DOCUMENT RESUME

ED 445 635 HE 033 308

AUTHOR	Yang, Faxian
TITLE	Using Survival Analysis To Analyze and Predict Students' Achievement from Their Status of Developmental Study.
PUB DATE	2000-05-00
NOTE	26p.; Paper presented at the Annual Meeting of the Association for International Research (40th, Cincinnati, OH, May 21-24, 2000).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Achievement; College Students; Community Colleges; *Developmental Programs; Educational Development; Educational Research; Grade Point Average; Higher Education; Predictor Variables; Racial Differences; Research Methodology; Student Characteristics; *Student Development; Two Year Colleges
IDENTIFIERS	*AIR Forum; *Survival Analysis

ABSTRACT

This study predicted student achievement based on need for developmental study, investigating different students' types of censoring; variables influencing student achievement; and whether developmental study was appropriate for realizing student achievement. Participants were 403 community college students. Achievers were students who within four years were awarded a degree/certificate; transferred to another public state college; or were sophomores with at least a 2.0 grade point average (GPA). Data came from the college's student system and its institutional research database. Constant variables were gender, ethnicity, age, and need for taking a developmental course. Time-dependent variables were number of credit hours and GPA. Based on developmental study needs/enrollment from 1992-96, students were divided into four categories: needed no developmental study; needed one developmental course; needed two developmental courses; or needed more than two courses. Results indicated that the fewer developmental courses students took, the higher the probability for success. White students had higher probability of achievement than minority students. Students with higher GPAs were more successful than students with lower GPAs. Gender and age did not predict achievement. Achievement significantly related to GPA in the 5th, 6th, and 7th semesters. Ten figures and 21 data tables are included. (Contains 19 references.) (SM)

Running head: ANALYZE AND PREDICT STUDENTS' ACHIEVEMENT

Using Survival Analysis to Analyze and Predict Students' Achievement from Their Status of Developmental Study

Faxian Yang

Research Analyst

Harford Community College
401 Thomas Run Road
Bel Air, Maryland 21015
(410) 836-4302
fyang@harford.cc.md.us

Paper presented at the Association for Institutional Research 40th Annual Forum, May 2000, Cincinnati, Ohio

Using Survival Analysis to Analyze and Predict Students' Achievement from Their Status of Development Study

Abstract

Students' achievement in community colleges has been an important issue in terms of retention and success of ...nan investment for legislatures, administrators, faculty, students, and the students' parents. Although many ...ariables may influence students' achievement, this study focuses on the effects from students' needs for developmental study. This study's purpose is to analyze and predict students' achievements from their status of developmental study. The significance of this study is to find a cutoff point of students' achievement and variables, which influenced students' achievement. Survival analysis is used in the study. The study's findings indicate that key elements in determining students' achievement are their persistence and GPA in 5th, 6th, and 7th semesters out of a total of 12 semesters.

Using Survival Analysis to Analyze and Predict Students' Achievement from Their Status of Developmental Study

Introduction

Community colleges' open-door/open-education policy offers everyone a chance to study at a college; this policy is one of the greatest strengths of the community colleges. To be able to study at a college is a dream of many people who are living in other countries. However, the open-door policy brings different kinds of students to community colleges based on their needs for developmental courses. College-prepared students do not need developmental education. On the contrary, many students are under prepared for college life from both psychological and academic aspects. Some students may take more than two developmental courses (Ignash, 1997). To help the under prepared students, community colleges have the responsibility to respond to their students' needs, since the students have the potential to be successful in college-level coursework if deficiencies are remedied (McMillan, Parke & Lanning, 1997).

Students' achievements in community colleges have been an important issue in terms of student retention and the success of human capital investment for legislatures, administrators, faculty, students, and the students' parents. Many variables—such as the needs for developmental study, student attendance status, student financial status, and the type of students based on whether s/he is a traditional learner, etc.—may influence students' achievement. This study is interested in the variable, the needs for developmental study. It tries to find the relationship between students' achievement and students' initial developmental study at a community college by using survival analysis.

Purpose of Study

The purpose of this study is to analyze and predict differential students' achievement based on the students' needs for developmental study.

Research Question

What *censoring* do different kinds of students have?

What variables do influence students' achievement?

Is developmental study a proper way to realize students' achievement?

The Definition of Student Achievement

As long as students belonged to one of the following groups, they are defined as achievers in the study:

1. Degree/Certificate awarded.
2. Transfer to another public college in the state.
3. Sophomore with a GPA of 2.0 or above.

If a student meets more than one criteria from list above, s/he is classified in a higher hierarchy.

Population and Data Source of Study

The population of the study includes 403 students at a medium-size community college in the northeast. The students were first-time, full-time, degree-seeking students entering the college in fall of 1992.

Although it is assumed that the students finish their associate degree within two years at the community college, many students needed a much longer time to be awarded than they should because of different reasons. However, as long as students achieved one of the statuses listed above within four years, they are counted as achievers in the study. Therefore, the ending study time for cohort of fall 1992 was summer 1996.

The data used for this study came from the college's Banner Student System and the database in the Office of Institutional Research of the college.

Limitations of Study

If students transfer to any college outside the state where the college is located, or any private college in the state, the college does not have their records. Therefore, the data source for the study is considered to be incomplete. However, the data are the only source available for this study.

The Significance of Study

This study tracks students' academic behavior and achievement for four years instead of one course and compares the behavior and achievement between developmental students and non-developmental students at a community college. Because this study employs discrete-time survival analysis, it includes all time-dependent variables in the model. Stepwise logistic regression can find most influential variables for students' achievement from the model. This approach is more quantitative than qualitative. Because same kind of students may have similar censoring and/or share common influential variables at different colleges, the results of this study may be meaningful to other colleges which have a developmental program. Other colleges can use the approach and methods of this study for their research to help their students to become more successful.

Review of Literature

Developmental Study

Weissman, Bulakowski, and Jumisko (1997) believed that

> The mission of the community college is directly linked to providing access to all students who can benefit and to enhancing opportunities for students to accomplish their academic and career goals (p.76).

Developmental study has been a part of community colleges' mission for a long time. Because of the open-door policy, it will continue to be an essential component of the offerings (Weissman, et al., 1997). Based on a National Center for Educational Statistics (NCES) report in 1996, all community colleges surveyed offered developmental courses, 40% of freshmen did not have adequate preparation in one of three basic skill areas of reading, writing, and mathematics. Although many students do not prepare to engage in college-level coursework, colleges still have responsibility to respond to students' needs as long as the students have the potential to be successful in college-level coursework (McMillan, Parke, & Lanning, 1997).

Some educational researchers compared grades of some initial college-level courses from developmental students to the grades of non-developmental students. Their findings and conclusions were different. Based on their study, Hopper, Taylor, and Wolford (1997) found that there was significant difference between developmental and non-developmental students to pass initial English composition at college-level. Non-developmental students had more chance to pass the initial English composition than the developmental students. However, Eanes (1992), Weissman, et al. (1997) and Lyons (1990) thought that after completing developmental courses, developmental students performed as well as or better than college-ready students. Based on their studies, no significant differences were found between developmental students and non-developmental students for their final course grades.

Survival Analysis in Study of Education

"Survival analysis is a class of statistical methods for studying the occurrence and timing of events" (Allison, 1995, p. 1). It was initially designed for biomedical science research. Today, the applications of the method go beyond biomedical research. A general purpose of survival analysis is to estimate causal or predictive models in which the risk of an event depends on covariates (Allison, 1995). Some educational researchers had employed the method to study students' or faculty's retention rates, dropout rates, degree attainment, and

employment status of handicap students for years (Sharron Ronco, 1994; Randall Schumacker and Kathleen Denson, 1994; Judith Singer and John Willett, 1993; Mike Tammada and Claudia Inman, 1997; Beiling Xiao, 1997, 1998). Singer and Willett are pioneers in applying discrete-time survival analysis into educational study.

Cox's regression model has promoted survival analysis and has been a major method employed in survival analysis since middle of 1970's. However, Schumacker and Denson (1994) believed that there are some important limitations in Cox's regression model. The basic assumption in Cox's model is the first and most significant limitation, which cancel the interaction of the variables with a time variable not in the equation. The lack of a term to represent unobserved heterogeneity in the model is another major limitation, which has been found to be especially significant when dealing with repeated events. As Singer and Willett stated in 1993, time itself is the fundamental time varying predictor, and it should not be left out in the model.

Variables may change as time changes. Survival analysis with time variables becomes very necessary in some studies. Discrete-time survival analysis meets the requirement of the analysis. It requires that one person has more than one observation, which is a major difference between standard survival analysis and discrete-time survival analysis. To meet the requirement, one-person, one-record data set (person data set) should be converted into one person, multiple-period data set (person-period data set) (Schumacker et al. 1994). If event times are truly discrete and measured coarsely, logistic regression is a preferable method to deal with the analysis (Allison, 1995).

Methodology

The Variables and Data Set

Two kinds of variables are included in the study: constant variables and time-dependent variables. Constant variables do not change as time changes, which include students' gender, ethnicity, age, and the needs for taking developmental course. Time-dependent variables include the students' number of credit hours and GPA, enrolled indicator and achievement indicator for each semester.

