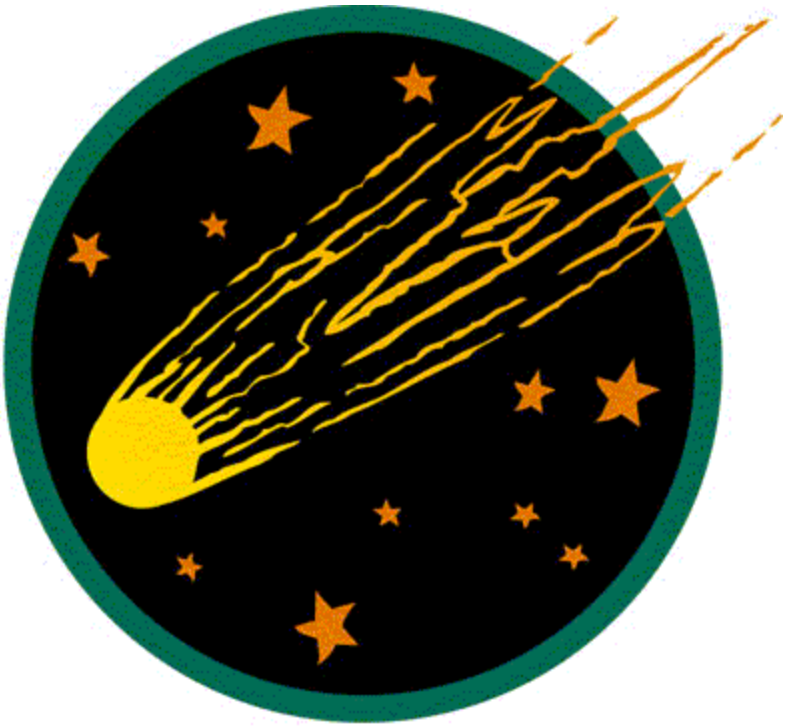
An Analysis of First-Time Freshmen Retention at UT Dallas Using the Entering Freshman Classes of 1996, 1997 and 1998



Prepared by:
Matthew J. Ramseyer
Office of Strategic Planning and Analysis

February 2000

An Analysis of First-Time Freshmen Retention at UT Dallas Using the Entering Freshman Classes of 1996, 1997 and 1998

Freshmen retention is regarded by many in the higher education community as a significant measure of customer satisfaction for undergraduate education. U.S. News and World Report, in its ranking of America's best colleges, attributes 20% of its national ranking and 25% of its regional ranking to student retention with 5% of its overall measure specifically directed to freshmen retention. Students who make it through their first year successfully have been found to have a much greater probability of graduating in a timely manner. In addition, getting students successfully through their first year of the curriculum in a manner that will promote their continued academic success as well as secure a positive future career is an important mission of any higher education institution. Because of this underlying reason, many states, including Texas, regard freshmen retention rates as a critical measure of institutional effectiveness when examining the performance of their public institutions.

Purpose

The purpose of this report is to examine the characteristics of the first time freshmen that entered in the Fall of 1996, 1997, 1998 and returned to classes in the Fall of 1999. The report also identifies the patterns in the characteristics of freshmen that were not retained during these three years. This analysis will provide information that will assist efforts to create, modify, or replace existing institutional structures in the effort to improve higher education outcomes for students, as well as improve the operations of the university.

Data

The data for this analysis is comprised of the freshman cohorts that entered in the Summer or Fall Semesters of 1996, 1997, or 1998. Summer first-time freshmen that did not remain in the fall semester were not included in the analysis and are considered transient with intentions to begin their academic career at other institutions. Students that entered in the summer semesters are only considered if they remained in the subsequent Fall semester.

All students are considered "first time in college" as defined by Texas Higher Education Coordinating Board Reports. Students coming to UT Dallas from the Texas Academy of Math and Sciences (TAMS) are also included in this study although these students have earned significant college credits prior to entering the university. Table 1 provides a summary of the freshmen cohorts in the analysis.

TABLE 1
BREAKDOWN OF FIRST-TIME FRESHMEN COHORTS

Year	First-Time Freshman Entering in the Summer[1]	First Time Freshmen Entering in the Fall	Total
1996	12	467	479
1997	7	472	479
1998	5	523	528
Total	**24**	**1462**	**1486**

[1] While there were more "First-Time Freshman" in the Summer Semesters, these students include only those who remained for the next Fall semester. The assumption is that "first-time freshman" who enroll in the summer but do not return in the fall intend to get a start on academic careers which they will pursue at other institutions. A total of 24 first-time freshman enrolled in the Summers of 1996, 1997, and 1998 that did not remain in the subsequent Fall semester.

The initial source of data for building the cohorts is the Texas Higher Education Coordinating Board Student Report (CBM001). The Texas Higher Education Coordinating Board reports are the official state data on enrollment and other institutional measures. Additional supplemental data was taken from the Freshman Database that is compiled by the Office of the Dean of Undergraduate Studies. Further information regarding student first year grade point averages and courses enrolled for in the first year were taken from historical files.

Analysis

The analysis initially consists of a series of cross tabulations that examine retention rates along the various variables in the data. This level of analysis however is limited in that it can capture at most three dimensions of student retention. Due to this limitation, a logit regression model will be utilized in order to generate predicted probabilities of student retention controlling for all available dimensions in the data.

Retention by Year

As can be seen in Table 2, the overall freshman retention rate for all three years cohort is 72.81%. This rate however has fluctuated slightly over the past three years.

