Barriers Within: Examining Gender, Race, and Participation in the TRIO Program with College Success

BY

C2007
Sarah J. Frazelle

Submitted to the graduate degree program in Economics and the
Faculty of the Graduate school of the University of Kansas
in partial fulfillment of the requirements of the degree of
Master's of Arts

_____
Chairperson

Committee members _____

_____

Date Defended April 20, 2007

The Thesis Committee for Sarah J. Frazelle certifies
That this is the approved Version of the following thesis:

Barriers Within: Examining Gender, Race, and Participation in the TRIO Program with College Success

Committee:

_____
Chairperson

Date approved: April 20, 2007

Abstract

In today's society earning a college education plays an extraordinarily important role in obtaining a living wage and thus is of importance for every individual. However, cross disciplinary research has found that gender, ethnicity, and academic and socio-economic backgrounds play a role in how well a student performs in college. The TRIO program is a federally-funded, national effort to help students from disadvantaged backgrounds complete their degree. Given these students are from an already disadvantaged background, this study looks inside the cohort and finds that race and gender have significant correlations with cumulative GPA for first year participants, regardless of adjusting for academic and socio-economic backgrounds. The study further examines TRIO students and their length of participation in the TRIO program and finds that participation is significantly and positively correlated with increases in cumulative GPA, especially for students from the most disadvantaged backgrounds.

I. Introduction

In November 2006, the state of Michigan banned the use of race and gender in the University of Michigan's undergraduate admission process thereby preventing admission to the university under affirmative action. Some people support the University's move, claiming that affirmative action is not suited for academia since students should be accepted based on standards of merit that are not related to race and gender. Supporters assume that after adjusting for factors such as low-income, parent education level, and disabilities, the playing field in academia between differing ethnicities should level out. Perhaps, however, there are other unquantifiable factors influencing the performance of different ethnic groups, thus implying that affirmative action could still be beneficial, especially if it can be shown that interventions and strategies can help students with disadvantaged status improve their success rates in college.

There is substantial documented evidence that first generation students and students who come from low income backgrounds have more obstacles in obtaining a college degree, regardless of the type of degree sought. According to national standards, students from first generation and low income backgrounds have lower GPAs, more difficulty staying in school and have lower graduation rates (Betts & Morell, 1999; Keane & Wolpin, 2000; Thayer, 2000) than students who do not have these factors in their background. Barriers to success could start very early, as the majority of these students enter college with lower ACT scores and high school GPAs. However, the problem needs attention if equal opportunity to education, regardless of societal background, is to be realized.

Intervention programs, such as TRIO, a group of federally granted intervention programs for students at risk of not completing their educational goals, have been organized to help students bridge the gap from high school to college and to help students achieve

college degrees. Since many at-risk students drop out of college within the first two years of their entrance, TRIO Student Support Services (SSS) focuses mainly on the freshman and sophomore populations, though they offer continued support for higher level students as well. The main facets of support that University of Kansas's SSS branch offers are personal advisors, with student/advisor ratios much smaller than the campus average, and personal tutoring. The goals of the SSS program are to raise awareness of campus procedures and available resources and to help with academics in math, science, and English areas.

Recent national efforts, stemming from the Government Performance and Results Act of 1993 and the U.S. Department of Education's strategic plan, have called for a more precise accounting of the effect and/or success of TRIO and SSS programs (Profile of Undergraduate in U.S. Postsecondary Education Institutions, 2006). Doing so, however, is more easily said than done. Historically, the accountability of accurate data recording for programs across the nation has been extremely variable between participating universities and colleges. Since the 1998-1999 school year, however, the national TRIO program has required more complete datasets for their annual project reports that each program turns in (Profile of the Student Support Services Program, 2004). Since that time, the data has become more representative of how the programs are actually doing, but accuracy still suffers in some departments due to time constraints and the work load of the employees.

The TRIO program at the University of Kansas differs, in some sense, to its neighboring programs, in that the director of the KU TRIO program, Ngondi Kumatuka, began a regional initiative for more accountability of the programs before the national effort started. Consequently, the director of the University of Kansas' SSS program has a relatively complete Access Database following its students dating back to 1998 with approximately 20 variables. There is also a paper trail with mostly complete data as far back as 1990; however,

the conversion to the new database appears to be a tedious and lengthy project, which has yet to be completed.

Though having the data is one part of the battle, the previous methods used by the TRIO for evaluation of the SSS program and judging its "efficiency" have some flaws. The main issues revolve around how to calculate the progress/successfulness of the students using a consistent, accurate, method. It is tempting to compare SSS students with national GPAs, retention rates and graduation rates of average American students. However, doing so is dangerous in that the SSS group have lower averages regardless of whether they entered the program or not. Therefore, expecting a fair comparison between students participating in the program and the national average is an unreasonable request of the program as the goals have been simply to improve the SSS students' success rates in college via higher cumulative GPAs and a higher percentage of degree completions.

