program		Did Not Complete Educational Program		Total	
	Default Rate	Sample Size	Default Rate	Sample Size	Default Rate	Sample Size
Proprietary	24.2	907	47.1	480	32.1	1,387
Public Two-Year	20.5	508	30.5	661	26.1	1,169
Private Two-Year	13.7	91	39.8	55	23.5	146
Public Four-Year	8.0	1,685	17.2	1,025	11.5	2,710
Private Four-Year	6.9	1,253	15.1	549	9.4	1,802
Total	12.5	4,444	25.6	2,770	17.5	7,214

NOTE: The default rate is defined as the weighted proportion of respondents who said that they had defaulted on their loan(s). Weights were used to offset differential sampling rates and nonresponse.

TABLE 8

DEFAULT RATES BY TYPE OF POSTSECONDARY INSTITUTION MOST RECENTLY ATTENDED AND BY SOURCE OF INFORMATION ON THE GSL PROGRAM

Source of Information about GSL Program	Type of Postsecondary School						Total	
	Proprietary		Two-Year		Four-Year			
	Default Rate	Distribution (%)	Default Rate	Distribution (%)	Default Rate	Distribution (%)	Default Rate	Distribution (%)
Lender Only	18.7	8.5	15.0	9.8	6.9	10.9	10.2	10.2
Postsecondary School Only	35.7	56.5	32.6	36.0	14.7	30.6	24.1	36.6
High School Only	25.7	6.3	32.6	10.7	8.0	8.4	16.3	8.4
Friend/Relative Only	25.8	8.1	25.2	12.9	8.7	16.1	13.4	14.0
Other	30.4	3.1	21.2	4.2	9.0	3.3	15.5	3.4
More Than One Source	29.4	17.5	19.2	26.3	8.7	30.6	13.1	27.3
Total	31.6	100.0	25.9	100.0	10.3	100.0	17.2	100.0

NOTE: The default rate is defined as the weighted proportion of respondents who said that they had defaulted on their loan(s). Weights were used to offset differential sampling rates and nonresponse.

TABLE 9

DEFAULT RATES BY TYPE OF POSTSECONDARY INSTITUTION MOST RECENTLY ATTENDED AND BY SOURCE OF INFORMATION ON GSL REPAYMENT

Source of Information about GSL Program	Type of Postsecondary School						Total	
	Proprietary		Two-Year		Four-Year			
	Default Rate	Distribution (%)	Default Rate	Distribution (%)	Default Rate	Distribution (%)	Default Rate	Distribution (%)
No Information	41.2	5.5	40.9	6.9	22.4	4.4	31.1	5.1
Only from Postsecondary Institution When Loan Was Made	39.4	30.1	32.6	18.8	15.1	15.1	26.1	18.8
Only from Postsecondary Institution When Leaving Program	36.7	4.4	25.5	3.2	9.8	4.3	17.8	4.1
Only from Lending Institution When Loan Was Made	25.4	33.3	20.0	43.1	9.1	41.8	13.9	40.4
Only From Lending Institution during or When Leaving Program	28.7	17.9	31.2	20.8	9.0	24.4	15.9	22.4
From Both Postsecondary Institution and Lending Institution	28.4	8.7	22.0	7.1	10.6	10.0	15.6	9.2
Total	31.8	100.0	26.5	100.0	10.7	100.0	17.9	100.0

NOTE: The default rate is defined as the weighted proportion of respondents who said that they had defaulted on their loan(s). Weights were used to offset differential sampling rates and nonresponse.

were highest when no information about repayment was received, or when information was received only from postsecondary institutions.

E. THE IMPORTANCE OF VARIOUS FACTORS THAT LED TO DEFAULT

Respondents who reported that they had defaulted were asked to indicate the importance of various factors that led to their default: unemployment; low income; the presence of more important loans to repay; dissatisfaction with their educational program; intervening personal problems; confusion about the repayment process; and the misperception that the loan need not be repaid. For each of these factors, defaulters indicated whether they felt that the factor was very important, somewhat important, or not important in leading to default. More than one factor could be cited as a very important cause for defaulting. While these responses do not help explain why some borrowers default and others do not, they do provide useful insights into the circumstances that defaulters felt contributed most to their failure to repay their loans.

Table 10 shows the proportion of respondents who cited a reason as "very" or "somewhat" important in leading to their default, cross-tabulated by type of school. For proprietary-school defaulters, 83 percent reported that being unemployed and without income was an important factor that led to their default. Intervening personal problems was the next most frequently cited reason for default among this group. For two-year-school defaulters, the two most important factors for default were unemployment and insufficient funds. For four-year school defaulters, the two most important factors for default were insufficient funds and unemployment. Confusion about the repayment process was also cited by a significant proportion of defaulters in each of the types of schools. Relatively few defaulters reported that they were unaware that they had to repay their loans, though those who did were more likely to be in proprietary or two-year schools.

TABLE 10

IMPORTANCE OF FACTORS RELATED TO DEFAULT, BY TYPE OF POSTSECONDARY INSTITUTION

Reason for Default	Proportion Reporting That the Reason Was "Very" or "Somewhat" Important			
	(1) Proprietary School	(2) Two-Year School	(3) Four-Year School	(4) Total
Unemployed and Without Income	83	74	64	73
Working but Had Insufficient Funds	57	62	69	63
Repaying More Important Loans	37	39	39	38
Intervening Personal Problems	62	59	49	56
Confused by Repayment Process	38	41	42	41
Did Not Realize That Loan Had To Be Repaid	19	18	12	16

NOTE: Estimates are weighted to offset differential sampling rates and nonresponse.