The 1992 cohort included 403 first-time, full-time, degree-seeking students. Of these, 176 were male and 227 female, or 355 white and 48 ethnic minority students. In the same cohort, 360 students' ages were 16 to 20; 26 students' ages were 21 to 30; and 17 students' ages were older than 30.

Converting Data from Student to Student-Semester

Because no student enrolled in all 12 semesters from fall 1992 to summer 1996, new variables were

required to identify the semester(s) in which s/he had enrolled. Because each student's GPA and credit hours in each semester were different, each student needed more than one observation to show his or her records as long as he or she enrolled more than one semester. Therefore, the records of each student must be converted from student to student-semester. The number of observations of each student depended on the number of semesters in which the student enrolled. Although the first semester for all students was fall 1992, the rest of the semester(s) may vary for different students. For example, the second semester may be spring 1993 for student A. It may also be summer 1996 for student B if he or she dropped the college from spring 1993 and return to the college in summer 1996. (see Tables 1, 2). This conversion is the first and an important step of analysis for *time-dependent covariates.*

Table 1: Original Student Data (Demo)

Student	Semester	Gender	Ethnicity	Age	Needs	GPA	Hour	Good
AAA	4	0	W	21	0	3.00	12	1
BBB	1	0	W	18	1	2.56	11	0
CCC	5	1	W	31	2	2.22	9	0
DDD	7	0	M	25	0	3.13	11	1
EEE	4	1	W	16	>2	2.09	9	0
FFF	10	0	M	22	>2	2.11	10	1

Note:
Semester: The last semester in which he or she enrolled.
Gender: 1 = Male, 0 = Female.
Ethnicity: W = White, M = Minority.
Needs: The needs for developmental course.
GPA: The GPA earned in his or her last semester.
Hour: The credit hours that he or she earned in last semester.
Good: Status of a student's achievement in his or her last semester;
1 = Achievement, 0 = Not Achievement.

The Definition: the Need for Developmental Study

Based on the needs/enrollment for developmental study from fall 1992 to summer 1996, the students are divided into four categories:

1. Did not need any developmental study.
2. Only needed one developmental course.
3. Needed two developmental courses.
4. Needed more than two developmental courses.

These four categories are the foundation of analysis of this study. According to the four categories of students and the definition of achievement, the researcher determines if the categories are reasonable indicators for predicting students' achievement and which category is the most successful.

Table 2: Student-Semester Data Set (Demo)

Student	Semester Indicator												Semester	Gender	Needs	GPA	Good
	1	2	3	4	5	6	7	8	9	10	11	12					
AAA	1	0	0	0	0	0	0	0	0	0	0	0	4	0	0	3.02	0
AAA	0	1	0	0	0	0	0	0	0	0	0	0	4	0	0	2.97	0
AAA	0	0	1	0	0	0	0	0	0	0	0	0	4	0	0	3.25	0
AAA	0	0	0	1	0	0	0	0	0	0	0	0	4	0	0	3.00	1
BBB	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	2.56	0
FFF	1	0	0	0	0	0	0	0	0	0	0	0	12	0	>2	1.80	0
FFF	0	1	0	0	0	0	0	0	0	0	0	0	12	0	>2	2.05	0
FFF	0	0	1	0	0	0	0	0	0	0	0	0	12	0	>2	2.20	0
FFF	0	0	0	1	0	0	0	0	0	0	0	0	12	0	>2	2.20	0
FFF	0	0	0	0	1	0	0	0	0	0	0	0	12	0	>2	2.10	0
FFF	0	0	0	0	0	1	0	0	0	0	0	0	12	0	>2	1.95	0
FFF	0	0	0	0	0	0	1	0	0	0	0	0	12	0	>2	2.00	0
FFF	0	0	0	0	0	0	0	1	0	0	0	0	12	0	>2	2.25	0
FFF	0	0	0	0	0	0	0	0	1	0	0	0	12	0	>2	2.50	0
FFF	0	0	0	0	0	0	0	0	0	1	0	0	12	0	>2	2.00	0
FFF	0	0	0	0	0	0	0	0	0	0	1	0	12	0	>2	2.50	0
FFF	0	0	0	0	0	0	0	0	0	0	0	1	12	0	>2	2.11	1

Note:

Semester Indicator: Indicate which semester(s) enrolled by a student.
Semester: The last semester which he or she enrolled.
Gender: 1 = Male, 0 = Female.
Needs: The needs for developmental course.
GPA: The GPA of a student in each semester.
Good: The semester in which the student achieved; 1 = Achievement, 0 = Not Achievement.

Survival Analysis

Survival analysis is used to find the time-dependent variables related to students' achievement. If a student is a non-achiever in this study or dropout, he or she is censored. Otherwise, he or she is not censored. The researcher expects to know which semester(s) played a more important role in students' achievement, and what variables caused the change to happen.

Non-parametric method. "The Kaplan-Meier Method (KM) estimator is a non-parametric method. It is the most widely used method for estimating survivor functions" (Allison, 1995, p.30). It is useful to compare two or more different treatment groups. The KM estimator is defined as

$$\hat{S}(t) = \prod_{j:t_j \le t} [1 - \frac{d_j}{n_j}] \text{ for } t_1 \le t \le t_k.$$

Note:

n_j : all students at time t_j.
d_j : the number of students who are achievers at time t_j.

The KM is used to find censoring, as well as survival and hazard distribution in this study (see Table 3). $\frac{d_j}{n_j}$ is column K, achieving rate. $\prod_{j:t_j \le t}[1-\frac{d_j}{n_j}]$ is column I, non achiever. Column I equals to non achiever rate in previous row multiplied by non achieving rate (1 – K) in this row. Column C is the number of students at the beginning of each semester. Column D to F or Column G are students not censored in the study. Column H is students censored in the study.

Table 3: Application of the Kaplan-Meier Method

A	B	C	D	E	F	G	H	I	J	K
	Semester	Head Count	Achievement				Dropped Out	Proportion of		
			Degree	Transfer	Sophomore	Sum		Non Achiever	Achiever	Achieving
1	Fall 1992	403	0	0	0	0	47	1*(1-K1)	1 – I1	G1/C1
2	Spring 1993	356	0	4	1	5	41	I1*(1-K2)	1 – I2	G2/C2
3	Summer 1993	310	0	0	0	0	1	I2*(1-K3)	1 – I3	G3/C3
4	Fall 1993	309	0	0	5	5	30	I3*(1-K4)	1 – I4	G4/C4
5	Spring 1994	274	18	0	13	31	37	I4*(1-K5)	1 – I5	G5/C5
6	Summer 1994	206	8	0	3	11	10	I5*(1-K6)	1 – I6	G6/C6
7	Fall 1994	185	20	1	2	23	14	I6*(1-K7)	1 – I7	G7/C7
8	Spring 1995	148	20	0	4	24	20	I7*(1-K8)	1 – I8	G8/C8
9	Summer 1995	104	11	0	2	13	4	I8*(1-K9)	1 – I9	G9/C9
10	Fall 1995	87	6	1	4	11	18	I9*(1-K10)	1 – I10	G10/C10
11	Spring 1996	58	12	0	8	20	30	I10*(1-K11)	1 – I11	G11/C11
12	Summer 1996	8	1	0	2	3	5	I10*(1-K12)	1 – I12	G12/C12
			96	6	44	146	257			

Parametric method. Logistic regression is a form of statistical modeling that is often appropriate for categorical outcome variables. "It describes the relationship between a categorical response variable and a set of explanatory variables" (Stokes, Davis, and Koch 1995, p165). The logistic regression is defined as

$$\log(\frac{\pi(x)}{1-\pi(x)}) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

$$\pi(x) = \frac{1}{1 + e^{-(\alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x)}}$$

Note:
$\pi(x)$: the conditional probability that individual *i* has an event at time *t*, given that an event does not already occur to that individual.

The model above indicates that $\pi(x)$ is related to the covariates by a logistic regression equation. This model is most appropriate when events can only occur at regular, discrete points in time (Allison, 1995). Because each student was counted maximally four academic years (or 12 semesters) since he or she entered the college, the semesters are counted as discrete points in time. This study's model includes the following variables:

GOOD,	GENDER,
TERM1-TERM10,	RACE,
GPA1-GPA10,	YOUNG,
HOUR1-HOUR10,	IND01.

GOOD is the dependent variable in this study. GOOD =1 means that a student achieved in the semester. GOOD = 0 indicates students not achieving. Because the maximum semesters enrolled by students in the cohort were 10, maximum time-dependent variables are 10. *TERM1 - TERM10* name the 10 semester indicators. If a student enrolled in a semester, the indicator for him or her is 1. Otherwise, it is 0. *GPA1-GPA10* and *HOUR1-HOUR10* are real GPA and credit hours for the students in each semester. Both *GENDER* and *RACE* have two categories. Male and white are 1. Female and minority are 0. *YOUNG* is name of student's age in model. It is divided into three categories, "1" for ages 16 to 20, "2" for ages 21 to 30, and "3" for ages above 30. The categories of *IND01* is the same as the definition for the needs for developmental courses above.

Analysis of Data

Descriptive Analysis

Basic information. More than 41% (166 of 403) of students did not need to take a developmental course in the cohort of fall 1992. About 24% (98 of 403) needed one, 17% (67of 403) needed two, and 18% (72 of 403) required more than two developmental courses (see Table 4). In other words, more than 58% (237 of 403) of students took at least one developmental course in the cohort. The students totally took 509 developmental courses in four academic years. Each student who had the course equally took 2.15 developmental courses.