TABLE 2
FIRST-TIME FRESHMAN RETENTION RATES BY YEAR

Year	Retention Rate	N
1996	72.23%	479
1997	71.40%	479
1998	74.62%	528
All	**72.81%**	**1486**

The difference increase in retention rate for the larger Fall 1998 class is promising. From a statistical point of view however it is not a highly significant difference. Difference of means tests between retention rates for 1998 and 1997 reveal a z-score of 1.15, only a 75% probability of being significant different.

Retention by Full-Time/Part-Time Status

The majority of first time freshman at UT Dallas are full-time students. Part-time status could indicate employment, family obligations or some other extraneous responsibility that keeps students from enrolling full-time. Retention rate based on full-time/part-time status for the first Fall semester enrolled is presented in Table 3.

TABLE 3
FIRST-TIME FRESHMAN RETENTION RATE BY FULL-TIME/PART-TIME STATUS

	Retention Rate	N
Full-Time	74.03%	1371
Part-Time	58.26%	115
All	**72.81%**	**1486**

Clearly there is a significant difference in the retention rates of Full-Time Freshmen and Part-Time Freshmen. Difference of means test reveal a z-score of 3.31, significant at the 0.001 level. This finding indicates that part-time students, who are likely to have responsibilities outside of their education, are less likely to be retained than full-time students.

Standardized Test Scores

SAT Scores were available for 1358 of the 1486 freshmen. Of the students lacking SAT scores, 77 reported ACT Scores. A total of 1435 students out of 1486 had data on standardized test scores. ACT Scores were converted to SAT scores using methods established by the College Board (Marco et al, 1995). The breakdown of freshmen retention by SAT Verbal, SAT Math and SAT Total Scores is presented in Table 4. There is a different N for SAT Verbal-Math scores and total scores, because the total score breakdown includes the converted ACT scores. There is no Verbal-Math breakdown for converted ACT scores.

TABLE 4
AVERAGE SAT SCORES FOR RETAINED AND NOT RETAINED FIRST TIME FRESHMAN

	Retained	**Not Retained**
Overall SAT (n=1435)	(n=1046)	(n=389)
Average	1202.01	1162.73
Standard Deviation	147.41	135.29
Minimum	800	770
Maximum	1570	1600
Math SAT	(n=995)	(n=363)
Average	618.43	594.38
Standard Deviation	80.15	78.02
Minimum	410	380
Maximum	800	800
Verbal SAT	(n=995)	(n=363)
Average	585.53	571.87
Standard Deviation	95.23	81.67
Minimum	270	200
Maximum	800	800

As can be seen, retained freshmen had higher SAT Score Averages then did students that were not retained. Using a difference of means test on the overall SAT scores between retained and not retained freshman reveals a statistically significant difference with a z-score of 4.78, a probability of significance beyond 99.9%.

High School Class Rank

Percentile rank in high school graduating class was available on 1361 of the 1486 students in the Freshman cohorts. Table 5 presents a comparison of the average class rank of retained students and those not retained.

TABLE 5
AVERAGE HIGH SCHOOL CLASS RANKS OF RETAINED AND NOT RETAINED FIRST TIME FRESHMEN

	Retained	**Not Retained**
Class Rank	(n=999)	(n=362)
Average	18.41%	23.97%
Standard Deviation	15.77%	17.73%
Minimum	0.16%	0.23%
Maximum	88.9%	100%[2]

The average class rank for retained students was the top 18.41% while those not retained had an average of 23.97%. Difference of means tests indicate that the difference in class rank between retained and not retained students is statistically significant with a probability of more than 99.9% (z-score = 5.26).

[2] This individual was 20th in a class of 20 from Townview, however they had a 1350 SAT Score.

Tabulation of Class Rank and SAT Score

By putting SAT Scores and High School Class rank into categories and cross tabulating on retention rates an informative picture of how retention rates are distributed by these two variables can be seen. Table 6 presents the tabulation of retention rates by High School Class Rank and SAT Scores. The total number of students in each cell is in parentheses. Cells that have greater than the average retention rate (72.8%) and have more than thirty observations are in green, and cells that have less than average retention rates with more than thirty observations are in red. Not reported cells are not included in the color-coding.

TABLE 6
CROSS TABULATION OF FRESHMAN RETENTION ON CLASS RANK AND SAT SCORE
FALL 1996 TO FALL 1998 COHORTS

Class Rank	Not Reported	< 1000	>=1000 < 1100	>=1100 < 1200	>=1200 < 1300	>=1300 < 1400	>=1400 < 1500	>=1500	Total
Not Reported	70.2% (47)	50.0% (8)	33.3% (9)	65.0% (20)	60.0% (10)	80.0% (15)	71.4% (14)	100.0% (2)	66.4% (125)
>75%			100.0% (1)		25.0% (4)	0.0% (2)			28.6% (7)
<=75% >50%		50.0% (6)	47.1% (17)	66.7% (15)	68.8% (16)	66.7% (6)	0.0% (1)		**59.0% (61)**
<=50% >25%	0.0% (1)	90.9% (22)	**65.7% (70)**	**65.5% (116)**	**63.9% (97)**	**80.6% (36)**	90.9% (11)	100.0% (2)	**69.0% (355)**
<=25% >10%	100.0% (1)	**66.7% (48)**	**66.0% (106)**	**69.8% (116)**	**71.7% (106)**	**85.2% (54)**	77.3% (22)	60.0% (5)	**71.2% (458)**
<=10% >5%		76.9% (26)	**71.9% (32)**	**61.5% (39)**	**75.5% (49)**	**85.7% (42)**	76.2% (21)	75.0% (4)	**74.6% (213)**
<=5%	100.0% (2)	77.8% (18)	72.7% (22)	**83.0% (47)**	**83.3% (54)**	**94.3% (70)**	**92.2% (51)**	66.7% (3)	**86.5% (267)**
Total	70.6% (51)	**72.7% (128)**	**65.0% (257)**	**68.8% (353)**	**70.1% (336)**	**85.8% (225)**	**83.3% (120)**	75.0% (16)	**72.8% (1486)**

Class rank can be seen as good indicator of student retention. Looking at the Total column for class rank it can be seen that as students' high school class rank improves their likelihood of retention increases. Students in the 75% to 50% range have significantly lower than average retention rates, and these increase to well above average levels as the students move to class ranks of better than 10 percent.