III. A MULTIVARIATE ANALYSIS OF DEFAULT ON GUARANTEED STUDENT LOANS

This chapter presents the methodology and the results of our multivariate analysis of default on guaranteed student loans. The multivariate results confirm a number of findings from the descriptive analysis of default.[1] Several factors that were associated with default in the descriptive analysis are found to be statistically significant determinants of default in the multivariate analysis, including earnings, family size, marital status, race/ethnicity, the size of the GSL monthly payment, whether loan recipients completed high school, and whether loan recipients completed their postsecondary program.

We begin the chapter by specifying a conceptual framework for default, to facilitate understanding default behavior and to provide a basis for interpreting the results of the empirical analysis. We then discuss our choice of an empirical model and the independent variables for the model. Estimation results for the full sample and for key subgroups are then presented and discussed.

A. A CONCEPTUAL FRAMEWORK FOR DEFAULT

As a preliminary step towards understanding the decision to default on a student loan, it is useful to examine the factors underlying the decision to obtain a student loan. Individuals who are deciding whether to obtain a postsecondary education will compare its benefits and costs, and will decide to obtain a postsecondary education if the benefits exceed the costs. The benefits of a postsecondary education are derived primarily in the form of enhanced skills and thus higher earnings. Other benefits include the cultural and social amenities that might be enjoyed while in school, and the greater enjoyment of life activities that may be derived from a more extensive education.

---

[1]Multivariate analysis was conducted in order to better understand the determinants of default. The descriptive analysis presented earlier is useful for indicating associations between default and certain characteristics. Multivariate analysis is required to confirm whether the associations indicated in the descriptive analysis are due to the characteristics themselves or to the underlying correlations of the characteristics with other determinants of default.

The costs of postsecondary education include foregone earnings from not working, tuition costs, and out-of-pocket expenses, such as transportation and child care. For most individuals, the benefits of a postsecondary education over their lifetimes will more than repay the costs of obtaining the education. However, the benefits of education will be enjoyed in the future, after schooling is completed, whereas the costs of schooling must be paid in the present, while schooling is ongoing. This imbalance is unimportant for wealthy individuals or youths whose parents have adequate finances to absorb the costs of schooling, but individuals who are less financially secure must supplement their own contributions or parental contributions with other funds to meet the costs of schooling.

The Guaranteed Student Loan program was created specifically to provide a mechanism for enabling persons from low and moderate-income backgrounds to meet the costs of schooling. In principle, because the loan need not be repaid until schooling ends, students can draw on the increased future earnings generated by their educational experience to make their repayment. Loan recipients can benefit if their enhanced earnings more than outweigh their loan payments, and the economy as a whole can benefit from the student loan program if it creates a more productive workforce.

Two points can be made about the impact of student loans on access to postsecondary education. First, by lowering the effective cost of schooling relative to the benefits of schooling, the student loan program should increase the proportion of persons who receive a postsecondary education.[2] Second, to the extent that student loans are granted to those who are then able to enhance their skills and future earnings significantly, loan payments would represent only a modest financial burden, and the risks of default should likewise be modest. However, to the extent that student loans are granted to those who are able to benefit only marginally from schooling and whose earnings rise little or not at all, loan payments may represent a more serious financial burden, and the risks of default for these

[2]No direct evidence is available on the effect of the GSL program on college enrollment. However, Manski and Wise (1983) found that the Basic Education Opportunity Grant (BEOG) program had significant effects on college enrollments after its introduction in 1973. Most of the estimated impact on enrollments was concentrated among proprietary and two-year schools.

borrowers may be high. Moreover, if borrowers who are at greater risk of defaulting due to their own earnings are also those from less financially secure backgrounds, it would be difficult for the parents of these youths to assist with loan payments or provide financial assistance in other respects.

Because there is no perfectly accurate mechanism for distinguishing which students who receive loans are likely to enhance their future earnings significantly, it is difficult to specify loan qualification standards that would reduce the risk of default without simultaneously denying access to postsecondary education for some persons who would benefit greatly. The student loan program thus accomplishes its objective of increasing access to postsecondary education at the price of providing relatively risky loans to some individuals who may benefit little from additional education.

Understanding the factors underlying the types of individuals who receive student loans is helpful for understanding the factors underlying default on student loans. Students from low and moderate-income backgrounds who do well academically in high school are more likely to attend and graduate from four-year colleges, rather than two-year colleges or proprietary schools. Because the skills imparted in a four-year education are generally more valuable in the labor market than those imparted in a two-year or vocational education, we would thus expect that students who graduate from four-year colleges would enhance their earnings significantly and default less frequently. The same reasoning applies for students who go on to postgraduate education, and we would thus expect that postgraduate students would default less frequently than would four-year graduates who enter the workforce immediately.

For two reasons, students who attend two-year and proprietary schools may be more likely to default than students who attend four-year schools. First, proprietary and two-year students are likely to choose to attend these types of schools because they feel that their academic abilities are better matched with the less rigorous academic levels of the schools. Because such students are less able to learn the more advanced skills that are well-rewarded in the labor market, an increase in their

future earnings would likewise be smaller, which may give rise to a burdensome loan payment and a higher likelihood of default.

Second, the highly specialized nature of vocational training relative to four-year educational programs creates a risk that students may leave school and be unable to find employment in their particular area of training. Though four-year students face this employment risk as well, four-year educational programs are more likely to impart a more diverse range of basic skills that enable graduates to better adapt to the demands of various types of jobs. Conversely, proprietary and two-year school graduates are generally trained to perform specific functions, and a lack of demand for these specific functions could hinder employment prospects greatly. This greater risk of unemployment or underemployment leads directly to a greater risk of default. The situation is most severe for students who fail to complete their educational program, because they will face loan obligations despite the fact that their incomplete training hinders their earnings opportunities to make repayment.