Another method to judge program efficacy has been to compare SSS student averages with a random sample of other students from similar backgrounds. However, given the strong presence of the SSS program on most campuses, especially with their advertising directed to the said population of students, one may wonder what would compel students to enter the program. The KU SSS program offers little to no monetary rewards or incentives, which implies that these students could have a stronger persistence to succeed than their counterparts who do not attempt to enter the program. This could give an upwards bias to the results when comparing the KU SSS students to the underprivileged population as a whole. However, looking at a potential downward bias, many of the students that are accepted into the program have more than one factor playing against them. Some come from only a low-income, first-generation, or disabled background, however, others come from a background

including two or three of these situations. This could imply that their averages would be lower than their counterparts that come from only one of the above populations.

Demonstrating academic need is another requirement for admittance into the SSS program. Demonstrating academic need, as defined by TRIO means that students must additionally fit into one or more of the following categories: low high school GPA, low SAT or ACT scores, low college GPA prior to entering the program, failing grades for the current enrolled semester, out of academia for five or more years, or have a low predictive indicator score, which is a composite variable created by TRIO to estimate a student's potential success in college. The academic need category, therefore, implies that students who participate in the program are not only from disadvantaged family backgrounds, but are also students who would otherwise be categorized as having a low probability of completing an undergraduate degree.

One way to isolate the data is to use regression and panel regression techniques among the group. Comparing these students against their own previous year performance might lessen some of the bias that arises from comparing these students against the larger population of disadvantaged students or, worse, with the national averages of the "typical" student. One method would be to allow GPAs to be the dependent variable and the other twenty-four categorical variables, along with a variable for length of participation in the program, to the independent variables. Since data is collected about the students for the duration of their time in the program, it should be possible to link length of time in the program with changes in student GPAs.

This study begins by looking inside a group of first year KU SSS students and finds correlations between variables such as race, gender, age, enrollment status, grade level and academic need with cumulative college GPA, a proxy variable for success in college. The

study next follows these first year students through their subsequent years in the program, and finds in most cases, spending more years in the program has a positive, significant correlation with cumulative GPA.

II. Data

Data for this study was collected from the University of Kansas SSS department. The Human Subjects Committee of Lawrence granted permission to use the data. Participant information was gathered from existing KU SSS data dating from 1996 to 2005. It was a cumbersome but necessary process to manipulate the accumulated data into variables able for use. For details on how the data was collected and extrapolated, see the appendix.

Table 1 provides a demographic and GPA comparison between all KU undergraduates in 2005, the first year SSS cohort and the fourth year SSS cohort used in this study.

Table 1:

Percentage of Students	KU – Fall 2005	First Year SSS Cohort	Fourth Year SSS Cohort
Male	49.1%	37.6%	59.5%
Female	50.9%	62.4%	40.5%
Black	3.4%	20%	22.3%
Asian	3.6%	6.6%	8.3%
Hispanic	3.4%	7.3%	6.5%
White	79.3%	59.4%	58.6%
Other	4.7%	6.5%	4.2%
Freshman	21.7%	34%	-
GPA	2.917	2.438	2.56

Groups that stayed in the program for the entire four years may be different than those who only spent a few years. For example, the percentage of women in the program drops drastically. Given that females were the most advantaged group of first year students (as shown in Model 1, Table 2) and the program is voluntary, they might have found few advantages from continuing in the program. This may partly explain why the correlation

between being female and obtaining relatively higher cumulative GPAs was reversed when looking at the SSS cohort over time.

Another interesting trend was the number of students who leave the program each year. During the second year, 44% of the students left the program. The third year, another 20% were gone. By the fourth year, only 16% of the program's original students were still participating. Part of the reason may have pertained to students graduating. Since students may enter the program at any grade level, some first year students may only have had a few years left until graduation. Later entry into the program is a reasonable scenario since it could have taken a while for students to find out about the program and realize they needed extra help. Another possibility for the relatively low continuation in the program may be that very academically poor students dropped out of college after only a few semesters. If this is the case, the cohort for the third and fourth year may have been better students and thus reflected higher GPAs regardless of the benefits of the program.

III. Modeling the First Year

To capture the correlation between race, gender, and year in school and the discrete variable, cumulative GPA, a series of linear regressions were run with Model (1) being chosen based on the adjusted R-squared.