SAT scores display a pattern with some notable differences. Looking again at the totals row, students with SAT Scores of less than 1000 have a higher probability of being retained than those with SAT scores between 1000 and 1300. Beyond the 1000 to 1100 range, the average retention level begins to increase; however, the 1400 to 1500 range has lower retention than the 1300 to 1400 range. Finally the scores above 1500 have a less than average retention rate although it should be noted that this category represents only 16 students.

The counterintuitive finding that students with SAT scores of less than 1000 have a higher likelihood of being retained than those with SAT scores of 1000 to 1300 is interesting. Much of the difference in retention rate for students with SAT scores of less than 1000 can be attributed to the high retention rates of students in the 50% to 25% of high school class, as well as those in the top 10% or better of their high school class Part of the explanation maybe attributed to study skills obtained in high school that put them in the top of their class, and the limited opportunities to attend other top ranked schools because of being "place bound" or not being accepted do to low SAT scores.

Scholarships

Scholarships are awarded to new freshmen with outstanding SAT scores and excellent standing in their high school graduating class. The correlation coefficient on scholarships and SAT scores is .6624. As would be expected, retention rates for students on scholarship are higher than for students not on scholarships. The retention rate for students receiving any type of scholarship is 86.0% (n=457) while those that did not receive a scholarship had a retention rate of only 67.0% (n=1029). A difference of proportion test indicates that the difference in retention rates of students on scholarship from those not receiving a scholarship is statistically significant at a probability of more than 99.9% (z=3.875).

Application Date

The hypothesis that there is a relationship between application date and retention was tested. Due to difficulty in getting precise application date information from the student information system, application date was only available for 1173 of the 1486 freshmen in the 3-year cohort. Table 7 provides retention rates by number of months between the time students application was received and when the student began classes.

TABLE 7
FIRST-TIME FRESHMAN RETENTION RATES BY NUMBER OF MONTHS FROM APPLICATION BEING RECEIVED AND BEGINNING OF CLASSES

	N	Retention Rate
Less than 3 Months	548	70.26%
3 months to less than 6 months	462	76.84%
6 months to less than 9 months	124	79.84%
More than 9 months	39	82.05%

Based on the data available it is apparent that students who apply well in advance of their first semester are more likely to be retained than those who do not apply in advance. Table 7 also shows that students that applied less than three months prior to the beginning of the semester had lower than average retention rates. This result is most likely a reflection of the student's level of interest in coming to UT Dallas and beginning college in general.

Gender

Table 8 details retention rates by gender and gives the gender breakdown of the 1996, 1997 and 1998 first time freshman classes. While the majority of first time freshmen students at UT Dallas are male, there is no significant difference in retention by gender.

TABLE 8
FIRST-TIME FRESHMAN RETENTION RATES BY GENDER FOR FALL 1996 TO FALL 1998 COHORTS

	N	% of All Cohorts	Retention Rate
Female	637	42.87%	71.15%
Male	849	57.13%	74.09%
Total	1486	100.0%	72.81%

Ethnicity

Table 9 details retention rates by student ethnicity as reported on the CBM001 reports. The category International represents students from other countries. These students can be of any ethnic background.

TABLE 9
FIRST-TIME FRESHMAN RETENTION RATES BY GENDER FOR FALL 1996 TO FALL 1999 COHORTS

	N	% of All Cohorts	Retention Rate
African-American	63	4.24%	63.49%
Anglo	929	62.52%	70.08%
Asian	318	21.40%	84.28%
Hispanic	126	8.48%	70.63%
International	42	2.83%	69.05%
Native American	8	0.54%	62.50%
Total	1486	100.00%	72.81%

As demonstrated in Table 9, Asian students have the highest retention rate of all ethnic groups. All other ethnic groups have lower than average retention rates. Some explanation of the data on this single, but complex element is warranted. Ethnicity is a complex variable that consists of many factors. For example, Asian students are composed of a widely varied group of individuals from a multiplicity of cultural backgrounds. Asians may consist of first generation or sixth or seventh generation individuals. "Hispanic" and other ethnic groupings also have similar complexities that are not captured by the categories available. In addition, the interaction that ethnicity has with other key variables such as Full-Time/Part-Time status, High School, and program of study all will reveal a more detailed understanding of the factors influencing retention at UT Dallas. From this point forward, examination of freshmen retention will by done on a multiple variables.

Full-Time/Part-Time Status by Ethnicity

Table 10 provides retention rates by ethnicity and full-time/part-time status. Totals for the cells are given in parentheses.

TABLE 10
FIRST TIME FRESHMAN RETENTION BY ETHNICITY AND FULL-TIME/PART-TIME STATUS

	Full-Time	Part-Time	All
African-American	66.10% (59)	25.00% (4)	63.49% (63)
Anglo	70.90% (873)	57.14% (56)	70.08% (929)
Asian	86.21% (283)	68.57% (35)	84.28% (318)
Hispanic	73.45% (113)	46.15% (13)	70.63% (126)
International	69.44% (36)	66.67% (6)	69.05% (42)
Native American	71.43% (7)	0.00% (1)	62.50% (8)
All	74.03% (1371)	58.26% (115)	72.81% (1486)

The low retention rate of part-time students has the impact of lowering the average retention rates for all ethnicity classifications. Even part-time Asian students show a significantly lower rate of retention than do their full-time counterparts. Clearly the difference in retention for full-time and part-time students is present in all ethnic categories.