Our conceptual framework for understanding default can thus be summarized briefly as follows. By its nature, the GSL program is attractive to individuals who would otherwise be unable to finance their postsecondary education. However, some individuals will benefit more than others from their postsecondary education. Individuals from low or moderate-income backgrounds who, for whatever reason, do not benefit greatly from their education are at greater risk of default. Individuals from high-wealth backgrounds are less likely to obtain student loans, but if they do obtain loans they are more likely to attend four-year schools and enhance their earnings significantly, thus leading to a lower risk of default.

On the basis of our framework, we can sketch several empirical hypotheses about the factors associated with default. First, we would expect that low parental income is positively associated with default. Low parental income may be correlated with lower academic ability and attendance at secondary schools of lower quality, both of which may prompt individuals to choose or be able to gain

acceptance only to proprietary or two-year schools, which in turn would enhance their future earnings only by modest amounts relative to students at four-year schools. Low parental income also means that loan recipients who find themselves in financial difficulties may not be able to rely on their parents for financial assistance. In general, minority students have low parental income (see Table 1) and, consequently, would be expected to default more often.

Second, for similar reasons, we would expect that attendance at proprietary and two-year schools is positively correlated with default. Attendance at proprietary and two-year schools is correlated with lower academic ability, which leads to modest increases in earnings and thus a higher risk of default. Moreover, proprietary and two-year school training may embody greater unemployment risk, which would also lead to a greater risk of default.

Third, failing to complete a postsecondary educational program should be positively correlated with default. Loan recipients who fail to complete their educational program will face loan repayment obligations but will generally not enjoy the earnings enhancements necessary to meet the loan obligations without burden.

In our descriptive analysis, we observed that loan recipients from minority groups, those with low parental income, those with low earnings, those who failed to complete their programs, and those who had attended proprietary and two-year schools generally had higher default rates. These findings are consistent with the predictions from our conceptual framework. The remainder of this chapter examines the determinants of default in a multivariate setting. As we will see, the results of the multivariate analysis are also consistent with our conceptual framework.

B. AN EMPIRICAL MODEL OF DEFAULT

The purpose of our multivariate analysis is to assess the relative importance of various characteristics of borrowers that are associated with default. They include primarily demographic and socioeconomic characteristics, educational attainment characteristics, and earnings and ability- to-pay

characteristics. We will also examine whether sources of information about the GSL program and the GSL repayment process affect the likelihood of default.

The variable that we wish to explain is default status, which can assume the value of zero for individuals who have not defaulted, and one for individuals who have defaulted. We assume further that the likelihood that a particular individual is a defaulter depends on his or her characteristics, the set of which we abbreviate as X. A useful and computationally convenient representation of the empirical relationship between the probability that an individual defaults and his or her characteristics is the following model:

$$(1)\quad \text{Prob(default)} = 1/\{1 + \exp(-Xb)\},$$

where "exp" denotes the base of the natural logarithms. The model represented by equation (1) is termed a "logit" model (Maddala, 1983). The set of parameters b are typically estimated with maximum likelihood techniques when the available data are at the individual level.

C. VARIABLES FOR THE EMPIRICAL MODEL OF DEFAULT

In this section we discuss the variables that comprise our empirical model and the various methods by which the data were transformed to arrive at the final form of these variables.

Roughly 60 percent of the SLRS sample were respondents who were repaying their loans. That is, these respondents had neither defaulted nor paid in full at the time the survey sample was drawn. However, for the logit model to be applicable, we must be able to code whether or not respondents defaulted. One possible strategy is to drop cases in which respondents are still in repayment. However, considerable data would be lost if this strategy were adopted. An attractive alternative is to use a particular definition of default that helps avoid this timing problem. Our definition is whether respondents have defaulted within two years after their repayment was due to begin.[3] All

[3]In our preliminary analysis, we found that, among individuals who had defaulted, more than 60 percent reported that they had defaulted within two years after their repayment was scheduled to begin.

individuals currently in the repayment period were coded as nondefaulters if they had a repayment period of at least two years and had not defaulted during that period. We excluded from the empirical analysis individuals who had been in repayment for less than two years and individuals who were still in school when the survey was administered.[4]

1. Independent Variables

Our choices for independent variables fall roughly into four groups: (1) demographic and family background variables (age, gender, ethnicity, parental income, and parental education), (2) ability-to-pay variables (earnings, monthly GSL payment, marital status, and family size), (3) educational characteristics (whether recipients completed high school, whether they last attended a proprietary school, whether they last attended a two-year school, and whether they completed their postsecondary program), and (4) GSL information characteristics (whether respondents learned about the program only from a postsecondary institution, whether respondents learned about the program only from a lending institution, whether respondents learned about the program from another source, whether respondents did not receive information about repayment, whether respondents learned about repayment from a postsecondary institution only, and whether respondents learned about repayment from a lending institution only). Means and standard errors for the model variables are shown in Table 11, and the simple correlation matrix for the model variables is shown in Table 12.

The pattern of correlation between the default variable and the independent variables is analogous to the results found in our descriptive analysis. Default is positively correlated with race/ethnicity, low parental income, not completing high school, not completing the most recent postsecondary program, and attending a proprietary or two-year school. Earnings are negatively associated with default, race/ethnicity, not completing high school or the most recent postsecondary school program, and attending a proprietary or two-year school.

---

[4]We excluded 803 respondents who were still in school and 1,539 respondents whose repayment period was less than two years. For the sample as a whole, 4,752 cases were in repayment and would have been excluded under the more stringent default criterion.