Model 1:

$$Cumgpa_i = \beta_0 + \beta_1 age_i + \beta_2 female_i + \beta_3 asian_i + \beta_4 black_i + \beta_5 latino_i + \beta_6 other_i + \beta_7 full_i + \beta_8 less_i$$

$$N \quad = \quad (1274) \quad (795) \quad (84) \quad (256) \quad (94) \quad (83) \quad (919) \quad (98)$$

$$+ \beta_9 freshna_i + \beta_{10} freshatt_i + \beta_{11} lact_i + \beta_{12} predin_i + \beta_{13} colgpa_i + \beta_{14} failgr_i + u_i$$

$$(224) \quad (216) \quad (210) \quad (382) \quad (291) \quad (95)$$

Gender was represented by the variable *female*, leaving male as the omitted variable. All students were divided into five race categories: Asian, African American, Hispanic (non-white), other, and White. Due to a small number of observations of some races, the variable

other was created to represent American Indians, Alaskan Natives, Native Hawaiian, Pacific Islanders, and those who reported more than one race. Whites were left as the omitted variable category since they are the predominantly advantaged group as found by tests run during this study.

Since not all students took the same number of courses each semester, it was important to adjust for the fact that some students in the KU SSS program were attending the university full-time and thereby carrying a heavier studying load. Other students attended less than half time, perhaps indicating working a full-time job while taking classes. Therefore, variables for enrollment status were incorporated for full-time status and less than half-time status[1], leaving other enrollment statuses as the omitted variable.

The next variable adjusted for grade level. Nelson (2003) points out that freshman student GPAs may be lower since they are acclimating themselves to new environments and may be inefficient in studying. She states, "As students learn how to study smarter, establish a wider knowledge base and learn the ropes of college life, they should become more efficient students." Thus two variables for freshman were created leaving the other undergraduates as the omitted variable[2]. Freshman were divided into two categories, *freshna*, which represented students who had never taken any college courses, and *freshatt*, which represented freshman who had some previous experience in college.

[1] For the first year students, robust OLS tests were run for students enrolled three-quarter time and half-time, however, the results were not statistically different from students reporting variable enrollment. When taking the students over time, however, three-quarter, half and less than half/varied enrollment variables were used. Full time was then left as the omitted variable category since it was found students enrolled full time were the most advantaged group of first year students.

[2] When taking the students over time, freshman who had no previous college exposure were taken as the omitted variable. Testing for significance found no difference between the freshman and sophomores. Furthermore, juniors and seniors were not found to be statistically different and were found to do significantly better than freshmen and sophomores.

The last group of variables incorporated the academic need category into the estimation of cumulative GPA. Including the academic need variable allowed for the possibility of seeing which academic background factors were most correlated with low cumulative GPAs. Caution was taken in analyzing these variables since they are binary variables, with students being placed in the one category which most accurately described their area of academic weakness. The most that this model could capture is a ranking of which variables best predict a low GPA. The variable *lact* was for students with ACT score less than 20. *Predin* was the variable for students scoring low on TRIO's predictive indicator assessment. *Colgpa* was for students who enter the program with a low college GPA. *Colgpa* should not be confused with the variable *failgr*, which was reserved for students who entered the program with failing grades for the current semester. Students placed in the *failgr* category may or may not have had a low cumulative GPA; they were merely performing poorly in the particular semester the data was recorded. Since students are allowed to start the program at any point during the semester, some would still have had time to improve their current grades so that their final grades were not failing. On the other hand, the variable for low college GPA would have a stronger correlation with low cumulative GPA since it could take many high grade semesters to bring up a low cumulative GPA. The omitted variable was mainly devoted to students who absolutely could not be categorized by one particular variable; however, due to the relatively small number of observations for these areas[3], it also included students with low high school GPAs, low diagnostic scores, and those out of academia for more than 5 years.

[3] The number of first year students with low high school GPAs was 12, low diagnostic scores was 1, and those out of the academic pipeline for more than five years was 6. The remaining 277 students were categorized as having more than one area of academic weakness.

IV. Modeling the Next Years

To capture the correlation between years in the program and cumulative GPA, three panel series models were used. First, a pooled OLS regression, model 2, was run.

Model 2:

$$cumgpa_{it} = \beta_0 + \beta_1 progyear_{it1} + \gamma_1 female_i + \gamma_2 asian_i + \gamma_3 black_i + \gamma_4 latino_i + \gamma_5 other_i + \gamma_6 full_i + \gamma_7 less_i + \gamma_8 freshatt_i + \gamma_9 badhs_i + \gamma_{10} predin_i + \gamma_{11} colgpa_i + \gamma_1 failgr_i + \varepsilon_{it}$$

where γ_i are the coefficients for the set of explanatory variables as defined for Model (1). This was the simplest way of capturing the notion that two observations from the same person were more alike than two observations from two different people. However, the pooled OLS model ignores the structure of the panel data.