School of Major by Ethnicity

Table 11 presents first-time freshman retention by school and ethnicity classification. The total number of people in a particular cell is given in parentheses. Cells highlighted in red have more than 30 students with retention scores less than the University average. Cells highlighted in green have more than 30 students with retention scores greater than the University average.

TABLE 11
FIRST-TIME FRESHMAN RETENTION BY SCHOOL OF MAJOR AND ETHNICITY

	Anglo	**African-American**	**Hispanic**	**Asian**	**Native American**	**International**	**All**
Arts and Humanities	**71.05% (38)**	50.00% (2)	66.67% (6)	75.00% (4)			**70.00% (50)**
Engineering and Computer Science	**74.63% (268)**	57.14% (14)	68.97% (29)	**84.82% (112)**	66.67% (3)	66.67% (15)	**75.97% (441)**
General Studies	45.45% (11)	0.00% (1)	100.00% (1)				46.15% (13)
Human Development	**66.15% (65)**	60.00% (5)	75.00% (8)	75.00% (8)		100.0% (1)	**67.82% (87)**
Management	**69.49% (118)**	58.33% (12)	80.00% (20)	**86.11% (36)**	100.0% (1)	57.14% (7)	**72.69% (194)**
Natural Science and Math	**72.99% (137)**	75.00% (12)	83.33% (18)	**87.5% (64)**		85.71% (7)	**78.15% (238)**
Social Science	73.08% (26)	83.33% (6)	75.00% (4)	100.00% (8)	100.0% (1)	100.0% (1)	**80.43% (46)**
Undeclared	**65.79% (266)**	63.64% (11)	**60.00% (40)**	**80.23% (86)**	33.33% (3)	63.64% (11)	**67.87% (417)**
All	**70.08% (929)**	**63.49% (63)**	**70.63% (126)**	**84.28% (318)**	62.50% (8)	**69.05% (42)**	**72.81% (1486)**

What is added to the analysis by this table is that students with majors in the Schools of Engineering and Computer Science, Social Science, and Natural Science and Math, all had retention scores higher than the University average. Students in the School of Management, Human Development, and Arts and Humanities all had retention rates below the University average. Students with majors in the School of General Studies also had below average retention scores, but due to small numbers, inferences regarding that school are limited.

The largest group of students with retention rates below the university average was undeclared students. Not only did these students have lower than average retention rates, but for Anglo, Hispanic, Asian, International, and Native American students, those with undeclared majors had lower retention rates then those students within those ethnic groups who had declared majors. This finding demonstrates that students with undeclared majors bring down the average retention rate for the University as a whole, as well as the retention rates for almost all of the ethnic categories.

Major by Ethnicity

Further breaking down Table 11, and looking at the actual major of students provides some insight into the overall retention rates of the school. Due to the detail of the category and the limited number of observations in most cases it is difficult to make many generalizations. Most notable is the lower than average retention rates of Psychology majors (66.1% for 59 total students). The low rate of retention for this major is the major influence for the lower than average retention rates in the School of Human Development. Students in Computer Science (79.65%) had a significant number of students (280) and retention rates that were higher than the university average. Asian students

in Electrical Engineering (90% with 40 students) as well as Business Administration (86.67% with 30 students) had higher than average retention rates overall as well as within the Asian ethnicity classification.

Retention Rates by High School

Of the 1486 students in the retention cohort, high school was successfully matched on 1441students. Of these 1441 students 921 or 63.9% came from 30 schools that had at least 8 students enroll at UT Dallas. Table 12 presents the retention rates at the 30 high schools which had the most first-time freshmen enroll in the Fall of 1996, 1997 and 1998. The total number is given in parentheses.