TABLE 11

DESCRIPTIVE STATISTICS FOR THE MULTIVARIATE ANALYSIS
Means (Standard Errors)

Variable	Sample Definition						
	(1) Entire	(2) White	(3) Black	(4) Hispanic	(5) Proprietary	(6) Two-Year	(7) Four-Year
Earnings	$14,793 ($12,490)	$15,411 ($12,740)	$10,096 ($10,184)	$11,241 ($9,437)	$9,437 ($8,786)	$10,596 ($8,471)	$16,840 ($13,284)
GSL Monthly Payments	$75.59 ($49.20)	$76.35 ($49.17)	$66.73 ($39.73)	$76.51 ($62.85)	$60.90 ($35.80)	$55.91 ($23.00)	$82.81 ($53.64)
Number of Dependents	0.54 (1.08)	0.48 (1.02)	1.00 (1.35)	0.74 (1.33)	0.72 (1.21)	0.71 (1.27)	0.46 (1.00)
Married	0.25 (0.43)	0.26 (0.44)	0.20 (0.40)	0.28 (0.45)	0.23 (0.34)	0.24 (0.43)	0.26 (0.44)
Black	.12 (.32)	--	--	--	.26 (.44)	.11 (.31)	.07 (.26)
Hispanic	0.04 (0.20)	--	--	--	.08 (.27)	.05 (.22)	0.03 (0.18)
Parental Income <$17,000	0.25 (0.44)	0.21 (0.41)	0.59 (0.49)	0.48 (0.50)	0.38 (0.49)	0.36 (0.48)	0.21 (0.40)
Parental Income $17,000-$30,000	0.31 (0.46)	0.21 (0.41)	0.24 (0.43)	0.26 (0.44)	0.31 (0.46)	0.33 (0.47)	0.30 (0.46)
Did Not Complete High School	0.05 (0.22)	0.04 (0.19)	0.13 (0.33)	0.17 (0.37)	0.17 (0.37)	0.08 (0.27)	0.02 (0.12)
Did Not Complete Most Recent Postsecondary Program	0.23 (0.42)	0.21 (0.41)	0.36 (0.48)	0.31 (0.46)	0.25 (0.43)	0.42 (0.49)	0.18 (0.38)
Proprietary School	0.18 (0.38)	0.14 (0.35)	0.42 (0.49)	0.34 (0.47)	--	--	--
Two-Year School	0.14 (0.35)	0.15 (0.35)	0.13 (0.34)	0.15 (0.36)	--	--	--

TABLE 11 (continued)

Variable	Sample Definition (1) Entire	(2) White	(3) Black	(4) Hispanic	(5) Proprietary	(6) Two-Year	(7) Four-Year
GSL Information Only from Lender	0.36 (0.48)	0.33 (0.47)	0.52 (0.50)	0.51 (0.50)	0.54 (0.50)	0.35 (0.48)	0.31 (0.46)
GSL Information Only from Post-Secondary Institution	0.11 (0.31)	0.12 (0.32)	0.08 (0.27)	0.10 (0.30)	0.10 (0.30)	0.12 (0.33)	0.11 (0.32)
No GSL Repayment Information	0.03 (0.17)	0.03 (0.16)	0.05 (0.24)	0.04 (0.19)	0.03 (0.17)	0.06 (0.23)	0.03 (0.16)
GSL Repayment Information Only from Lender	0.67 (0.47)	0.69 (0.46)	0.53 (0.50)	0.55 (0.50)	0.55 (0.50)	0.68 (0.47)	0.70 (0.46)
GSL Repayment Information Only from Postsecondary Institution	0.21 (0.41)	0.19 (0.40)	0.34 (0.47)	0.34 (0.47)	0.32 (0.47)	0.19 (0.40)	0.19 (0.39)
Sample Size	4,304	3,606	521	198	911	683	2,710

TABLE 12

SIMPLE CORRELATION MATRIX

Var. #	(1) Default	(2) Earnings	(3) GSL Monthly Payment	(4) Depen-dents	(5) Married	(6) Black	(7) Hispanic	(8) Parent Income $17K	(9) Parent Income $17-$30K	(10) Did Not Complete High School	(11) Did Not Complete Post-secondary	(12) Pro-prietary	(13) Two-Year	(14) Program Info/ Lender Only	(15) Program Info/Post-secondary Only	(16) No Repayment Info	(17) Repayment Info/Lender Only	(18) Repayment Info/Post-secondary
1.	1.00																	
2.	-0.16	1.00																
3.	-0.02	0.21	1.00															
4.	0.15	0.03	-0.00	1.00														
5.	-0.02	0.06	0.03	0.54	1.00													
6.	0.24	-0.12	-0.07	0.14	-0.04	1.00												
7.	0.06	-0.04	-0.02	0.04	0.01	-0.04	1.00											
8.	0.14	-0.09	-0.08	0.21	0.08	0.23	0.11	1.00										
9.	-.01	-0.00	-0.01	-0.03	-0.02	-0.04	-0.03	-0.41	1.00									
10.	0.22	-0.11	-0.07	0.17	0.03	0.11	0.11	0.12	-0.01	1.00								
11.	0.16	-0.15	-0.14	0.04	-0.05	0.11	0.04	0.07	0.04	0.11	1.00							
12.	0.15	-0.20	-0.15	0.08	-0.02	0.21	0.09	0.13	0.00	0.25	0.03	1.00						
13.	0.06	-0.12	-0.16	0.06	-0.01	-0.01	0.00	0.10	0.03	0.05	0.19	-0.20	1.00					
14.	0.09	-0.03	-0.06	-0.05	0.06	0.11	0.07	0.11	-0.02	0.15	0.02	0.18	-0.01	1.00				
15.	-0.04	-0.01	-0.03	-0.09	-0.03	-0.03	-0.01	-0.06	0.00	-0.04	-0.01	-0.02	0.01	-0.26	1.00			
16.	0.07	-0.03	-0.02	0.04	0.00	0.05	0.01	0.07	-0.05	0.03	0.07	-0.00	0.06	-0.01	0.00	1.00		
17.	-0.09	0.01	-0.00	-0.07	-0.00	-0.10	-0.06	0.11	0.03	-0.09	-0.03	-0.11	0.01	-0.24	0.12	-0.25	1.00	
18.	0.07	-0.04	-0.04	0.04	-0.02	0.10	0.07	0.08	-0.03	0.09	0.03	0.12	-0.02	0.27	-0.09	-0.09	-0.75	1.00

NOTE: Variable numbers on the left correspond to column numbers across the top of the table. All correlations are weighted.