About 63% (143 of 227) of female students and 53% (94 of 176) of male students took at least one developmental course (see Table 4). About 55% (195 of 355) of white students and 97% (29 of 30) of African American students had at least one developmental course (see Table 5). About 79% (34 of 43) of students whose

age were older than 20 and 56% (203 of 360) of students with ages between 16 and 20 took the courses (see Table 6). In general, more than 58% (237 of 403) of students in the cohort at the college were not ready for college education when they first entered the college.

Table 4: Gender and the Needs for Developmental Study

Group	Gender		Total
	Male	Female	
No Needs	82	84	166
Need One	38	60	98
Need Two	28	39	67
More Than Two	28	44	72
	176	227	403

Table 5: Ethnicity and the Needs for Developmental Study

Group	Ethnicity						Total
	Black	Indian	Asian	Hispanic	White	Unknown	
No Needs	1	0	2	1	160	2	166
Need One	7	1	0	1	86	3	98
Need Two	10	0	2	0	55	0	67
More Than Two	12	0	4	1	54	1	72
	30	1	8	3	355	6	403

Table 6: Age and the Needs for Developmental Study

Group	Age			Total
	16 - 20	21 -30	>= 31	
No Needs	157	7	2	166
Need One	88	7	3	98
Need Two	54	6	7	67
More Than Two	61	6	5	72
	360	26	17	403

Achievement. Achievement analysis is based on the needs for developmental course. For students who did not require developmental courses, 28.31% (47 of 166) of them were awarded an associate degree or a certificate; 3.01% (5 of 166) transferred to other public colleges in the state where the college is located; and 12.05% (20 of 166) were sophomores with 2.0 or higher GPA (see Tables 7, 8, 9). In general, 43.37% (72 of 166) of students achieved in four academic years. This group was the most successful group in all four groups.

As students in the cohort took only one developmental course, 22.45% (22 of 98) were awarded an associate degree or a certificate; 1.02% (1 of 98) transferred to other public colleges in the state; and 15.31% (15 of 98) were sophomores with 2.0 or higher GPA (see Tables 7, 8, 9). On the whole, 38.78% (38 of 98) of students achieved in four academic years.

If students took two developmental courses, their achievement rate was lower than students in the previous two groups. Degree-award rate was 20.89% (14 of 67); none of them transferred; 8.96% (6 of 67) were sophomore with 2.0 or higher GPA (see Tables 7, 8, 9). Overall, 29.85% (20 of 67) of students achieved in four academic years.

When students took more than two developmental courses, their achievement rate was the lowest in the four groups. Only a mere 13 out of 72 students (18.06%) were awarded an associate degree or a certificate; none of them transferred; 4.17% (3 of 72) of them were sophomore with 2.0 or higher GPA (see Tables 7, 8, 9). Taken as a whole, 22.22% (16 of 72) of students achieved in four academic years.

Table 7: The Needs for Developmental Study and Semester of Graduation

Needs	Semester of Graduation								Total Awarded	Total
	Spring 1994	Summer 1994	Fall 1994	Spring 1995	Summer 1995	Fall 1995	Spring 1996	Summer 1996		
None	13	6	10	9	4	2	3	0	47	166
One	5	2	5	0	6	0	4	0	22	98
Two	0	0	3	6	1	0	3	1	14	67
More	0	0	2	5	0	4	2	0	13	72
Total	18	8	20	20	11	6	12	1	96	403

Table 8: The Needs for Developmental Study and Semester of Transfer

Needs	Semester of Transfer			Total Transfer	Total
	Spring 1993	Fall 1994	Fall 1995		
None	4	1	0	5	166
One	0	0	1	1	98
Total	4	1	1	6	264

Table 9: The Needs for Developmental Study and Sophomore with 2.0 or Higher GPA

Needs	Last Semester Enrolled										Total Sopho -more	Total
	Summer 1993	Fall 1993	Spring 1994	Summer 1994	Fall 1994	Spring 1995	Summer 1995	Fall 1995	Spring 1996	Summer 1996		
None	1	1	7	3	1	2	0	1	4	0	20	166
One	0	3	4	0	1	0	2	2	3	0	15	98
Two	0	1	2	0	0	1	0	0	1	1	6	67
More	0	0	0	0	0	1	0	1	0	1	3	72
Total	1	5	13	3	2	4	2	4	8	2	44	403

Based on the analysis presented, it is clear that the fewer developmental courses a student takes, the more successful a students becomes. In other words, students' academic background in high school strongly influenced students' achievement at the college (see Table 10). The needs for developmental courses not only influenced students' achievement, but also impacted students' grades at the college. About 24% (40 of 166) of students earned

overall "A" or "B" in the group, in which students did not take developmental course; 7% (7 of 98) earned "B" in the group that needed one, 10% (7 of 67) in the group which required two, and none of them (0 of 72) in the group that need more than two. About 26% (44 of 166) of students' overall GPA, in the group that did not take developmental course, were "D", 26% (26 of 98) in the group that needed one developmental course, 31% (21 of 67) in the group which required two courses, and 46% (33 of 72) in the group that needed more than two courses (see Table 11).

Table 10: Summary of the Number of Achievers by the Needs for Developmental Courses

Group	Needs for Developmental Course	Total	Achieved	Non-Achieved	% Non-Achieved
1	No Needs	166	72	94	56.63
2	Need One	98	38	60	61.22
3	Need Two	67	20	47	70.15
4	More Than Two	72	16	56	77.78
		403	146	257	63.77

Table 11: Summary of Overall GPA by the Needs for Developmental Courses

Group	Needs for Developmental Course	Overall GPA					Total
		A	B	C	D	F	
		4.000	3.000 - 3.999	2.000 - 2.999	1.000 - 1.999	0.000 - 0.999	
1	No Needs	1	39	64	44	18	166
2	Need One	0	7	40	26	25	98
3	Need Two	0	7	18	21	21	67
4	More Than Two	0	0	25	33	14	72
		1	53	147	124	78	403

Survival Analysis

The Kaplan-Meier method. One of best features of the Kaplan-Meier method is plotting by strata. By estimating survival curve, plotting can summarize the patterns of response, and survival curves can be visually compared (Greenhouse, Stangl, Bromberg, 1989). Survival distribution function in following figures described the probability of retention in each semester. In contrast, Hazard function in the figures portrayed the probability of achievers in each semester.

The null hypothesis is a way to test difference among/between different strata by their nature. The needs for developmental study, gender, ethnicity, age, and overall GPA are tested by the method discussed here (see Table 12). For example, null hypothesis states that the rate of achievement for needs for developmental courses is the same for all groups, at $\alpha = 0.05$ level. However, based on Mantel-Cox statistic (log-rank), χ (df = 2) = 23.0078, *'probability of > Chi-Square'* < 0.0001. Therefore, the hypothesis is rejected.

Based on the analysis presented here, differential needs for developmental course, ethnicity, and overall GPA made a difference for students' achievement. However, students' gender and their age did not make any distinction (also see Figures 1 to 10).

Table 12: Log Rank Test

Log Rank Test	Chi-Square	Degree Freedom	Pr > Chi-Square	Result
The need for developmental study	23.0078	2	<.0001	Reject
Gender `	0.2018	1	0.6533	Non Reject
Ethnicity	5.3313	1	0.0209	Reject
Age	1.4680	2	0.4800	Non Reject
Overall GPA	119.2418	3	<.0001	Reject

Quartile estimates of 50 % of students in the group with no developmental-study needs enrolled eight semesters at the college, 10 semesters in the group with one or two developmental courses, and 12 semesters in the group with more than two developmental courses. Overall, quartile estimates of 50 % of students enrolled 10 semesters in the cohort (see Table 13).

Table 13: Summary Statistics for Time Variable *Semester*, by Needs for Developmental Courses

Group	Mean	Standard Error	Quartile Estimates		
			25%	50%	75%
Total	9.2056	0.1707	7.0000	10.0000	12.0000
No Needs	8.1213	0.2496	6.0000	8.0000	11.0000
1-2 Needs	9.2850	0.2609	7.0000	10.0000	11.0000
More than Two	10.8341	0.2633	10.0000	12.0000	.

Students who dropped out are counted as non-achievers and censored, which happened after each semester and before the next semester began. The numbers in the column of achievers (see Tables 14, 15, 16, 17) indicate the cumulative probability for each semester range that students achieved up to that semester. The column of achievers plus the column of non-achievers in each row always equals one. For example, as all students in the cohort, 41 students became achievers up to spring 1994, which was counted 13.97% cumulative probability for achievers. Among the students who did not take any developmental course, 26 became achievers—consisting of 20.74% in the group, 15 in the group which took one or two the course (s), comprising 13.44% in the group. No students became achievers up to spring 1994 in the group which took more two the courses (see Tables 14, 15, 16, 17).

The students who enrolled in more than two developmental courses required the longest time to be awarded degree/certificate of all the groups. In contrast, if students did not need developmental coursework at all, they required the shortest time to be awarded a degree/certificate (see Figures 1 and 2).