TABLE 12
FIRST-TIME FRESHMAN RETENTION BY HIGH SCHOOL AND ETHNICITY FALL 1996 TO FALL 1998

	Anglo	African American	Hispanic	Asian	International	Total
Rowlett	44.44% (9)					44.44% (9)
Irving	50.00% (4)		33.33% (3)	66.67% (3)		50.00% (10)
N Mesquite	37.50% (16)		100.00% (2)	100.00% (1)		52.38% (21)
Jesuit	57.14% (21)		50.00% (2)	100.00% (2)		58.33% (24)
Rockwall	57.14% (14)		100.00% (1)			60.00% (15)
R.L. Turner	16.67% (6)			100.00% (7)		61.50% (13)
Lake Highlands	58.82% (17)	100.00% (2)	50.00% (2)	60.00% (5)		61.54% (26)
Mesquite	40.00% (5)			100.00% (2)	100.00% (1)	62.50% (8)
Townview	71.42% (7)	66.67% (12)	46.67% (15)	100.00% (4)		63.16% (38)
Bishop Lynch	57.14% (14)			100.00% (3)		64.71% (17)
Allen	68.00% (25)	0.00% (1)	100.0% (2)	0.00% (1)		65.52% (29)
Colony	66.67% (9)			66.67% (3)		66.67% (12)
S Garland	71.43% (14)	0.00% (1)	100.00% (1)	50.00% (2)		66.67% (18)
W.T. White	66.67% (6)		50.00% (2)	100.00% (1)		66.67% (9)
Duncanville	60.00% (5)	100.00% (1)	100.00% (1)	100.00% (1)		75.00% (8)
Naaman Forest	73.68% (19)	50.00% (2)	100.00% (2)	80.00% (5)		75.00% (28)
Poteet	70.00% (10)		75.00% (4)	77.78% (18)		75.00% (32)
Richardson	69.57% (23)	0.00% (1)	100.00% (2)	88.89% (9)	100.00% (2)	75.68% (37)
Lakeview Centennial	75.00% (12)	100.00% (2)	0.00% (1)	77.78% (9)	100.00% (1)	76.00% (25)
Pearce	72.97% (37)		100.00% (3)	75.00% (4)	100.00% (2)	76.09% (46)
Garland	70.59% (17)	0.00% (1)	85.71% (7)	88.89% (9)		76.47% (34)
Coppell	83.33% (12)			72.73% (10)		78.26% (22)
Berkner	73.33% (30)	0.00% (1)	100.00% (3)	82.86% (35)	100.00% (1)	78.57% (70)
Newman Smith	70.00% (30)	66.67% (3)	100.00% (1)	88.89% (27)		78.69% (61)
Plano Senior	77.32% (97)		62.50% (8)	86.84% (38)	80.00% (5)	79.05% (148)
Wylie	77.78% (9)			100.00% (2)		81.82% (11)
Plano East	80.00% (55)	100.00% (2)	57.14% (7)	100.00% (16)	100.00% (2)	82.93% (82)
Skyline	72.72% (11)	66.67% (3)	100.00% (4)	100.00% (7)		84.00% (25)
Bryan Adams	75.00% (8)		100.00% (1)	100.00% (4)		84.62% (13)
N Garland	75.00% (8)	100.00% (2)	66.67% (3)	94.12% (17)		86.67% (30)

It is difficult to make conclusive analysis on many of these schools due to the few observations at each institution. For Schools such as North Garland (n=30), Plano East (n=82), Plano Senior High (n=148), Newman Smith (n=61), Berkner (n=70), Garland (n=34), Pearce (n=46), and Richardson (n=37) it can be said that students from these schools are more likely to be retained than students from other schools. These eight schools had 508 or 35.2% of the total freshman that enrolled at UT Dallas, and represent the high schools or the school districts in the immediate service area of the university.

While some schools have markedly lower retention rates than average, caution should be used in extrapolating conclusions regarding the curriculum and its ability to prepare students for UT Dallas. Nonetheless, as new cohorts are added to the database, retention from school districts with retention rates below the University average should be monitored to determine if advising should be targeted to individuals coming from these schools.

First Year Courses

Table 13 provides the grade point averages of retained and not retained first time freshmen in their first year (Fall and Spring semester) courses.

TABLE 13
FALL-SPRING GPA OF STUDENTS BY RETENTION STATUS

	Fall-Spring GPA	Standard Deviation of GPA
Retained	2.52	0.8027
Not-Retained	1.78	1.0684
Overall	2.33	0.9362

Clearly, students that were not retained had a lower grade point average than those that were not retained. In fact, on the average, students that were not retained did not leave the university in good standing. Of course as a cause of retention first year grade point average is not a good measure. If after students do not maintain a 2.0 grade point average in their first semester, they must have a 2.2 grade point average in their next semester, or they will be suspended and required to reapply to the university. It is likely a great deal of the effect of students that were not retained having lower than average grade point averages is a result of this policy.

By changing the unit of analysis from students to courses, courses in which students were enrolled that had more students that were not retained can be identified. Analysis along this line will provide insight into the relationship between some courses of the first year curriculum and freshmen retention. It is hoped that analysis of courses of first year students will reveal courses in which students had a “bad course experience” resulting in the student making a decision to leave the university.

The three-year cohort took a total of 381 different courses during their first year of enrollment. Out of these 381 courses the average retention rate of first time students per course was 78.11%. This means that all else held equal, the average course that a first time freshman was enrolled would have 78.11% of students return to the university the following fall semester. Many of these courses however had less than 50 first time freshmen students enroll in them over the three-year period. When looking at only courses that had at least a total of 50 first-time freshmen enrolled during the three-year period the average retention rate per course changes to 78.57, a relatively small difference from the overall average. Table 14 provides detail on freshmen retention by course for the 1996 to 1998 freshmen cohorts.

TABLE 14
FRESHMEN RETENTION BY COURSE FOR 1996 TO 1998 FRESHMEN COHORTS

	Number of Courses	Average Retention per Course	Standard Deviation of Retention
Total Enrollment >=50	42	78.57%	7.62%
Total Enrollment < 50	339	78.06%	27.49%
All Courses	381	78.11%	26.05%

As would be expected based on the central limit theorem, courses with enrollment of less than 50 have a greater variation in retention rates than those with enrollments greater than 50. Overall, however, the average retention rate per course differs little. Because of the variation in courses that had a low total enrollment of first-time freshmen, the analysis from this point on will focus on courses in which at least 50 total freshmen were enrolled. This will allow for the examination of courses in which the most reliable inferences can be made as well as direct to focus on courses that are part of the majority of first-time freshman curriculum. Table 15 presents first time freshmen retention rates by course for courses that had at least 50 students enrolled over the three-year period.