D. HYPOTHESIS TESTS IN THE MULTIVARIATE MODEL

The estimated coefficients in our empirical model can be tested to determine whether they differ significantly from zero based on a t-test.[5] If an estimated coefficient divided by its standard error exceeds the critical value of the t-statistic, we would conclude that the observed coefficient value is not likely to differ from zero solely by chance.[6]

We examine three hypotheses in particular. The first is the ability-to-pay hypothesis, which states roughly that the income level of loan recipients is an important determining factor for default. Because the ability to make a GSL loan payment may also depend on marital status and family size, we also examine the t-statistics for these variables as evidence for the ability-to-pay hypothesis.

Considerable policy attention has focused recently on the problem of high default rates among students of proprietary and two-year schools. The second hypothesis to be examined is whether default rates differ by type of school. We test the extent to which a particular type of school is associated with default by using t-tests on the two school-type variables (proprietary schools and two-year schools).[7]

The third hypothesis that we examine is whether sources of information about the GSL program and the repayment process are related to default. The results of this test are of interest because policymakers may be able to control to some degree the channels by which potential loan recipients are informed about their responsibilities as borrowers and the penalties for defaulting. Greater

---

[5]Formally, the hypothesis tests that we use here are asymptotically correct (that is, they are derived on the basis of very large sample sizes). Hypothesis tests for such models as equation (1) based on small sample sizes generally do not exist.

[6]For these sample sizes, the hypothesis that a coefficient equals zero can be rejected at the 95 percent confidence level if its t-statistic is greater than 2.00. At the 99 percent confidence level, the critical t-statistic is 2.60.

[7]What we are testing is whether the type of school attended has an effect on default independent of the other variables in the model, including earnings. Obviously, the type of school attended also affects earnings, which affect default, but we did not estimate this type of indirect relationship.

emphasis could be placed on sources of information that appear to be more effective at reducing default.

E. ESTIMATION RESULTS

The results of the basic estimation are shown in Table 13 for two variants of our empirical model.[8] Column 1 of Table 13 shows the estimation results for a basic linear specification (Model 1) of the empirical model. For the linear specification, we include all variables but allow no interactions among the separate variables. Eleven of the independent variables are statistically significant at the 95 percent confidence level. Average earnings in the two-year repayment period and whether respondents were married when their loans came due were negatively associated with default, while the monthly GSL payment, whether respondents did not complete high school, whether respondents did not complete their postsecondary program, the number of their dependents when their loan came due, whether they were black or Hispanic, whether they had last attended a proprietary school or a two-year school, and whether their parents' income was less than $17,000 per year were positively related to default.[9]

In column 2 of Table 13, we create a nonlinear specification by interacting the two race/ethnicity variables with the proprietary school variable and with earnings (we refer to this version as the interaction model, or Model 2). The interaction facilitates determining whether the effect of these variables differs across racial subgroups. The magnitudes of the estimated interaction coefficients suggest that earnings and proprietary-school effects do differ for different race/ethnicity subgroups. However, the statistical significance of the results is low. Hispanic students from proprietary schools appear to be more likely to default than black or white students from proprietary schools. Hispanic

---

[8]The model was estimated with SAS Proc LOGIST. Cases were weighted to offset differential sampling probabilities and nonresponse.

[9]In preliminary runs, we included 162 default cases for which we had no information about the time period of default, under the assumption that default occurred within two years for these cases. We found only a slight effect of including these cases on the estimated coefficients.

TABLE 13

ESTIMATION RESULTS FOR THE ENTIRE SAMPLE
(t-statistics in parentheses)

DEPENDENT VARIABLE: DEFAULT WITHIN TWO YEARS = 1

Variable	(1) Linear Model	(2) Interaction Model
Earnings	-6.34 (-4.8)	-6.49 (-4.6)
GSL Monthly Payment	0.48 (3.6)	0.48 (3.5)
Number of Dependents	0.33 (5.6)	0.34 (5.7)
Married	-0.55 (-3.1)	-0.56 (-3.2)
Black	1.24 (7.9)	1.19 (4.7)
Hispanic	0.53 (2.20)	0.55 (1.43)
Parental Income <$17,000	0.50 (2.4)	0.49 (2.3)
Parental Income $17,000-$30,000	0.28 (1.4)	0.28 (1.3)
Did Not Complete High School	1.12 (5.9)	1.10 (5.8)
Did Not Complete Most Recent Postsecondary Program	0.77 (5.8)	0.77 (5.8)
Proprietary School	0.48 (2.9)	0.49 (2.50)
Two-Year School	0.40 (2.3)	0.40 (2.3)
GSL Program Information Only from Lender	0.03 (0.2)	0.03 (0.2)
GSL Program Information Only from Postsecondary Institution	-0.41 (-1.5)	-0.42 (-1.5)

TABLE 13 (continued)

Variable	(1) Linear Model	(2) Interaction Model
No GLS Repayment Information	0.50 (1.50)	0.50 (1.5)
GSL Repayment Information Only from Lender	-0.24 (-1.1)	-0.23 (1.0)
GSL Repayment Information Only from Postsecondary Institution	-0.02 (-0.1)	-.03 (-0.1)
Black x Proprietary	--	-0.06 (-0.2)
Hispanic x Proprietary	--	0.18 (0.4)
Black x Earnings	--	1.34 (0.7)
Hispanic x Earnings	--	-1.65 (0.5)
Percent Correctly Predicted	88.8	88.8
Sample Size	4,304	4,304
Number of Defaulters	487	487

NOTE: Weights are used in the estimation to offset differential sampling rates and nonresponse. Other independent variables included in all runs but not reported here are age, gender, the proportion of months with no earnings in the two-year repayment period, two dummy variables for whether the respondent's parents dropped out of high school or finished high school, missing data flags for earnings, the monthly GSL payment, and parental income, and an intercept. The estimated coefficients for these variables were statistically insignificant in all runs.

students are less likely to default than white students with equivalent post-schooling earnings, whereas black students are more likely to default than white students with equivalent post-schooling earnings. Below, we analyze subgroup models in more detail, to assess whether other differences are evident across race/ethnicity categories and across different types of schools.