Figure 1: The KM Method - the Needs for Developmental Course, Survival Distribution

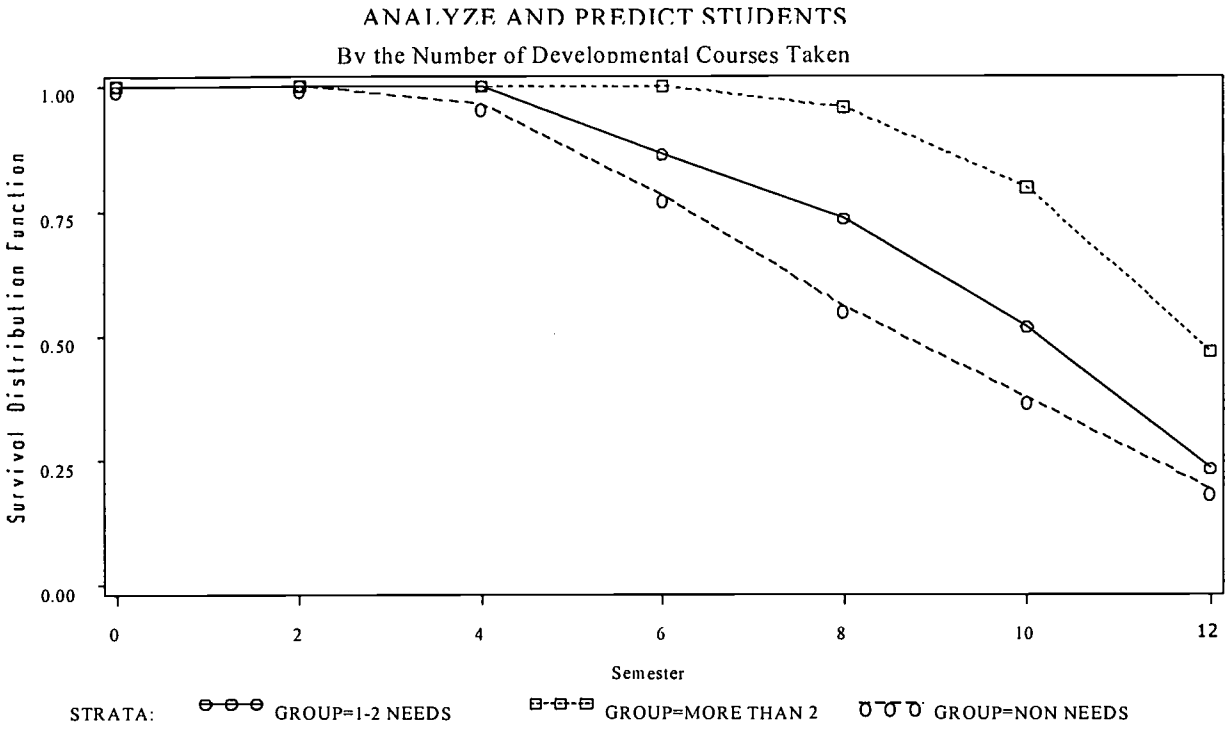

Figure 2: The KM Method - the Needs for Developmental Course, Hazard Distribution