TABLE 15
FRESHMEN RETENTION RATES IN FIRST YEAR COURSES THAT HAD AT LEAST 50 FIRST-TIME FRESHMEN ENROLLED OVER LONG SEMESTERS AY 1996 TO AY 1998

Course Prefix	Course Number	Total Students Enrolled	Average Freshmen Retention Rate	Standard Deviations from the Mean
SOC	2301	67	58.21%	2.67
MATH	2312	245	68.16%	1.37
MATH	1314	356	68.26%	1.35
MATH	2190	104	68.27%	1.35
PSY	2301	246	70.73%	1.03
GOVT	2305	719	71.77%	0.89
CHM	1311	568	71.83%	0.88
CHM	1111	556	71.94%	0.87
CS	1115	91	72.53%	0.79
BA	2301	118	72.88%	0.75
NATS	1311	79	73.42%	0.68
RHET	1101	1436	73.82%	0.62
NATS	1111	77	74.03%	0.60
MATH	1325	226	74.34%	0.56
ECO	2301	191	74.35%	0.55
RHET	1301	1055	75.07%	0.46
IS	4183	69	75.36%	0.42
NATS	1411	206	75.73%	0.37
ACCT	2301	108	75.93%	0.35
HST	1301	720	75.97%	0.34
CS	1315	536	76.12%	0.32
A&H	1301	767	76.27%	0.30
MATH	1471	1218	76.85%	0.23
GOVT	2306	429	78.79%	0.03
A&H	2306	452	78.98%	0.05
ECO	2302	114	79.82%	0.17
CHM	1312	296	82.43%	0.51
CHM	1112	294	82.99%	0.58
NATS	1412	120	83.33%	0.63
STAT	1342	132	84.09%	0.73
HST	2301	197	84.77%	0.81
MATH	1326	93	86.02%	0.98
MATH	1472	773	86.93%	1.10
MATH	2420	72	87.50%	1.17
NATS	1112	51	88.24%	1.27
NATS	1312	51	88.24%	1.27
PHYS	2325	335	88.36%	1.29
ENGR	1100	159	88.68%	1.33
CS	2315	230	89.13%	1.39
MATH	2418	56	89.29%	1.41
EE	2310	88	89.77%	1.47
PHYS	2125	308	90.58%	1.58

Sociology 2301 stands out as the course which had the lowest retention rate of students that were enrolled in it. Full 2.67 standard deviations beyond the mean, the course Race, Gender and Class, an elective and not part of the general education core of the university, had a significantly higher number of freshmen enrolled in it over three years that

did not stay the following fall then did the average course. Math 2190, 2312 and 1314 (Pre-Calculus Lab, Pre-Calculus, and College Algebra) all had more students that were not retained then did the average course. These core courses can be seen as "gateways" which prevented some students, most likely lacking adequate math skills, from moving on to their next year at UT Dallas.

Logit Models

By employing regression analysis the multidimensional nature of the variables that influence student retention rates can be examined. Multiple regression allows for examination of the impact of multiple independent variables on the dependent variable. By controlling for the multiple influences which are being measured, a better assessment of the significance and strength of underlying influences which are measured by the independent variables can be considered. For example, by considering a model which includes high school class rank, SAT scores, ethnicity and full-time/part-time status, the variables which have the strongest influence on retention can be seen. In addition, variables which appear to have strong influence in one or two dimension cross-tabulations, may be masking the effects of other variables. For example, part-time students may have a tendency to be undeclared students, so that when looking at just non-degree seeking students the effect of being part time may be masking the impact of being non degree seeking. Coupling this example with SAT and High School Class rank variables goes even further in providing an explanation of what is influencing retention. In a sense, by employing regression methods, the true impact of variables can be seen by controlling for other variables. In addition, regression models can present this data in a fashion that simply cannot be done with cross tabulations.

When using regression analysis to examine the impact of independent variables on a dependent variable that is dichotomous (taking a value of "yes" or "no"), such as whether an individual was retained or not, the assumptions of OLS regression are violated. While the details of the assumptions and the violations of OLS can be examined in other literature[3], the violation which is most obviously problematic is the likelihood of generating predicted probabilities of the dependent variable which are greater than one, or less than zero. By utilizing the logistic function and maximum likelihood estimation, estimation of coefficients will produce predicted values that are between the values of zero and one.

Table 16 presents the results of six different maximum likelihood logit models on the probability of being retained for first-time freshmen.

[3] See Aldrich and Nelson, Linear Probability, Logit, and Probit Models. Quantitative Applications in the Social Sciences: Sage University Paper Number 45 and J. Scott Long, Regression Models for Categorical and Limited Dependent Variables. Advanced Quantitative Techniques in the Social Sciences no. 7.

TABLE 16
MAXIMUM LIKELIHOOD LOGIT RESULTS ON BEING RETAINED

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Coef.	Z	Coef.	Z	Coef.	Z	Coef.	Z	Coef.	Z	Coef.	Z
SAT	0.0015	3.33	0.0020	4.09	0.0015	3.19	0.0014	3.08	0.0014	2.95	0.0019	3.62
Class Rank	-1.7508	-4.78	-1.6587	-4.38	-1.7040	-4.62	-2.1461	-5.39	-1.6604	-4.71	-1.9500	-4.61
Part-Time[4]	-0.4965	-2.09	-0.6244	-2.54	-0.5288	-2.22	-0.5090	-2.09	-0.5040	-2.11	-0.6496	-2.57
Female[5]			-0.1608	-1.21							-0.1575	-1.07
Hispanic[6]			0.3117	1.39							0.3622	1.54
African-American			0.0041	0.01							0.0757	0.24
Asian			1.0812	5.72							0.9663	4.86
International			0.7892	1.38							0.6578	1.12
Engineering and Computer Science[7]					0.1593	0.62					-0.6814	-0.24
Human Development					-0.2928	-0.88					-0.2007	-0.58
Management					0.0425	0.15					-0.0602	-0.20
Natural Science and Math					0.2091	0.74					0.0824	0.28
Non-Degree					-0.0955	-0.38					-0.2204	-0.82
Jesuit							-0.4441	-1.02			-0.4451	-1.00
Lake Highlands							-0.2555	-0.61			-0.2280	-0.53
Townview							0.0545	0.15			-0.4784	-0.12
Richardson							0.2968	0.75			0.2523	0.62
Pearce							0.7742	1.88			0.8073	1.94
Garland							0.3465	0.83			0.2383	0.56
Berkner							0.6066	1.96			0.3330	1.04
Newman Smith							0.6540	1.97			0.3489	1.02
Plano Senior							0.7516	3.28			0.6891	2.92
Plano East							0.7700	2.50			0.7225	2.32
North Garland							0.9564	1.74			0.6877	1.22
Math 1314									-0.1835	-1.25	0.3592	0.22
Math 2312									-0.1968	-1.20	-0.1251	-0.71
Soc 2301									-0.8795	-2.80	-0.8526	-2.60
Constant	-0.3551	-0.65	-1.1547	-1.82	-0.3476	-0.57	-0.3844	-0.69	-0.9758	-0.17	-0.9381	-1.29
Log Likelihood	-763.27		-742.51		-760.63		-748.64		-757.74		-727.38	