1. Hypothesis Tests and Discussion

With respect to our three hypotheses of particular interest, the results in Table 13 show clearly that a borrower's earnings level after leaving school is a powerful determinant of default. The t-statistics for the earnings coefficient of 4.6 in Model 1 is highly significant at any reasonable confidence level. Moreover, family size and marital status are significant predictors of default in both models, which is evidence that resources and needs play a crucial role in the default decision. The results in Table 13 also confirm the hypothesis that the type of school is a determinant of default. In both the linear and interaction models, the proprietary and two-year school coefficients are positive and statistically significant. These results provide evidence that the type of school is related to default after the characteristics, background factors, and information sources of students are controlled for.

The third hypothesis is that sources of information about the GSL program and the repayment process are associated with default. As detailed in our previous descriptive discussion (in Section II.C and Tables 8 and 9), default rates were generally higher among loan recipients who reported that they had received information on the GSL program only from a postsecondary institution, as well as among loan recipients who reported that they had not received information about GSL repayment or had received it only from their postsecondary institution at the time the loan was made. However, our results about the influence of channels of information on default are mixed when other factors are taken into account. In Model 1, the estimated coefficients for learning about the GSL program only from a lender and for learning about the GSL program only from a postsecondary school indicate effects that are opposite from the effects revealed by the descriptive findings. Neither

coefficient is statistically significant. The estimated coefficients for not receiving any information about repayment or receiving information only from a lender are consistent with the results of the descriptive analysis, but, again, neither is statistically significant. It appears that whatever the impact of the various sources of information on default, it is being absorbed by the characteristics of borrowers and other factors in the model.

2. Subgroup Results

We estimated Model 1 for six defined subgroups--three by race/ethnicity (white, black, and Hispanic) and three by type of school (proprietary, two-year, and four-year). The results are shown in Tables 14 and 15.

In general, the smaller sample sizes for the subgroups reduce the power of our statistical inferences. Nonetheless, several interesting differences emerge from a comparison of the subgroup results. For the race/ethnicity subgroups, it is evident that the effect of earnings on default is lower for blacks and higher for Hispanics. That is, after various other factors that affect default are controlled for, black borrowers are more likely to default than white borrowers with equivalent earnings, and Hispanic borrowers are less likely to default. A similar pattern was found in the interaction Model (Model 2). The effects of attending a proprietary school and failing to complete the postsecondary program on default are also noticeably larger for Hispanic borrowers.

As shown in Table 15, it is clear that the effects of a number of factors on default differ by type of school. For example, the relationship between GSL monthly payment and default is stronger for two-year school students than for proprietary and four-year-school students. The relationship between the number of dependents, whether borrowers were married, and whether borrowers completed high school was weaker for proprietary-school students than for students from four-year schools. Hispanic borrowers were more likely to default if they had attended proprietary schools, which mirrors the result noted for the Hispanic subgroup model (column 3 of Table 14) and in the interaction model (column 2 of Table 13).

TABLE 14

ESTIMATION RESULTS FOR RACE/ETHNICITY SUBGROUPS

DEPENDENT VARIABLE: DEFAULT WITHIN TWO YEARS = 1

Variable	(1) White	(2) Black	(3) Hispanic
Earnings	-6.37	-5.12	-9.87
	(-4.0)	(-2.1)	(-1.7)
GSL Monthly Payment	0.53	0.56	0.49
	(3.3)	(1.8)	(1.0)
Number of Dependents	0.33	0.36	0.23
	(4.2)	(3.7)	(1.1)
Married	-0.51	-0.92	-0.06
	(-2.3)	(-2.7)	(-0.1)
Parental Income <$17,000	0.77	-0.59	0.49
	(3.1)	(-1.4)	(0.5)
Parental Income $17,000-$30,000	0.42	-0.56	-0.19
	(1.7)	(1.2)	(-0.2)
Did Not Complete High School	1.22	1.01	0.87
	(5.0)	(3.2)	(1.4)
Did Not Complete Most Recent Postsecondary Program	0.92	0.25	1.54
	(5.6)	(1.1)	(2.7)
Proprietary School	0.46	0.45	1.08
	(2.2)	(1.7)	(1.6)
Two-Year School	0.40	0.36	0.90
	(1.9)	(1.0)	(1.2)
GSL Program Information Only from Lender	0.14	-0.26	0.18
	(0.7)	(-0.8)	(0.3)
GSL Program Information Only from Postsecondary Institution	-0.41	-0.08	-2.06
	(-1.2)	(-0.2)	(-1.4)
No GSL Repayment Information	0.13	1.00	2.46
	(0.3)	(1.7)	(1.4)

TABLE 14 (continued)

	(1) White	(2) Black	(3) Hispanic
GSL Repayment Information Only from Lender	-0.41 (-1.6)	0.8 (0.2)	1.38 (0.90)
GSL Repayment Information Only from Postsecondary Institution	-0.24 (-0.8)	0.38 (-0.8)	1.37 (0.9)
Percent Correctly Predicted	92.1	69.3	82.8
Sample Size	3,606	521	198
Number of Defaulters	278	173	43

NOTE: Weights are used in the estimation to offset differential sampling rates and nonresponse. Other independent variables included in all runs but not reported here are age, gender, the proportion of months with no earnings in the two-year repayment period, two dummy variables for whether the respondent's parents dropped out of high school or finished high school, missing data flags for earnings, the monthly GSL payment, and parental income, and an intercept. The estimated coefficients for these variables were statistically insignificant in all runs.