To better capture the structure of the data, a random effects model, model 3, was run.

Model 3:

$$cumgpa_{it} = \beta_0 + \beta_1 progyear_{it1} + \varepsilon_{it}$$

where $\varepsilon_{it} = \alpha_i + u_{it}$

and α_i are time-invariant, person specific, and uncorrelated with years in program

and u_{it} is the random error.

The Hausman test was run but the random effects model fitted on this data failed to meet the asymptotic assumptions of the Hausman test. The rejection of the random effects model was to be expected since the first year cohort significantly showed a correlation between time-invariant variables and years spent in the program. Therefore the model was not used when analyzing the data.

TRIO is a voluntary program and, therefore, it is natural to assume that there would be a selection bias for participants. To correct for the selection bias a fixed effects model, model 4, was used.

Model 4:

$\Delta cumgpa_{it} = \beta_1 \Delta progyear_{it1} + \varepsilon_{it}$

where $\varepsilon_{it} = \alpha_i + u_{it}$

and α_i are the unobserved fixed effects (i.e. the individual effect which adjust for selection) and u_{it} is the random error.

A fixed effect estimator is robust to the omission of any relevant time-invariant variables. It obtains a consistent estimate of parameters despite the possible correlation of omitted effects. It allows individuals to control for themselves by basically including a different dummy variable for each individual. Therefore, finding a positive, statistically significant correlation with years in the program on cumulative GPA using a fixed effects model implies that the selection bias has been adjusted for.

The SSS sample was broken into two subgroups, one including individuals whose years in the program varied between one and nine years and a second group removing all individuals whose recorded participation in the program exceeded more than four years. These students were removed because those who finish in more than four years might perform worse academically throughout the program and thereby downward bias the effects of years in the program on GPA.

After running the models (1) through (3), each subgroup was further broken down into male, not white or Asian[4], black, and disabled but not low-income. A pooled OLS and fixed effect regression was then run on each of theses divided-sub-groups and tested for correlation between years in the program and cumulative GPA.

[4] White and Asian-Americans are placed together since they were not found to be significantly different from each other in previous tests used in this study.

V. Estimation Results

Tables 2-6, located in the appendix, display the results from the four models discussed for the different subgroups and divided-sub-groups. Table 2 displays the results for the first year students in Model (1). Table 3 and 4 display the results of all models run before extracting any students whose years of participation exceeded four years. Table 3 is for the pooled OLS models and table 4 is for the fixed effect models. Tables 5 and 6 display the results for the models run after extracting the longer participating students. Table 5 is for the pooled OLS models and table 6 is for the fixed effects model.

The first year cohort, Model (1), was tested to see if a student's background, whether it be low-income, first-generation, disabled, or low-income and first-generation, made a significant difference in cumulative GPA. The test did not find any significant difference between these groups and thus were removed from the model. Model (1) was run on all first year students, including the ones whose participation exceeded four years.

Initially, variables for age and years in the program were included in all pooled OLS and fixed effects estimates. Including both variables reported that the neither coefficient was significant in many of the models. Since a persons' age typically increases with each subsequent year in the program, age and year in the program variables are highly correlated. Therefore, age was removed from the models. Removing age from the models increased the significance of the variable years in the program in most estimates.

The first model adjusts for the first year students' gender, age, race, and academic background. Contrary to Betts and Morell's (1999) conclusion that adjusting for high school achievement shows no difference between whites and blacks or other races (non-whites) and whites, the model finds that adjusting for poor academic backgrounds does not diminish race effects between whites, blacks, and Hispanic non-whites, table 2. The correlations are still

highly significant and different from whites and Asians. The R-squared value implies that after adjusting for academic background the model explains 24% of the variation in cumulative GPA for this cohort. Although much is left unexplained, it is important to note these strong correlations of gender and race with GPA that are still apparent in the estimations for students in their first year of the SSS program.