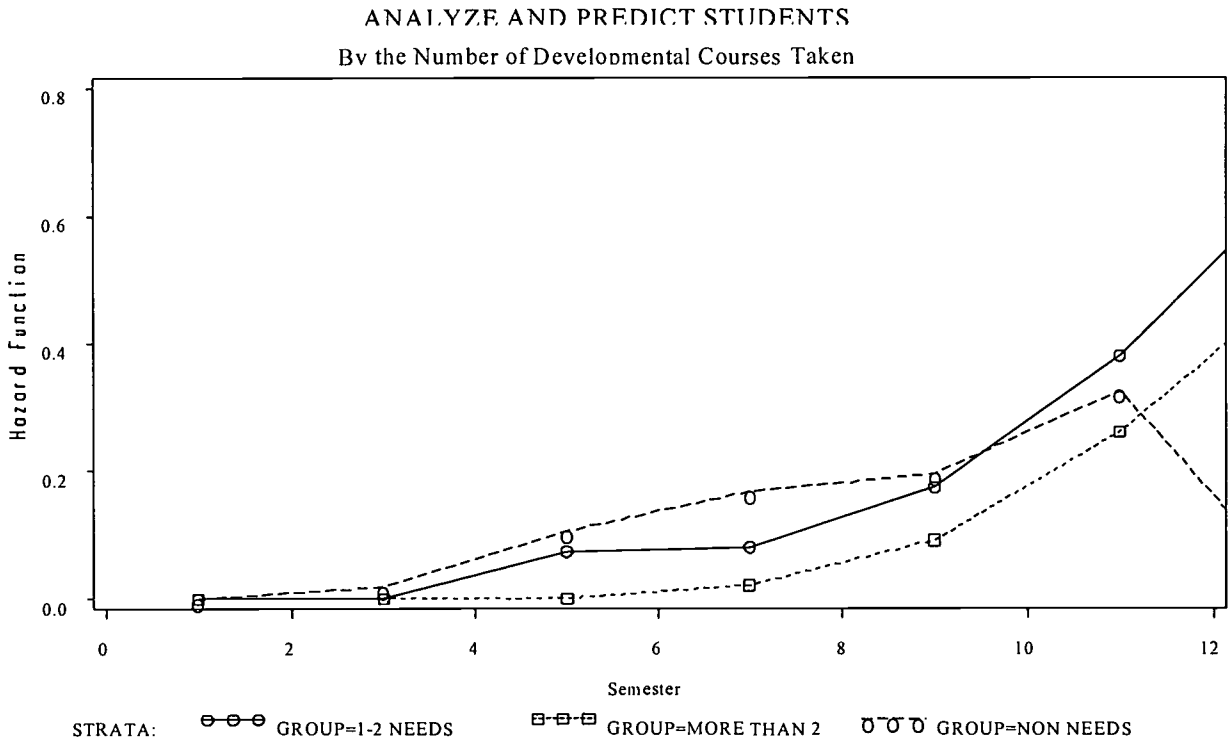

Figure 3: The KM Method - Gender, Survival Distribution

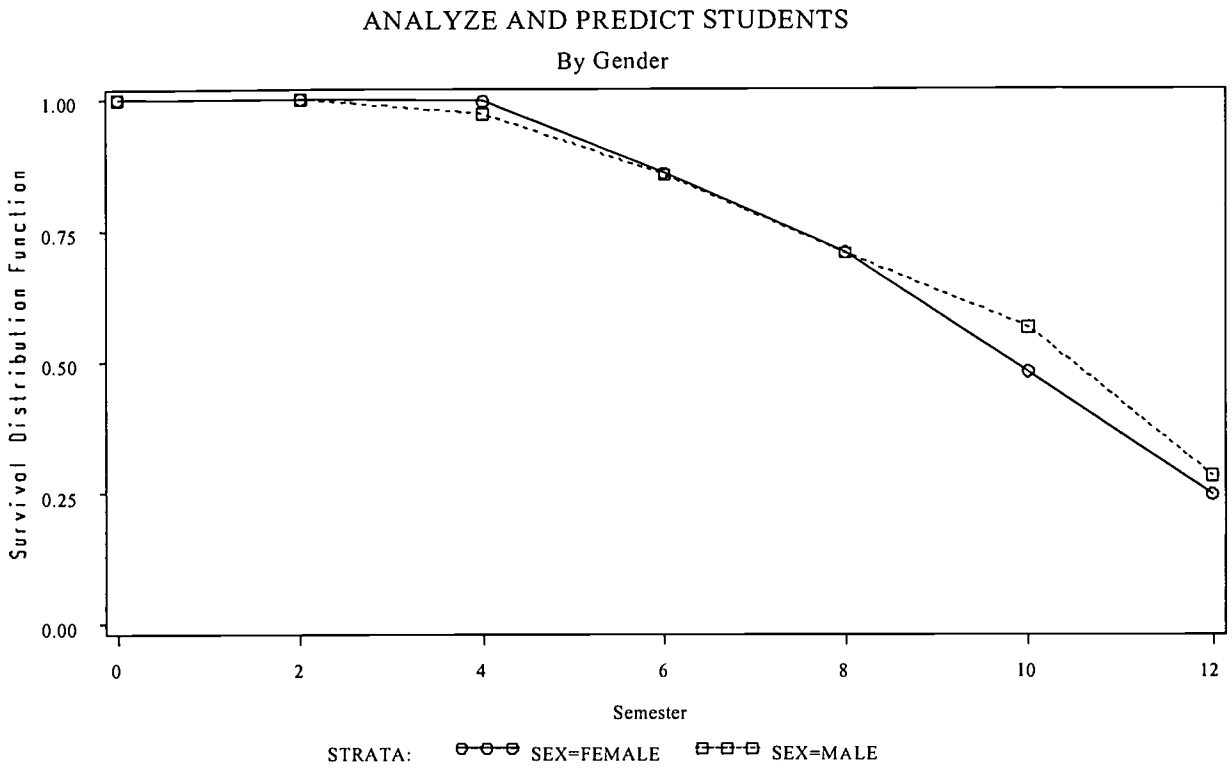

Figure 4: The KM Method – Gender, Hazard Distribution

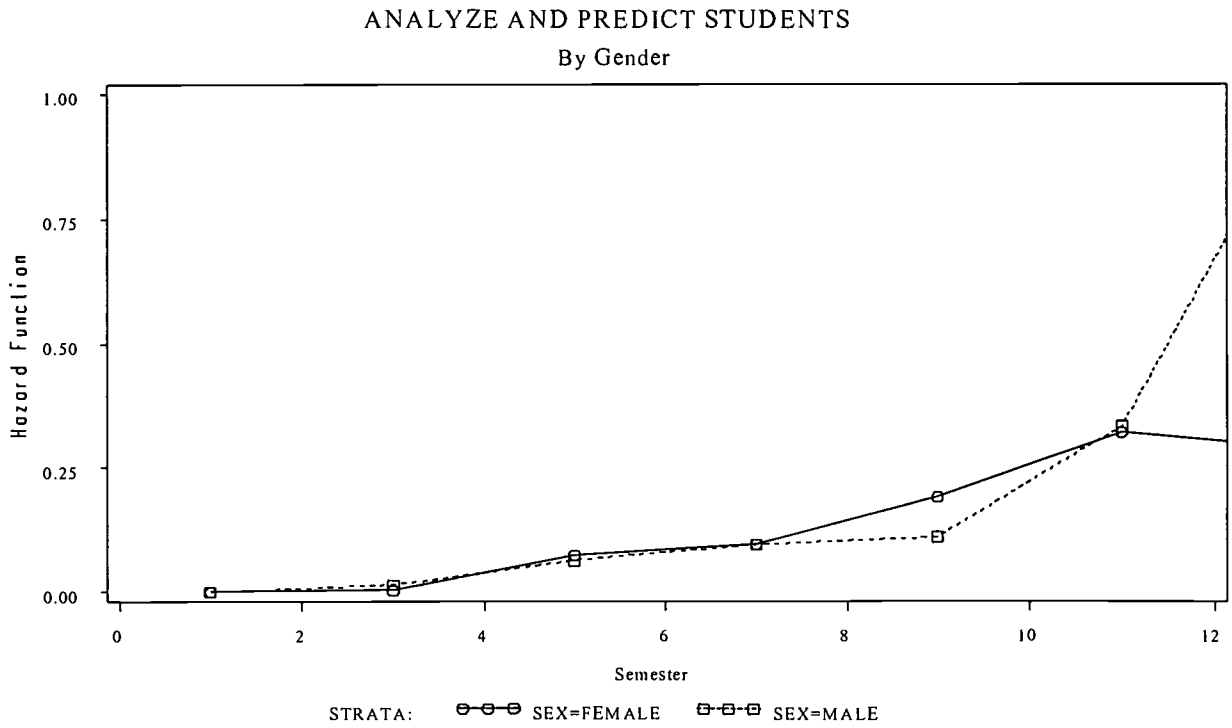

Figure 5: The KM Method - Ethnicity, Survival Distribution

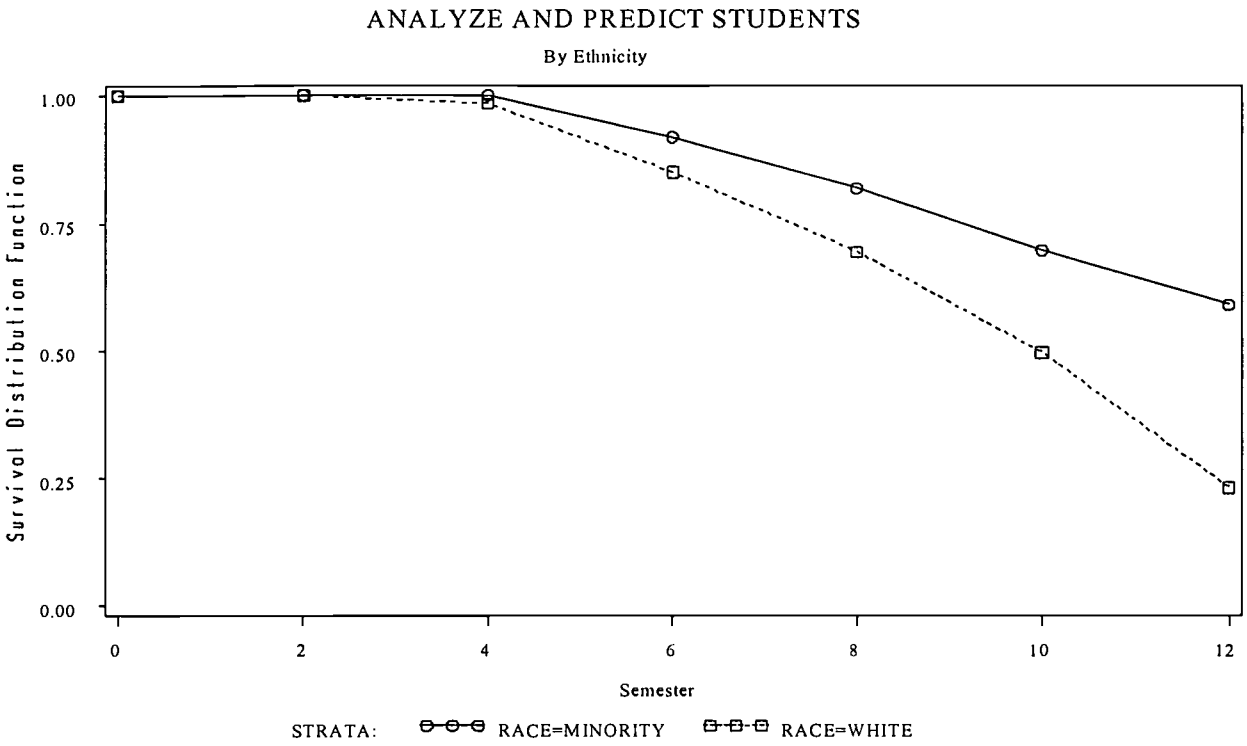

Figure 6: The KM Method - Ethnicity, Hazard Distribution

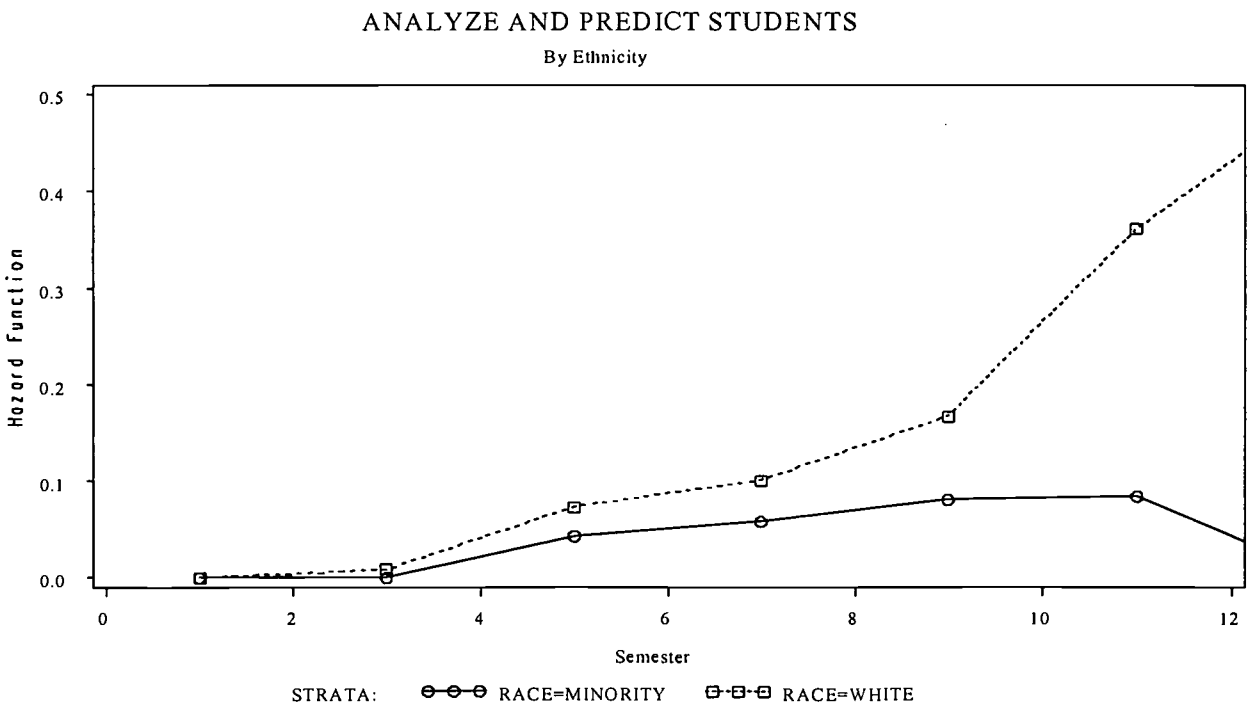

Figure 7: The KM Method - Age, Survival Distribution

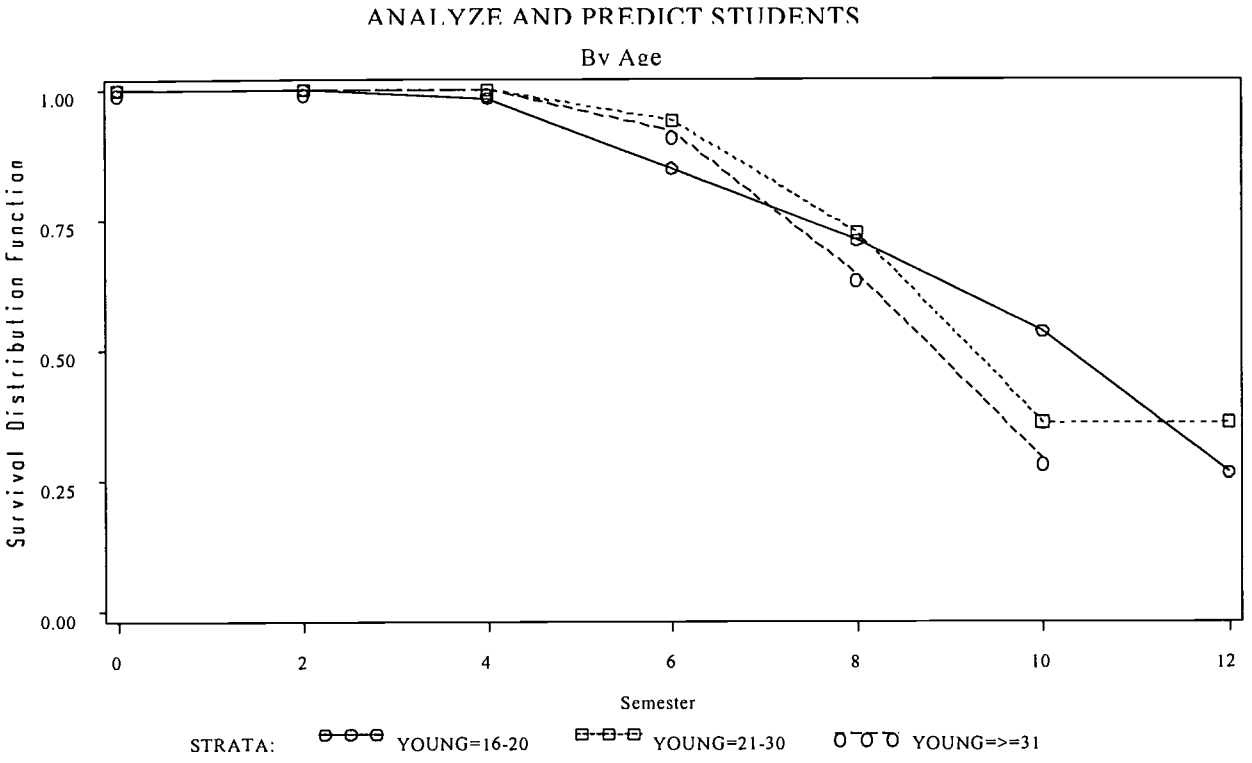

Figure 8: The KM Method - Age, Hazard Distribution

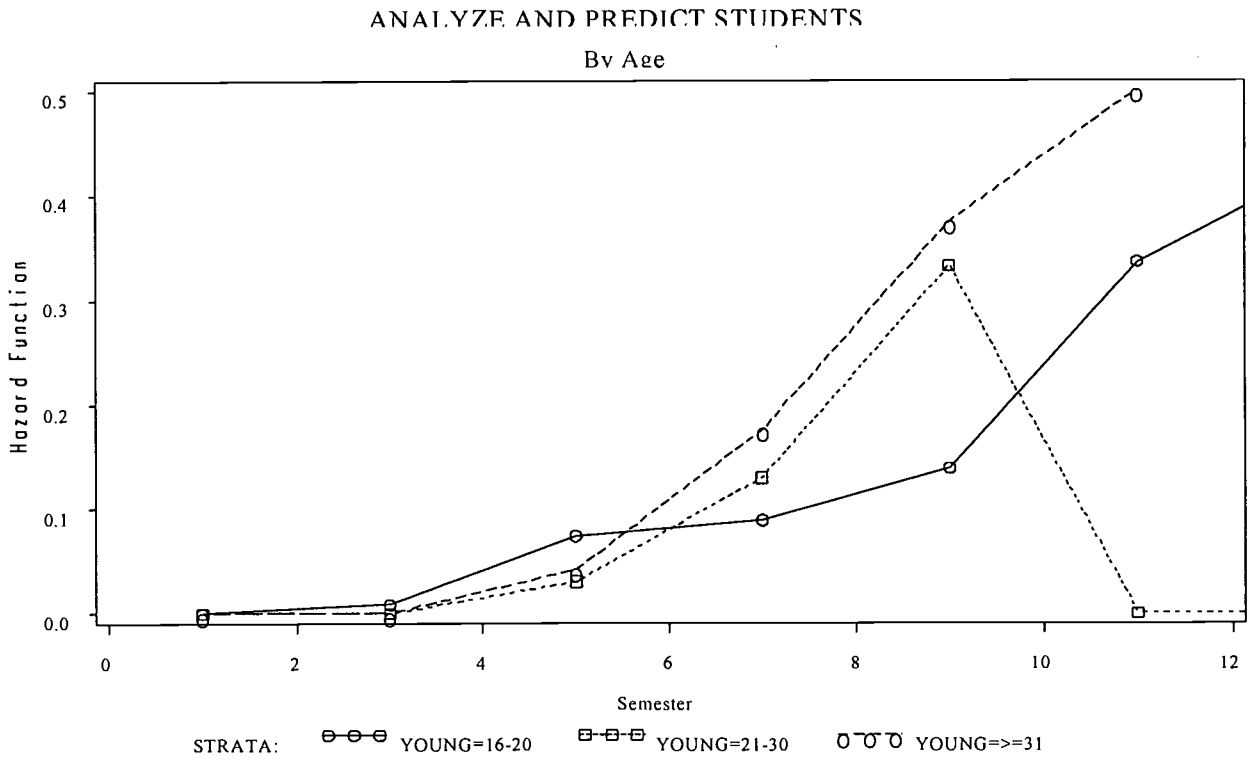

Figure 9: The KM Method - Grade, Survival Distribution

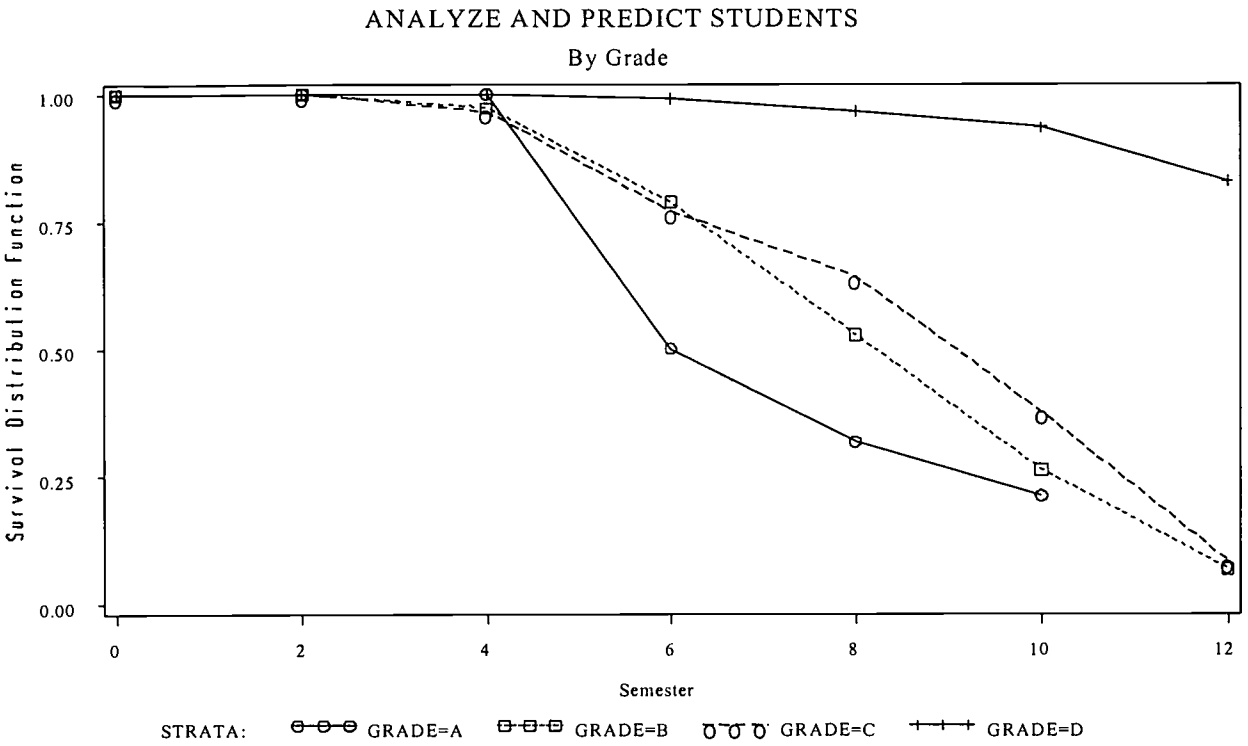

Figure 10: The KM Method - Grade, Hazard Distribution

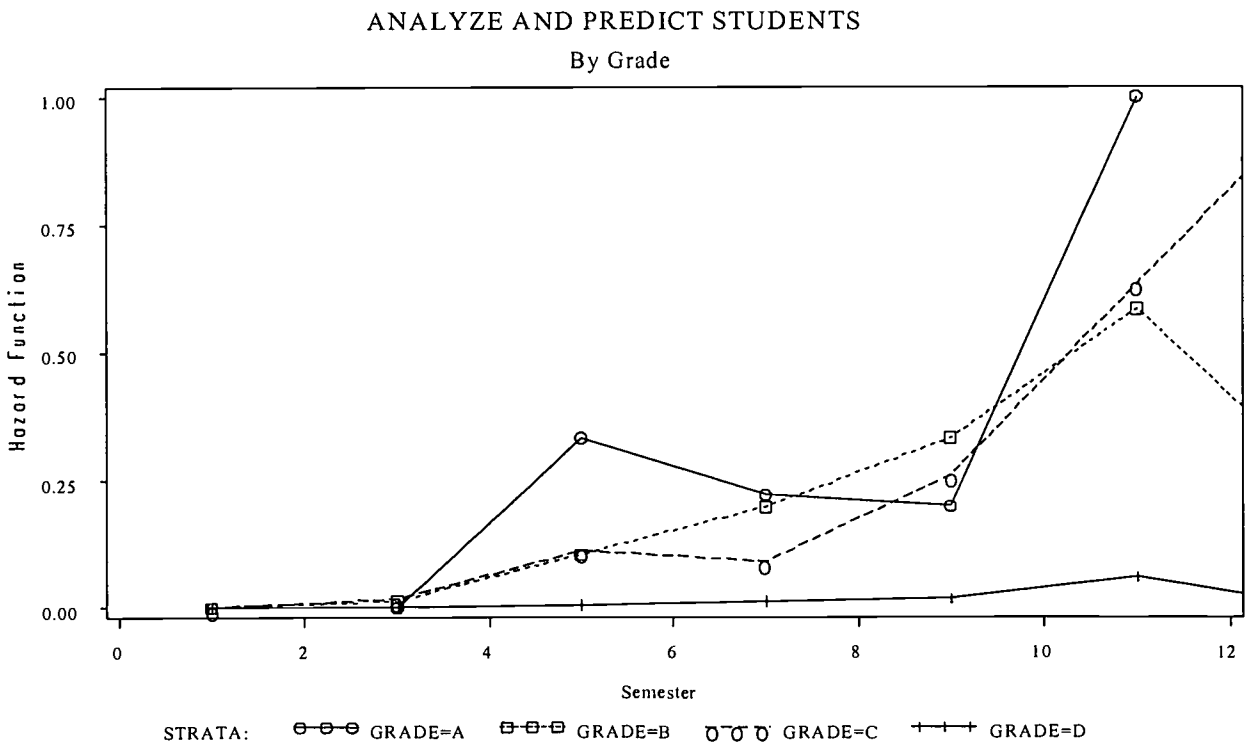

Table 14: Achievement Rate by the Needs for Developmental Courses, Total

Semester	Head Count	Achievement				Dropped Out	Proportion of		
		Degree	Transfer	Sophomore	Sum		Non Achiever	Achiever	Achieving
Fall 1992	403	0	0	0	0	47	100.00%	0.00%	0.00%
Spring 1993	356	0	4	1	5	41	98.60%	1.40%	1.40%
Summer 1993	310	0	0	0	0	1	98.60%	1.40%	0.00%
Fall 1993	309	0	0	5	5	30	97.00%	3.00%	1.62%
Spring 1994	274	18	0	13	31	37	86.03%	13.97%	11.31%
Summer 1994	206	8	0	3	11	10	81.43%	18.57%	5.34%
Fall 1994	185	20	1	2	23	14	71.31%	28.69%	12.43%
Spring 1995	148	20	0	4	24	20	59.74%	40.26%	16.22%
Summer 1995	104	11	0	2	13	4	52.28%	47.72%	12.50%
Fall 1995	87	6	1	4	11	18	45.67%	54.33%	12.64%
Spring 1996	58	12	0	8	20	30	29.92%	70.08%	34.48%
Summer 1996	8	1	0	2	3	5	18.70%	81.30%	37.50%
		96	6	44	146	257			

Table 15: Achievement Rate by the Needs for Developmental Courses, No Needs for Developmental Courses