TABLE 15

ESTIMATION RESULTS FOR POSTSECONDARY SCHOOL SUBGROUPS

DEPENDENT VARIABLE: DEFAULT WITHIN TWO YEARS = 1

Variable	(1) Proprietary	(2) Two-Year	(3) Four-Year
Earnings	-4.54	-4.41	-8.01
	(-1.6)	(-1.4)	(-4.3)
GSL Monthly Payment	0.38	1.20	0.52
	(1.1)	(2.3)	(3.2)
Number of Dependents	0.21	0.46	0.46
	(2.4)	(3.7)	(4.5)
Married	-0.42	-0.38	-0.96
	(-1.6)	(0.9)	(-3.2)
Black	1.21	0.95	1.45
	(5.4)	(2.4)	(5.7)
Hispanic	0.66	0.65	0.34
	(1.9)	(1.3)	(0.8)
Parental Income <$17,000	0.53	0.82	0.54
	(1.3)	(1.6)	(1.8)
Parental Income $17,000 - $30,000	0.75	0.44	0.05
	(1.9)	(0.8)	(0.2)
Did Not Complete High School	0.87	1.61	1.33
	(3.7)	(4.3)	(2.9)
Did Not Complete Most Recent Postsecondary Program	0.45	0.67	0.93
	(2.1)	(2.4)	(4.5)
GSL Program Information Only from Lender	-0.13	0.54	0.12
	(-0.5)	(1.3)	(0.5)
GSL Program Information Only from Postsecondary Institution	-0.53	0.42	-0.62
	(-1.1)	(0.8)	(-1.5)
No GSL Repayment Information	0.76	0.51	0.41
	(1.3)	(0.8)	(0.8)

TABLE 15 (continued)

Variable	(1) Proprietary	(2) Two-Year	(3) Four-Year
GSL Repayment Information Only from Lender	-0.44 (-1.2)	-0.28 (0.5)	-0.08 (-0.2)
GSL Repayment Information Only from Postsecondary Institution	-0.01 (0.0)	-0.07 (0.1)	-0.11 (-0.3)
Percent Correctly Predicted	79.3	86.4	93.1
Sample Size	911	683	2,710
Number of Defaulters	197	101	189

NOTE: Weights are used in the estimation to offset differential sampling rates and nonresponse. Other independent variables included in all runs but not reported here are age, gender, the proportion of months with no earnings in the two-year repayment period, two dummy variables for whether the respondent's parents dropped out of high school or finished high school, missing data flags for earnings, the monthly GSL payment, and parental income, and an intercept. The estimated coefficients for these variables were insignificant.

F. IMPACT ESTIMATES

The estimated coefficients in our empirical model represent the effect of an explanatory variable on the logarithm of the odds that a respondent defaults. We have converted a number of the more important coefficients into impact estimates that are more intuitive and relevant for policy purposes.[10] Impact estimates represent how the default probability would be affected if all sample members were in the category of the variable represented by zero or the category of the variable represented by one (for example, proprietary school versus nonproprietary school, did not complete high school versus completed high school, and white versus black), with all other individual characteristics remaining the same.

The (weighted) average probability of default within two years for our empirical sample is 8.1 percent. As shown in Table 16, increasing annual earnings by 50 percent reduces the probability of default within two years by 3.5 percentage points, or by 43 percent of the average default probability. For example, if a borrower in our sample had an earnings level that was 50 percent greater than the average earnings for the entire sample, and other characteristics of the borrower were "average," we would predict that the borrower would have a 4.6 percent probability of defaulting within two years, as opposed to the sample average of 8.1 percent. Similarly, increasing the monthly GSL payment by 50 percent increases the probability of default by 1.3 percentage points, or by 16 percent of the average default probability.

The largest impact in Table 16 is for the black variable. The change in the probability of default for blacks of 10.6 percentage points equals 131 percent of the average default probability. Not completing high school increases the predicted default probability by 9.5 percentage points, or by 117 percent of the average. Not completing the most recent postsecondary program increases the average

[10]The technique for computing impact estimates for variables that take on binary values entails using the estimated parameters and the values of individual characteristics to compute two predicted probabilities of default for each sample member, one for each value assumed by the binary variable. The two predicted probabilities are then averaged for the entire sample, and the difference in the average probabilities is reported.

TABLE 16

IMPACT ESTIMATES FOR THE LINEAR MODEL OF DEFAULT

Variable	(1) Comparison	(2) Change in the Probability of Default
Earnings	+50%	-3.5
GSL Monthly Payment	+50%	+1.3
Black	0 to 1	+10.6
Hispanic	0 to 1	+3.8
Proprietary School	0 to 1	+3.2
Two-Year School	0 to 1	+2.7
Did Not Complete High School	0 to 1	+9.5
Did Not Complete Postsecondary Program	0 to 1	+5.3
Black/Proprietary	0/0 to 1/1	+16.3
Married/No Dependents	0/0 to 1/3	+3.1

NOTE: All calculations are based on Model 1 (column 1 of Table 13).

default probability by 5.3 percentage points. or by 65 percent of the average. The proprietary school impact is 3.2 percentage points, or 43 percent of the average, and the two-year school impact is 2.7 percentage points, or 33 percent of the average.[11]

We can use the same conversion technique to analyze the impacts of combinations of variables. As shown in Table 16, the impact of combining the black variable with proprietary- school attendance is an increase in the predicted default rate of 16.3 percentage points, or 201 percent of the sample average.[12] And a married borrower with two children at the time repayment is scheduled to begin has a predicted default probability that is 3.1 percentage points higher than the probability for an unmarried borrower with no children (10.4 percent versus 7.3 percent).