Model 1 contains the base variables of SAT, Rank, and part-time status. These three variables were found to be significant in cross-tabulations and remain significant and stable as variables are added in Models 2 through 6. In Model 2, student demographic characteristics are added which improve the explanatory power of the model[8]. Most notable is the significant coefficient on Asian. This coefficient indicates, as revealed in Table 9, that Asian students are more likely to be retained than non-Asian students. Model 2 also reveals that there is no significant difference in being retained between the other ethnic groups. Although Table 9 indicated that there were differences between these groups, when controlling for part-time status, SAT scores, and High School class rank, the difference becomes insignificant. Model 3 removes the student demographic variables and includes dummy variables for school of

[4] Full-Time is the base category
[5] Male is the base category
[6] For ethnicity variables the base category is Anglo and Native American
[7] For School of Major the base category is Social Science, Arts and Humanities and General Studies
[8] Log Likelihood goes from –763.27 to –742.51. Please see references in note 4 regarding how goodness of fit and changes in explanatory power are determined.

major. None of these variables are significant and the change in the log likelihood is not very large. Model 4 includes SAT scores, class rank, and part-time status along with dummy variables for major high schools that had a large number of students come to UT Dallas. Of the high schools, Berkner, Newman Smith, Plano East, and Plano Senior High were seen to have significant (.05 level) impacts in increasing the probability of student retention. Model 5 takes out the dummy variables for high schools and uses the courses that were found to have lower rates of retention from Table 14. As expected, students who enrolled in Sociology 2301 had a statistically significant coefficient that demonstrated that they were less likely to be retained[9]. The other courses, Pre-Calculus (Math 2312) and College Algebra (Math 1314) did not have statistically significant coefficients indicating that controlling for class rank and SAT score, these courses are not likely to decrease the probability of being retained.

Model 6 includes all of the variables from the previous models. All of the significant variables maintain their significance and are relatively stable save for two of the high school dummy variables. Berkner and Newman Smith become insignificant and cannot be said to have any significant effect on the probability of students being retained. Looking back to Table 12 it can be seen that these high schools have different retention rates between their two largest populations (Anglos and Asians) and these populations are approximately the same size. Plano Senior High and Plano East also have fairly different retention rates between the same two populations; however, the Anglo population is much larger than the Asian population. In other words, differentiating between ethnic populations when including high school dummy variables removes the impact of the high school variables indicating that in the cases of Berkner and Newman Smith, the differences in retention are mitigated by differences in ethnicity, while at Plano Senior and Plano East, the differences are not. Of course Model 6 would be improved by an increased number of observations, and the significance of all variables may increase. However, it is unlikely given the number of observations in Model 6 (n=1357) that the variables included would be reduced in significance.

Predictions for individuals can be calculated by plugging individual characteristics into the logistic distribution function: $1/1+e^{-BX}$ where e is the natural logarithm, B is the vector of coefficients generated by the logit model, and X is the matrix of individual outcomes. For example, using Model 6, the average student (SAT = 1191.36, Class Rank = 0.1989) who is not part-time, of an ethnicity category other than Asian, did not go to either of the Plano High Schools, and did not enroll in SOC2301 has a predicted probability of being retained of 86.71%. If the same student however was Asian, the predicted probability would jump to 94.49%. If the average student was again, not Asian, but part-time and did enroll in SOC2301 the predicted probability of retention would drop to 59.23%. Table 17 presents predicted probabilities for other examples.

By using logit models to more thoroughly explore the determinants of retention at UT Dallas a more complete understanding of the relationships between variables can be achieved. What using this methodology has uncovered is that variables that appear to have an impact in cross-tabulations have less or no discernable impact in the presence of other explanatory variables. For example, although undeclared students appear to have low retention rates, when coupled with SAT scores, class rank, and part-time status, the overall negative impact on retention of being undeclared disappears. Some of the same significance appears to occur to when considering ethnic categories along with other variables. For all but the Asian group, SAT scores and High School class rank better explain the differences in retention rates between ethnic groups. Using this technique for analysis, it can be seen that helping students achieve full-time status, and taking a closer look at courses like Sociology 2301, will help to improve student success in their first year at UT Dallas.

[9] While the logit model controls for student characteristics such as SAT score and class rank some concern was expressed regarding student "self selection" into SOC 2301. Appendix A provides detail on the students that enrolled in SOC 2301 and demonstrates that the population of students in the course is not different than the population of all first-time freshmen examined.