Semester	Head Count	Achievement				Dropped Out	Proportion of		
		Degree	Transfer	Sophomore	Sum		Non Achiever	Achiever	Achieving
Fall 1992	166	0	0	0	0	21	100.00%	0.00%	0.00%
Spring 1993	145	0	4	1	5	15	96.55%	3.45%	3.45%
Summer 1993	125	0	0	0	0	1	96.55%	3.45%	0.00%
Fall 1993	124	0	0	1	1	7	95.77%	4.23%	0.81%
Spring 1994	116	13	0	7	20	19	79.26%	20.74%	17.24%
Summer 1994	77	6	0	3	9	5	70.00%	30.00%	11.69%
Fall 1994	63	10	1	1	12	2	56.66%	43.34%	19.05%
Spring 1995	49	9	0	2	11	4	43.94%	56.06%	22.45%
Summer 1995	34	4	0	0	4	2	38.77%	61.23%	11.76%
Fall 1995	28	2	0	1	3	7	34.62%	65.38%	10.71%
Spring 1996	18	3	0	4	7	8	21.16%	78.84%	38.89%
Summer 1996	3	0	0	0	0	3	21.16%	78.84%	0.00%
		47	5	20	72	94			

Table 16: Achievement Rate by the Needs for Developmental Courses, One or Two Developmental Courses

Semester	Head Count	Achievement				Dropped Out	Proportion of		
		Degree	Transfer	Sophomore	Sum		Non Achiever	Achiever	Achieving
Fall 1992	165	0	0	0	0	22	100.00%	0.00%	0.00%
Spring 1993	143	0	0	0	0	18	100.00%	0.00%	0.00%
Summer 1993	143	0	0	0	0	0	100.00%	0.00%	0.00%
Fall 1993	125	0	0	4	4	17	96.80%	3.20%	3.20%
Spring 1994	104	5	0	6	11	14	86.56%	13.44%	10.58%
Summer 1994	79	2	0	0	2	4	84.37%	15.63%	2.53%
Fall 1994	73	8	0	1	9	6	73.97%	26.03%	12.33%
Spring 1995	58	6	0	1	7	8	65.04%	34.96%	12.07%
Summer 1995	43	7	0	2	9	0	51.43%	48.57%	20.93%
Fall 1995	34	0	1	2	3	9	46.89%	53.11%	8.82%
Spring 1996	22	7	0	4	11	8	23.45%	76.55%	50.00%
Summer 1996	3	1	0	1	2	1	7.82%	92.18%	66.67%
		36	1	21	58	107			

Table 17: Achievement Rate by the Needs for Developmental Courses, More than Two Developmental Courses

Semester	Head Count	Achievement				Dropped Out	Proportion of		
		Degree	Transfer	Sophomore	Sum		Non Achiever	Achiever	Achieving