Next the study followed the original SSS cohort over time. The first series of models were run including all students whose participation in the program ranged from one to nine years. With both the pooled OLS and fixed effects estimation on the full sample, the variable for years spent in the program was significant at the 5% level. However, after dividing the students by gender, race, and disability and running separate pooled OLS regressions on each subgroup (table 3); little correlation was found between the number of years spent in the program and cumulative GPA. There was no statistically significant correlation for the subgroup of males or blacks. The other subgroups, non-white or Asian, and disabled, display a positive correlation with how long they participate on cumulative GPA but only display significance at the 5% level. Fixed effects estimates were also run on each subgroup to help adjust for selection bias (table 4). The fixed effects estimates for the variable years in the program was positive and significant at the 1% level for all subgroups except the disabled subgroup, which was significant at the 5% level. Since the categories representing males and other ethnicities, columns 2-4 in table 4, are significant at the 1% level when running the fixed effects model but not when running the pooled OLS model, this implies that being male or of another ethnicity must be positively correlated with one or more of the other categorical variables from the pooled OLS model which display negative correlations with cumulative GPA. Using the fixed effects model would eliminate any of these correlations and allow the

true effect of years in the program to be seen for these sub-groups. Blacks appear to gain the most benefit from the program, followed closely by those who are not white or Asian.

The next set of regressions was run on students who spent only four or less years in the program. After isolating these students from the rest of the SSS cohort, years of participation in the program was still positively and significantly correlated at the 5% level using both the pooled OLS and fixed effect estimate (table 5, column 1 and table 6, column one). The "less than 5 year" subgroup was then divided into the same subgroups as the "all years" group and pooled OLS and fixed effects estimates were run.

Table 5 displays the results of the pooled OLS regressions. After extracting those whose participation exceeded 4 years, positive and significant correlations were found between years in the program and cumulative GPA for males and not white or Asian groups. The non white or Asian, and black subgroups became more significant, with the males becoming significant at the 5% level and the not white or Asian group becoming significant at the 1% level. However, the estimation for years in the program with cumulative GPA became statistically insignificant for the disabled subgroup.

Table 6 displays the results from the fixed effects estimates on those who participated in the program less than five years. Removing the correlation between the explanatory variables and cumulative GPA increased the level of significance for the male subgroup to 1%. The black and not white or Asian subgroups stayed significant at the 1% level. Interestingly, however, the disabled subgroup's estimate became insignificant when extracting students exceeding four years in the program. It could be postulated the disabled group became insignificant because most of them had years of participation greater than four years; however, only 94 students, or approximately 18%, of this cohort participated over four years. This is a relatively small loss compared to the 25% of males, 34% of not white or

Asians, or 34% of blacks who were removed from the full sample estimates. Another possible explanation could be the disabled students who finish with only four or fewer years of participation are either rushing to finish (as compared to those who took longer) or are not participating in the program long enough to take advantage of the benefits of the program.

Strong negative correlations were established over all regressions and time for four main categories. First, freshmen were correlated with low GPAs at the 1% significance level in all regressions. Finding that freshman do worse than their counterparts supports Nelson's argument that academia takes a period of adjustment for many students (2003). Next, students who were enrolled in less than half time had negative correlations at the 1% level for all regressions. This may be a signal that either these students had other obligations, such as work, or were not as strongly committed as the rest of the cohort. Having a previous low cumulative GPA was also significantly negatively correlated at the 1% level over all regressions. Though it may be possible for some students to increase their GPA over time, a low college GPA when entering the program would be difficult to increase due to the number of semesters it might take to substantially raise the cumulative GPA. Thus, even if these individuals become better students, their GPA may not substantially increase. The last group, Blacks and Latinos, were negatively correlated at the 1% level for almost every regression run, which is unsettling since the reason for this phenomenon is unobservable in the data. On a positive side, when looking in the group of non white or Asian and the group of Blacks, years in the program is highly positively correlated with cumulative GPA in the fixed effects model at the 1% level.

V. Conclusion and Areas for Further Research

The results of this study overwhelmingly support the correlation between cumulative GPA and gender, race, enrollment status, and grade level for the first year students despite

adjusting for socio-economic and academic backgrounds. For students from disadvantaged backgrounds at KU, white, female, non-freshman students attending classes full-time were the most advantaged group. Much research supports the findings that whites, females and non-freshman are advantaged (Rose 2005, Nelson 2003), however, the debate persists whether race and gender results stem from the socio-economic status of the student's family background, the student's personal academic background or some unobservable phenomenon. Despite adjusting for personal backgrounds, this study finds gender and race are still significant, with Asians displaying no significant differences from whites.

Having established gender and race correlations for this cohort of SSS students, two further questions were examined. The first, whether gender and race correlations persisted over time, and the second, whether years in the program was positively correlated with GPA over time. Establishing evidence of these correlations would support affirmative action plans since they imply that non-white and non-Asian students are at a non-quantifiable disadvantage and, despite negative social and academic backgrounds, these students' GPAs could be improved if the students were directed toward helpful resources such as SSS.