TABLE 17
EXAMPLES OF PREDICTED PROBABILITIES USING LOGIT MODELS

Student	Probability of Being Retained
Ethnicity = Asian SAT = 1400 Rank = 5% High School = Plano East Status = Full-Time SOC2301 = Did not Enroll	.986
Ethnicity = Not Asian SAT = 1400 Rank = 5% High School = Plano East Status = Full-Time SOC2301 = Did not Enroll	.964
Ethnicity = Not Asian SAT = 1200 Rank = 15% High School = Plano East Status = Full-Time SOC2301 = Did not Enroll	.938
Ethnicity = Not Asian SAT = 1200 Rank = 15% High School = Other than Plano or Plano East Status = Full-Time SOC2301 = Did not Enroll	.879
Ethnicity = Not Asian SAT = 1000 Rank = 25% High School = Other than Plano or Plano East Status = Part-Time SOC2301 = Did not Enroll	.682
Ethnicity = Not Asian SAT = 1000 Rank = 25% High School = Other than Plano or Plano East Status = Part-Time SOC2301 = Enrolled	.478

Conclusion

Differences in retention at UT Dallas can largely be determined by traditional measures of student ability such as SAT scores, and rank in high school graduating class. However, this study has demonstrated that part-time students are much less likely to be retained even when controlling for SAT and class rank. This finding demonstrates that students who do not, or cannot pursue full-time study will have less of a chance of being successful in their first year regardless of their high school rank or SAT scores.

Even when controlling for other significant factors, there is strong indication that Asian students are more likely to be retained regardless of SAT and high school class rank. The power and significance of the dummy variable for Asian in Model 2 and Model 6 indicate that a further example of the underlying causes of this distinction is worth investigation. Asian students at UT Dallas may be tied to family in the area and less likely to leave home for school. Regardless, it may be useful to identify cultural patterns utilized by Asian students to persist apply them for supplemental instruction and support programs for students who can be identified as “at risk”. Also of interest

regarding the ethnic category variables is that aside from Asians, when controlling for other characteristics, retention rates between ethnic groups do not appear to be significantly different.

The inclusion of high school indicators in the model demonstrated that students coming from Plano East and Plano Senior high school were more likely to be retained all else being equal. Other schools in the immediate vicinity of UT Dallas did not have a measurable impact when controlling for other variables. This finding indicates that the curriculum at the Plano School District may better prepare students for the rigors of UT Dallas than other schools. Investigation of the differences in curriculum between local school districts would provide insight into the performance differences between students. In addition, comparing data on the performance of High School graduates from the area with other state colleges and universities may reveal useful information regarding the relationship between high school preparation and success at competitive universities.

By including courses in which students were enrolled and then were more likely to leave the university in the analysis, the impact of student interaction with faculty could be measured to some degree. Although when controlling for other variables such as those indicating student ability, the significance of most of these variables declined, one course, Sociology 2301, was identified to have a significant impact on decreasing the likelihood of students being retained. While the reasons for the low retention rates of students coming out of this class merit further investigation, such as examination of teaching evaluations, the methodology which identifies such courses is useful in providing information to Deans and Senior Administrators about courses and what curriculum students should enroll in.

There are additional variables that would lead to a better understanding of which factors impact a student's decision to persist from their first year to second. These would include, but not be limited to residency on campus, involvement in extracurricular activities, as well as the level of unmet financial need. The study could further be improved with follow up surveys of students regarding student perceptions of themselves and the university, as well as additional information regarding family background.

Appendix A
An Examination of Sociology 2301

Due to the implications of the findings regarding Sociology 2301, Race, Gender and Class a closer examination of the characteristics of the students enrolled in the course is merited. Not only will a closer examination validate the finding that first year students are less likely to be retained the following year, but it will also validate the methodology used in the study to control for student characteristics to identify factors impeding student success.

A total of 67 first time freshmen enrolled in Sociology 2301 during the Fall 1996, Fall 1997 and Fall 1998 semesters. Of these 67 freshmen, 28 were not retained the subsequent fall semester giving an overall retention rate for freshmen enrolled in the course of 58.2%. The retention rate however has not been the same from year to year. In 1996 it was 66.67% (n=24), in 1997 it was 44.0% (n=25) and in 1998 it was 66.67% (n=18). Details of student class rank and SAT scores are provided in Table A1.

TABLE A1
SAT AND CLASS RANK DETAIL FOR SOCIOLOGY 2301 FALL 1996, FALL 1997 AND FALL 1998

	Retained	Not Retained	All
SAT - Average	1215.26	1106.15	1170.94
SAT – Standard Deviation	165.69	120.47	157.48
Class Rank – Average	19.05%	25.47%	21.72%
Class Rank – Standard Deviation	17.21%	15.23%	16.59%

The first-time freshmen that enrolled in Sociology 2301 have similar SAT Scores and Class Rank as the rest of the students in the retention study. Although the breakdown in SAT Scores and Class Rank differ between retained and not retained students in the course, they are not dissimilar from the breakdowns of the entire population (see Tables 4 and 5).

With respect to ethnic breakdown, Sociology 2301 over the three fall semesters had 71.6% Anglo, 4.4% African-American, 7.5% Hispanic, 13.5% Asian, and 3% International. This overall breakdown is similar to the ethnicity breakdown of the overall population of first time freshmen although there are fewer Asian students and more Anglo students (see Table 9).

Based on SAT Scores, Class Rank, and Ethnicity there is no reason to believe that the population of students enrolled in Sociology 2301 differs strongly from the overall population of students in the freshman retention analysis. This finding indicates that something about the course, in and of itself, is having an impact on decreasing the probability that first-time freshmen that enroll will be retained.