Fall 1992	72	0	0	0	0	4	100.00%	0.00%	0.00%
Spring 1993	68	0	0	0	0	8	100.00%	0.00%	0.00%
Summer 1993	60	0	0	0	0	0	100.00%	0.00%	0.00%
Fall 1993	60	0	0	0	0	6	100.00%	0.00%	0.00%
Spring 1994	54	0	0	0	0	4	100.00%	0.00%	0.00%
Summer 1994	50	0	0	0	0	1	100.00%	0.00%	0.00%
Fall 1994	49	2	0	0	2	6	95.92%	4.08%	4.08%
Spring 1995	41	5	0	1	6	8	81.88%	18.12%	14.63%
Summer 1995	27	0	0	0	0	2	81.88%	18.12%	0.00%
Fall 1995	25	4	0	1	5	2	65.51%	34.49%	20.00%
Spring 1996	18	2	0	0	2	14	58.23%	41.77%	11.11%
Summer 1996	2	0	0	1	1	1	29.11%	70.89%	50.00%
		13	0	3	16	56			

Logistic regression. Testing Null Hypothesis was $\beta = 0$. The model for the analysis was

$$\log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \alpha + \beta_1 TERM1 + \beta_2 TERM2 + \beta_3 TERM3 + \beta_4 TERM4 + \beta_5 TERM5 +$$

$$\beta_6 TERM6 + \beta_7 TERM7 + \beta_8 TERM8 + \beta_9 TERM9 + \beta_{10} TERM10 + \beta_{11} GPA1 + \beta_{12} GPA2 +$$

$$\beta_{13} GPA3 + \beta_{14} GPA4 + \beta_{15} GPA5 + \beta_{16} GPA6 + \beta_{17} GPA7 + \beta_{18} GPA8 + \beta_{19} GPA9 + \beta_{20} GPA10 +$$

$$\beta_{21} HOUR1 + \beta_{22} HOUR2 + \beta_{23} HOUR3 + \beta_{24} HOUR4 + \beta_{25} HOUR5 + \beta_{26} HOUR6 + \beta_{27} HOUR7 +$$

$$\beta_{28} HOUR8 + \beta_{29} HOUR9 + \beta_{30} HOUR10 + \beta_{31} SEX + \beta_{32} RACE + \beta_{33} YOUNG + \beta_{34} IND01$$

Based on stepwise of logistic regression, three variables were selected from above model. They were GPA6, GPA7, and GPA5 (see Table 18). Rooted in *Wald Chi-Square* of this study, GPA6, i.e., GPA in 6th semester, had the strongest effect in predicting student success among the three variables. GPA5 was the weakest one (see Table 19).