[11]The calculated impacts of default rates for attending proprietary and two-year schools are based on actual post-school earnings levels. This may understate the effects of school type on default because earnings and school type are correlated. Appendix A discusses an alternative approach for assessing the effects of school type on default. The alternative approach allows predicted post-school earnings to vary depending on the type of school attended. The school type impact is then calculated on the basis of both the direct school type effect and the indirect school type effect by way of earnings. Using this alternative approach, the proprietary school impact is 4.9 percentage points, and the two-year impact is 3.7 percentage points.

[12]Due to the nonlinearity of the logit model, the combined impact of black and proprietary school does not equal the sum of the separate impacts for black and proprietary school.

REFERENCES

Greene, L. "An Economic Analysis of Student Loan Default." *Educational Evaluation and Policy Analysis*, vol. 11, no. 1, spring 1989, pp. 61-68.

Knight, S., et al. "National Postsecondary Student Aid Study, Student Loan Recipient Survey, Methodological Report and Users Manual." Chicago: National Opinion Research Corporation, November 1988.

Maddala, G. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge, UK: Cambridge University Press, 1983.

Manski, C., and D. Wise. *College Choice in America*. Cambridge, MA: Harvard University Press, 1983.

U.S. Department of Education. *FY1987 GSLP Data Book*. Washington, D.C., 1988.

U.S. General Accounting Office. "Preliminary Analysis of Student Loan Borrowers and Defaulters." Washington, D.C.: U.S. Government Printing Office, June 1988 (GAO/HRD-88-11-112BR).

Wilms, W., R. Moore, and R. Bolus. "Whose Fault is Default? A Study of the Impact of Student Characteristics on Guaranteed Student Loan Default Rates in California." *Educational Evaluation and Policy Analysis*, vol. 9, no. 1, spring 1987, pp. 41-54.

APPENDIX A

ADDITIONAL ANALYSIS OF THE EFFECT OF SCHOOL TYPE AND POST-SCHOOL EARNINGS ON DEFAULT

Chapter III presented the impacts on default of a number of factors considered separately. The impacts on default of several factors considered jointly were also presented. Two important factors that were not considered jointly in Chapter III were school type and earnings. Table 12 shows that the type of postsecondary school attended is correlated with post-school earnings and, therefore, with default. This *indirect* effect of school type on default can be considered to be part of the total effect of school type on default.

In this appendix, we present the results of an alternative approach that allows for direct and indirect effects of school type on default. The first step was to estimate a linear regression model of the effect of school type on earnings, controlling for a number of background characteristics--age, sex, race/ethnicity, parents' education, and high school and postsecondary school completion. The estimated school type coefficients in the earnings equation are statistically significant (see Table A-1) and indicate that students who last attended proprietary and two-year schools had annual earnings that were $5,000 and $3,700 less than students who last attended four-year schools, after adjusting for background characteristics. The second step was to combine the predictions from the estimated earnings model with the estimated default model.[1] This approach yielded the following estimates of the magnitude of the indirect effect of school type on default:

[1]If the default model had been estimated using ordinary least squares regression, the indirect effect of school type on default could have been estimated by multiplying the school type earnings differentials estimated from the earnings model by the estimated coefficient on earnings in the default model. However, since the default model is estimated using logistic regression, a two-step process is needed to generate the equivalent estimate. First, the estimated regression model of earnings is used to generate three predicted earnings levels for each borrower, under the assumptions that the borrower had attended a four-year school, a two-year school, and a proprietary school. Second, the predicted effects on the probability of default of attending each of the three school types are calculated using the estimated default model and the predicted earnings levels from step one.

TABLE A-1

ESTIMATION RESULTS FOR A REGRESSION MODEL
OF POST-SCHOOL EARNINGS

Explanatory Variable	Mean (S.E.)	Model
Constant	--	4.2 (4.5)
Last Attended a Proprietary School	0.18 (0.38)	-5.0 (-9.6)
Last Attended a Two-Year College	0.14 (0.35)	-3.7 (-7.0)
Age	24.8 (6.1)	0.3 (10.3)
Female	.50 (.50)	-3.9 (-10.3)
Black	.12 (.32)	-2.5 (-4.1)
Hispanic	.05 (.22)	-0.6 (-0.7)
Neither Parent Completed High School	.11 (.31)	-2.1 (-3.1)
At Least One Parent Completed High School	.57 (.50)	-1.0 (-2.5)
Respondent Did Not Complete High School	.05 (.22)	-2.9 (-3.7)
Respondent Did Not Complete Postsecondary Program	.25 (.43)	-2.8 (6.4)
R^2		.136
N		4,092

NOTES: Estimated coefficients are in units of thousands of dollars of earnings. t-statistics are in parentheses. The regression was run using sample weights to offset differential sampling rates and nonresponse.

- The predicted default probability for proprietary schools compared to four-year schools increased 1.6 percentage points when indirect effects were included, or half again as much as the direct effect estimated in Chapter III (see Table 16).

- The predicted default probability for two-year schools compared to four-year schools increases 1.3 percentage points when indirect effects were included, or half again as much as the direct effect estimated in Chapter III (see Table 16).

In other words, incorporating the effect of school type on earnings and, therefore, indirectly on default, widened the gap between predicted default rates associated with attendance at two-year and proprietary schools versus four-year schools by 50 percent.

These calculations may overstate the effects of school type on default if the decision to attend a proprietary or two-year school is influenced by unobserved characteristics that also affect future earnings. If this is the case, the regression model will overestimate the earnings *increase* of attending a four-year school, for proprietary and two-year school students, and will overestimate the earnings *reduction* of attending a proprietary or two-year school, for four-year school students. For this reason, impacts calculated using this alternative approach are best viewed as upper bound estimates of the true school type impacts.