When examining the group of students over time, it appears the female advantage, strong in the first year students, diminishes and becomes a disadvantage over time. As previously mentioned, though, this could be from academically strong females not finding an advantage to voluntarily continuing in the program. In other words, perhaps those who do not need the program choose not to participate for very long. The SSS department at the University of Kansas records a categorical variable for why students leave the program, however, for the available data, the coding was inconsistent when merging individuals over years, and thus was not useful.

Despite the diminishing female correlation, race correlations persist over time, as evidenced by the series of regressions run with the panel data for SSS. This disturbing correlation is evidence of a non-quantifiable disadvantage to Blacks and Latino-Americans. All students in the SSS group come from similar disadvantaged socio-economic backgrounds. Furthermore, acceptance in the program is based on displaying some academic need, thus all students in the cohort are academically disadvantaged as well. Therefore, if the effects of race truly diminished after adjusting for high school background as Betts & Morrell (1999) suggest, then black and Latino-Americans in the SSS cohort should perform no differently than whites. However, this study still finds a negative correlation with blacks and Latino-Americans on GPA. Another interesting study could look at racial effects within varying income levels to examine whether such strong negative correlations persist between race and educational outcomes.

The effects of the KU SSS program seem to help improve its students' cumulative GPAs. There was a positive, statistically significant correlation between years in the program and cumulative GPA, though the correlations seemed stronger for the non-disabled compared to others. Separating the individuals with less than five years in the program from the total SSS cohort increased the level of significance for the number of years a student participated program. Those who were Black and not white or Asian apparently gained the most benefit. It is important to realize, however, that this study does not observe the trends and correlations of a control group. Thus, any findings in this study are inconclusive in regards to a comparison between the SSS students and the rest of the University of Kansas. Future research should follow a control group of students eligible for SSS but who do not participate. This method would have some selection bias, since the students who do not participate in the SSS program may be less motivated to do well in college as those who do,

however, using a fixed effects model would help alleviate some of this bias and would at least yield a rough comparison between SSS participants and the rest of the disadvantaged students at the University of Kansas.

References

Betts, Julian R. and Darlene Morell. 1999. The Determinants of Undergraduate Grade Point Average: The Importance of Family Background, High School Resources, and Peer Group Effects. *The Journal of Human Resources*: 238-293.

Glen, David. 2005. Sociologists Debate Degrees, Dropouts & Faculty Beliefs. *Chronicle of Higher Education*: 52, no2, A24-25, S, 2.

Johnston, Jack and John Dinardo. 1997. Econometric Methods, 4th Ed. McGraw-Hill Companies, Inc.

Keane, Michael P. and Kenneth I. Wolpin. 2000. Eliminating Race Differences in School Attainment and Labor Market Success. *Journal of Labor Economics*: 614-652.

Maldonado, David; Robert Rhoads, Tracy Puenovista. 2005. The Student-Initiated Retention Project: Theoretical Contributions and the Role of Self-Empowerment. *American Educational Research Journal*. Vol 42, No 4, 605-635, Winter 2005.

Nelson, Rebecca. 2003. Student Efficiency: A Study on the Behavior and Productive Efficiency of College Students and the Determinants of GPA. *Issues in Political Economy*. Vol. 12.

Profile of Undergraduates in U.S. postsecondary Education Institutions. 2003-2004. Statistical Analysis Report. U.S. Department of Education, NCES, 2006-184

Profile of the Student Support Services Program. 2004. 1997-1998 and 1998-1999 with Select Data from 1999-2000. U.S. Department of Education, Office of Post Secondary Education, Federal TRIO programs.

Steele, Claude M. 1997. A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance. *American Psychologist*. June 1997: 613-629.

Rose, H. 2005. The Effects of Affirmative Action Programs: Evidence from the University of California at San Diego. *Educational Evaluation and Policy Analysis, 27*(3), 263-289.

Thayer, Paul. 2000. Retention of Students from First Generation and Low Income Backgrounds. *The Journal of the Council for Opportunity in Education.* May 2000.

Office of Institutional Research and Planning. 2005. Grading Report for Fall Semesters. http://www2.ku.edu/~oirp/profiles.shtml

Office of Institutional Research and Planning. 2005. Head Count Enrollment-University Summary. http://www2.ku.edu/~oirp/profiles.shtml

Appendix

Data Collecting process

The data for this project is a subset of data collected from the TRIO program at the University of Kansas and was approved for use by the Human Subjects Committee of Lawrence. The data was initially received in six Excel workbooks ranging from 1994 to 2005. From these workbooks I created a unique ID number by concatenating birthdates with gender, ethnicity, and eligibility variables. I then combined all of the data and sorted in ascending order by the ID numbers. In order to establish how many years each individual had been in the program, I then wrote a numbering code in SAS which assigned each new ID a year one and then numbered sequentially until finding a new ID number. This allowed me to divide the subjects by number of years in the program.