Table 18: Stepwise of Logistic Analysis

Step	Stepwise Procedure	-2 LOG L Intercept Only	-2 LOG L Intercept & Covariates	Chi-Square for Covariates	Residual Chi-Square	P Value
0					448.9231	0.0001
1	+ GPA6	993.836	936.069	57.7669	425.6566	0.0001
2	+ GPA7	993.836	884.688	109.1476	371.4261	0.0001
3	**+ GPA5**	**993.836**	**830.705**	**163.1310**	**371.9791**	**0.0001**
4	+ TERM10	993.836	804.421	189.4146	289.1338	0.0001
5	- TERM10	993.836	830.705	163.1310	371.9791	0.0001

Table 19: The Result of Logistic Analysis, Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
Intercept	1	-3.2126	0.1329	584.5597	0.0001		
GPA5	1	0.6383	0.0785	66.0672	0.0001	0.290967	1.893
GPA6	1	0.8512	0.0853	99.4853	0.0001	0.319235	2.343
GPA7	1	0.9547	0.1088	76.9461	0.0001	0.260652	2.598

Then the model equation can be written as follows:

$$\log\left(\frac{\pi(x)}{1-\pi(x)}\right) = -3.2126 + 0.6383\ GPA5 + 0.8512\ GPA6 + 0.9547\ GPA7$$

Table 20 lists the parameters' interpretations, and Table 21 displays the predicted logits and odds of students' achievement. The odds ratio with lower GPA verse higher GPA in the sixth semester, which was upgraded one level, is the ratio of predicted odds of students' achievement. It has been shown to

$$e^{\hat{\beta}_{16}} = e^{0.8512} = 2.343$$

Table 20: Interpretation of Parameters

Parameter	Estimate	Standard Error	Interpretation
α	-3.2126	0.1329	Log odds of student's achievement with low GPA5, GPA6, GPA7
β_{15}	0.6383	0.0785	Increment to log odds with upgrading one level of GPA5
β_{16}	0.8512	0.0853	Increment to log odds with upgrading one level of GPA6
β_{17}	0.9547	0.1088	Increment to log odds with upgrading one level of GPA7

Table 21: Model-Predicted Logits and Odds of Student's Achievement

GPA5	GPA6	GPA7	Logit	Odds of Achievement
↓	↓	↓	$\hat{\alpha}$ = -3.2126	$e^{-3.2126}$=0.04025
⇑	↓	↓	$\hat{\alpha} + \hat{\beta}_{15}$ = -3.3126 + 0.6383 = -2.5743	$e^{-2.5743}$=0.07621
↓	⇑	↓	$\hat{\alpha} + \hat{\beta}_{16}$ = -3.3126 + 0.8512 = -2.3614	$e^{-2.3614}$=0.09429
↓	↓	⇑	$\hat{\alpha} + \hat{\beta}_{17}$ = -3.3126 + 0.9547 = -2.2579	$e^{-2.2579}$=0.10457
⇑	⇑	↓	$\hat{\alpha} + \hat{\beta}_{15} + \hat{\beta}_{16}$ = -3.3126 + 0.6383 + 0.8512 = -1.7231	$e^{-1.7231}$=0.17851
⇑	↓	⇑	$\hat{\alpha} + \hat{\beta}_{15} + \hat{\beta}_{17}$ = -3.3126 + 0.6383 + 0.9547 = -1.6196	$e^{-1.6196}$=0.19798
↓	⇑	⇑	$\hat{\alpha} + \hat{\beta}_{16} + \hat{\beta}_{17}$ = -3.3126 + 0.8512 + 0.9547 = -1.4067	$e^{-1.4067}$=0.24495
⇑	⇑	⇑	$\hat{\alpha} + \hat{\beta}_{15} + \hat{\beta}_{16} + \hat{\beta}_{17}$ = -3.3126+0.6383+0.8512+0.9547 = -0.7684	$e^{-0.7684}$=0.46375

Students with higher GPA have over two times higher odds for achievement than students with lower GPA in the sixth semester in the study. For example, if a student enrolled 6th semester, and if he or she upgraded his or her grade one level, such as from "C" to "B", the student would be more than twice as likely to be an achiever than a student without upgraded grade. If a student's GPA in 5th, 6th, and 7th semesters was 2.15, 2.50, and 2.75, then the equation is

$$\log\left(\frac{\pi(x)}{1-\pi(x)}\right) = -3.2126 + 0.6383 \text{ X } 2.15 + 0.8512 \text{ X } 2.50 + 0.9547 \text{ X } 2.75 = 2.91317$$

$$\pi(x) = \frac{1}{1+e^{-2.91317}} = 0.9485$$

This means that the student had a 94.85% chance to become an achiever. However, if another student's GPA in 5th, 6th, and 7th semester was 1.35, 1.40, and 1.20, then equation becomes:

$$\log\left(\frac{\pi(x)}{1-\pi(x)}\right) = -3.2126 + 0.6383 \text{ X } 1.35 + 0.8512 \text{ X } 1.40 + 0.9547 \text{ X } 1.20 = -0.013575$$

$$\pi(x) = \frac{1}{1+e^{0.01358}} = 0.4966$$

Because of the difference of GPA in these three semesters, the probability of achievement for the second student was about half of that of the first student, no matter what GPA the student had in other semesters. The analysis indicated that students' persistence and grades in 5th, 6th, and 7th semesters were crucial factors. These crucial factors decided whether a student achieved in his or her college life. If students enroll and have satisfactory GPA in their 5th, 6th, and 7th semesters, they have more chance to success whether or not they needed developmental course.

Findings and Conclusions

The findings from the "Achievement" part of the analysis are easily understood. This is a common approach. However, it may be inefficient and biased because the result is based on a comparison of the proportion of subjects who had failed by the end of the study (Greenhouse, et al., 1989). The proportions are not correct as they are calculated using the number of students in each group from fall 1992 as a denominator. This method treats dropped out (censored) students as if they completed 12 semesters of study. However, up to the time when students dropped out (censored) or withdrew from the college, they had not failed their study. Therefore, total length of follow-up for all students—even when the length of follow-up differed among students—is required. Because of reasons discussed here, the findings from this part are for reference purpose only.

The findings using the Kaplan-Meier method indicate that the fewer developmental courses students took, the higher their probability for success. White students had higher probability to become achievers than minority students. Students with high overall GPA were more successful than students with low overall GPA. However, students' genders and their ages were not meaningful variables to analyze and predict students' achievement.

The findings using logistic regression is encouraging for students who enroll in developmental courses. They suggest that students' achievement is highly linked to their GPA in 5th, 6th, and 7th semesters rather than to other factors. As long as the students overcome the developmental courses, they may get through college education as well as college-prepared students do. The findings also suggest that educators should continue to offer developmental courses. In general, the open-door policy at the community college is successful.

Bibliography

Agresti, A. (1990). *Categorical data analysis.* John Wiley & Sons: New York.

Allison, S. (1995). *Survival analysis using the SAS system: a practical guide.* SAS Institute Inc: Cary, NC.

Eanes, R. (1992). Linking college developmental reading and English courses for general education courses. *ED352635.*

Ignash, J. (1997). Who should provide postsecondary remedial/developmental education. *New direction for community colleges. 100,* 5.

Greenhouse, J., Stangl, D., and Bromberg, J. (1989). An introduction to survival analysis: statistics methods for analysis of clinical trial data. *Journal of consulting clinical psychology, 57,* 4, 536-544.

Gruber, F. (1999). Tutorial: survival analysis---a statistic for clinical, efficacy, and theoretical applications. *Journal of speech, language, and hearing research, 42,* 432-447.

Hopper, P.; Taylor, R., Wolford, P. (1997). Success rate of developmental vs. nondevelopmental students in freshmen college level English. *ED409065.*

Lyons, D. (1990). Success of community college students completing developmental courses. *ED325194.*

McCullagh, P. & Nelder, J.A. (1983). *Generalized linear models.* Chapman and Hall: London.

McMillan, V., Parke, S. & Lanning, C. (1997). Remedial/developmental education approaches for the current community environment. *New direction for community colleges. 100,* 21.

Ronco, S. (1994). Meandering ways: studying student stop out with survival analysis. *ED373633.*

Schumacker, R., Denson, K. (1994). Interpreting significant discrete-time periods in survival analysis. *ED 383767.*

Singer, J. (1993). Are special educators' career paths special? Results from a 13-year longitudinal study. *Exceptional children 59,* 3, 262-279.

Singer, J., Willett, J. (1993). It's about time: using discrete-time survival analysis to study duration and the timing of events. *Journal of education statistics. 18,* 2, 155-195.

Stokes, M., Davis, C., and Koch, G. (1995). *Categorical data analysis using the SAS system.* SAS Institute Inc: Cary, NC.

Tamada, M., Inman, C. (1997). Survival analysis of faculty retention data: how long do they stay? *ED410864.*

Weissman, J, Bulakowski, C. & Jumisko, M. (1997). Using research to evaluate developmental education programs and policies. *New direction for community colleges. 100,* 73.

Xiao, B. (1997). Using discrete-time survival analysis to examine time to Master of Science degree attainment. Paper present at the Association for Institutional Research 37th Annual Forum, May 1997.

Xiao, B. (1998). Factors influencing Master's degree attainment in business, engineering, health and human sciences, and visual and performing arts. *ED424807.*

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

☑ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").