All subjects with "no response" to one or more of the categorical variables have been removed from the data subset. Most variables were numerically coded in a form that STATA would not recognize, therefore, I have decomposed these variables into individual dummy variables. After creating the series of categorical dummy variables, I took out all individuals who were not enrolled and those who were graduate students. For the models only measuring those individuals who spent less than 5 years in the program, I wrote a formula to extract all of the years of data for those whose participation in the program exceeded four years.

Variable Definitions

Program Year: time-variant variable representing the number of years in the program.

Female: a dummy variable coded 1 if female, 0 otherwise.

Asian: a dummy variable coded 1 if ethnicity was Asian American, 0 otherwise.

Black: a dummy variable coded 1 if ethnicity was black, 0 otherwise.

Latino: a dummy variable coded 1 if ethnicity was Hispanic American, 0 otherwise.

>1 race report: a dummy variable coded 1 if ethnicity was Hispanic American, 0 otherwise.

Full time: a dummy variable coded 1 if enrollment status was recorded full time, 0 otherwise.

< Half-time: a dummy variable coded 1 for enrollment status reported less than half time or varied, 0 otherwise.

Freshman, never attended: a dummy variable coded 1 for students who had no previous college experience, 0 otherwise.

Freshman, previous courses: a dummy variable coded 1 for a first year college student who had previously taken at least one course, 0 otherwise.

Bad High School Indicators: a dummy variable coded 1 for students accepted into TRIO based on at least one of the following categories; low SAT math score, low SAT verbal score, low ACT score, low high school GPA, or low GED score, 0 otherwise.

Predictive Indicator: a dummy variable coded 1 if accepted into TRIO based on scoring low on a set of composite variables compiled by TRIO, 0 otherwise.

Low College GPA: a dummy variable coded 1 if accepted into TRIO based on having a low college GPA, 0 otherwise.

Current Failing Grades: a dummy variable coded 1 if accepted into TRIO based on failing grades for the current semester, 0 otherwise.

Tables

Table 2: Dependent Variable: Cumulative GPA, All Years

Independent Variables	First Year OLS	Independent Variables	First Year OLS
Age	0.0122**	< Half Time	-0.3696**
	-0.003		-0.084
Female	0.1644 **	Freshman, never attended	-0.3269**
	-0.038		-0.057
Asian	0.0044	Freshman, previous courses	-0.4286**
	-0.078		-0.058
Black	-0.2246**	Bad High School Indicators	0.0077
	-0.047		-0.058
Latino	-0.2279**	Predicative Indicator	0.3154**
	-0.073		-0.052
> 1 race report	-0.1042	Low College GPA	-0.2967**
	-0.079		-0.078
Full Time	0.1996**	Current Failing grades	0.1558**
	-0.049		-0.078
Constant	2.12**	Observations	1274
	-0.105	R-Squared	0.2404

* significant at 5%; ** significant at 1%

Table 3 Dependent Variable: Cumulative GPA, Full Sample, Pooled OLS

	Pooled OLS-Full Sample	Pooled OLS-Full Sample, If Male	Pooled OLS-Full Sample, If Not White or Asian	Pooled OLS-Full Sample, If Black	Pooled OLS-Full Sample, If Disabled
ProgYear	0.022**	0.01	0.030*	0.029	0.050*
	[0.007]	[0.010]	[0.012]	[0.016]	[0.022]
Female	-0.093**	_	-0.076	-0.107*	-0.082
	[0.025]		[0.041]	[0.054]	[0.064]
Asian	-0.04	-0.023	_	_	-0.814**
	[0.052]	[0.064]			[0.139]
Black	-0.327**	-0.325**	-0.172**	_	-0.015
	[0.031]	[0.039]	[0.054]		[0.117]
Latino	-0.173**	-0.222**	-0.012	_	-0.375**
	[0.050]	[0.070]	[0.067]		[0.112]
other	-0.152**	-0.154*	_	_	-0.174**
	[0.051]	[0.065]			[0.053]
badhs	0.267	0.116	-0.124	-0.085	_
	[0.149]	[0.142]	[0.139]	[0.223]	
ColGPA	-0.300**	-0.314**	-0.234**	-0.242**	-0.350**
	[0.029]	[0.039]	[0.049]	[0.058]	[0.075]
Cur Fail Grades	-0.027	-0.019	-0.093	-0.152	0.028
	[0.051]	[0.068]	[0.083]	[0.104]	[0.172]
< Half Time	-0.415**	-0.389**	-0.400**	-0.396**	-0.285**
	[0.027]	[0.036]	[0.046]	[0.058]	[0.076]
Fresh	-0.251**	-0.273**	-0.248**	-0.259**	-0.304**
	[0.050]	[0.065]	[0.079]	[0.094]	[0.114]
Sen	0.337**	0.365**	0.344**	0.354**	0.300**
	[0.032]	[0.043]	[0.055]	[0.068]	[0.075]
Constant	2.723**	2.731**	2.500**	2.331**	2.451**
	[0.039]	[0.049]	[0.076]	[0.079]	[0.087]
Observations	4685	2838	1661	1028	519
R-squared	0.19	0.2	0.17	0.19	0.19

Robust standard errors in brackets

* significant at 5%; ** significant at 1%

Table 4 Dependent Variable: Cumulative GPA, Full Sample, Fixed Effect

	Fixed- Full Sample	Fixed- Full Sample, If Male	Fixed- Full Sample, If Not White or Asian	Fixed- Full Sample, If Black	Fixed- Full Sample, If Disabled
ProgYear	0.041**	0.040**	0.065**	0.071**	0.045*
	[0.005]	[0.007]	[0.008]	[0.010]	[0.020]
Constant	2.395**	2.435**	2.170**	2.076**	2.383**
	[0.014]	[0.018]	[0.024]	[0.028]	[0.051]
Observations	4685	2838	1661	1028	519
Number of ID	2118	1292	760	461	201
R-squared	0.02	0.02	0.06	0.09	0.02

Standard errors in brackets

* significant at 5%; ** significant at 1%

Table 5 Dependent Variable: Cumulative GPA, Less than Five Years, Pooled Effect					
	Pooled-yrs<=4	Pooled-yrs<=4 If Male	Pooled-yrs<=4 If Not White/Asian	Pooled-yrs<=4 If Black	Pooled-yrs<=4 If Disabled
ProgYear	0.051**	0.040*	0.067**	0.054	0.028
	[0.012]	[0.016]	[0.021]	[0.028]	[0.029]
Female	-0.102**	_	-0.101*	-0.109	-0.055
	[0.028]		[0.051]	[0.070]	[0.066]
Asian	0.018	0.045	_	_	-0.779**
	[0.056]	[0.068]			[0.148]
Black	-0.319**	-0.322**	-0.266**	_	0.055
	[0.036]	[0.043]	[0.062]		[0.134]
Latino	-0.206**	-0.266**	-0.14	_	-0.408*
	[0.056]	[0.074]	[0.076]		[0.160]
Other	-0.047	-0.001	_	_	-0.180**
	[0.058]	[0.070]			[0.057]
badhs	0.156	0.026	-0.194	0.11	_
	[0.159]	[0.148]	[0.238]	[0.105]	
ColGPA	-0.265**	-0.264**	-0.192**	-0.220**	-0.292**
	[0.032]	[0.042]	[0.058]	[0.070]	[0.069]
Cur Fail Grades	0.008	0.016	-0.067	-0.115	0.015
	[0.059]	[0.081]	[0.090]	[0.110]	[0.188]
< Half Time	-0.447**	-0.413**	-0.414**	-0.380**	-0.310**
	[0.031]	[0.039]	[0.054]	[0.070]	[0.087]
Fresh	-0.289**	-0.316**	-0.333**	-0.335**	-0.337**
	[0.058]	[0.073]	[0.096]	[0.115]	[0.127]
Sen	0.280**	0.295**	0.273**	0.310**	0.349**
	[0.036]	[0.048]	[0.070]	[0.091]	[0.088]
Constant	2.756**	2.759**	2.652**	2.381**	2.464**
	[0.042]	[0.053]	[0.086]	[0.092]	[0.095]
Observations	3416	2131	1098	670	425
R-squared	0.21	0.21	0.19	0.19	0.21

Robust standard errors in brackets

* significant at 5%; ** significant at 1%

Table 6 Dependent Variable: Cumulative GPA, Less than Five Years, Fixed Effect					
	Fixed-yrs<=4	Fixed-yrs<=4 If Male	Fixed-yrs<=4 If Not White/Asian	Fixed-yrs<=4 If Black	Fixed-yrs<=4 If Disabled
ProgYear	0.042**	0.044**	0.058**	0.065**	0.035
	[0.009]	[0.012]	[0.017]	[0.019]	[0.026]
Constant	2.509**	2.541**	2.338**	2.247**	2.426**
	[0.019]	[0.025]	[0.034]	[0.039]	[0.058]
Observations	3416	2131	1098	670	425
Number of ID	1660	1030	544	325	176
R-squared	0.01	0.01	0.02	0.03	0.01

Standard errors in brackets

* significant at 5%; ** significant at 1%