Decomposing the Racial Gap in STEM Major Attrition: A Course-Level Investigation

Matthew D. Baird, Moshe Buchinsky, and Veronica Sovero

RAND Labor & Population

WR-1171

November 2016

This paper series made possible by the NIA funded RAND Center for the Study of Aging (P30AG012815) and the RAND Labor and Population Unit.

RAND working papers are intended to share researchers' latest findings and to solicit informal peer review. They have been approved for circulation by RAND Labor and Population but have not been formally edited or peer reviewed. Unless otherwise indicated, working papers can be quoted and cited without permission of the author, provided the source is clearly referred to as a working paper. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors. RAND® is a registered trademark.

For more information on this publication, visit www.rand.org/pubs/working_papers/WR1171.html

Published by the RAND Corporation, Santa Monica, Calif.

© Copyright 2016 RAND Corporation

RAND® is a registered trademark

Limited Print and Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law. This representation of RAND intellectual property is provided for noncommercial use only. Unauthorized posting of this publication online is prohibited. Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Permission is required from RAND to reproduce, or reuse in another form, any of its research documents for commercial use. For information on reprint and linking permissions, please visit www.rand.org/pubs/permissions.html.

The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest.

RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

Support RAND

Make a tax-deductible charitable contribution at
www.rand.org/giving/contribute

www.rand.org

Preliminary and Incomplete
Please do not quote without permission

Decomposing the Racial Gap in STEM Major Attrition: a Course-Level Investigation

Matthew D. Baird¹

Moshe Buchinsky²

Veronica Sovero³

October 23, 2016

¹RAND Corporation

²UCLA and NBER

³Wake Forest University

Abstract

This paper examines differences in STEM retention between minority and non-minority undergraduate students. To do so, we use detailed student records of a student's courses, grades, and current major for every term the student was enrolled in a large public university. To examine the role of ability in the switching decision and timing, we estimate STEM and non-STEM ability, and then compare the joint distribution of students who switch out of STEM versus STEM stayers. Students with relatively greater non-STEM ability are more likely to switch out of STEM, but ability cannot completely account for the differences in switching patterns for Hispanic and Black students. In fact, Black and Hispanic students are more likely to persist in STEM after ability is taken into account. We also find evidence of switching behavior that appears motivated by a preference for graduation within four years.

1 Introduction

There has been growing concern in the United States about the steady decrease in student enrollment in the Science, Technology, Engineering, and Math (STEM) majors. The ability to increase the fraction of individuals seeking education in STEM major has been viewed by many (e.g. Carnevale, Smith, and Melton (2011)) as a vital element necessary for the growth of the U.S. economy, especially relative to the growing technological nations in Europe and the Far East. The President’s Council of Advisors on Science and Technology suggested in their annual report that the graduation rate from STEM majors must increase dramatically to be able to meet the growing demands for STEM professionals, and in particular for the key growing nanotechnological industries.⁴

President Obama’s statement highlights the importance with which the federal government regards STEM fields:

“We’ll reward schools that develop new partnerships with colleges and employers, and create classes that focus on science, technology, engineering, and math—the skills today’s employers are looking for to fill jobs right now and in the future.”

Increasing STEM graduation rates can also serve as an important policy tool in increasing socioeconomic mobility for minority groups. Individuals graduating from STEM majors have substantial earnings premium over their counterparts in other fields, a premium that has increased enormously over time (e.g. Gemici and Wiswall (2014)). Despite the high earnings premium, there is a persistent under representation of several minority groups in STEM fields. The U.S. Department of Education (see report by Chen and Soldner (2013)) reports vastly different attrition rates for individuals with different family background. More importantly the two types of individuals who leave STEM majors (those who switched to different majors and those who dropped out of college without earning a degree or certificate) have very different characteristics. Of those in STEM majors, more women than men switched to a non-STEM field, 32.4% and 25.5%, respectively. Moreover, 23.7% of men dropped out of college altogether, while only 14.2% of women dropped out of school. The differences for the various racial groups are even more staggering. The dropout rate for

⁴The U.S. Department of Education projections for job growth in STEM fields over the period from 2010 to 2020 far exceed the projected growth in all occupations. For all occupations the project growth is 14%, while for all STEM jobs the project rates of growth are more than 16%. For medical scientists the projected growth is 36%, and growth of jobs for biomedical engineers is 62%.

Asian is only 9.%, while they are much larger for any other group: 19.8% for whites, 23.1% for Hispanics, and 29.3% for African American.⁵ Indeed, the National Science Foundation (NSF), the National Institutes of Health (NIH), the Department of Education, among other federal agencies, devote considerable amount of resources for colleges across the U.S. in an attempt to increase the representation of such underrepresented groups in STEM majors.

In this paper, we use a rich dataset from a large selective public university to analyze STEM retention patterns over the years 1980 through 2006. We have a uniquely detailed dataset that allows us to observe the individuals' career path at the university, from the point at which they are admitted until they graduate (or, alternatively dropout). We use course and term level data to characterize the various routes to career paths that individuals with varying characteristics take. Of particular importance is explaining the STEM retention gap between minorities and non-minorities. Examining the raw differences in STEM retention by race reveal that minority students are more likely to switch out of STEM, but they do so later in their student career. This is puzzling because minority STEM students have lower SAT math scores and high school GPA than their non-minority counterparts, factors which would make them less likely to successfully graduate as a STEM major. To examine the role of ability in a student's decision to switch out of STEM, we recover a student's STEM and non-STEM ability from GPA regressions. The GPA regressions include a rich set of course level controls to account for factors that may affect a student's academic performance. In particular, we control for various peer effects such as average ability and average demographic characteristics.

After we recover student's STEM and non-STEM ability, we examine how relative ability in a major plays a role in a student's graduation outcomes. We find evidence that students with relatively higher non-STEM ability are more likely to graduate because they are more likely to switch to a non-STEM major. It also appears that minority students have a stronger preference for STEM majors relative to non-STEM majors. Conditional on being admitted as a STEM major, minority students are less likely to graduate with a non-STEM major and more likely to not complete any major within four years. A term by term switching regression model reveals that minority students are less likely to leave STEM. Controlling for unobserved ability significantly reduces but does not completely eliminate the STEM graduation gap between minority and non-minority students.

We also investigate whether a student's socioeconomic background plays a role in the

⁵STEM attrition rates also vary considerably by other students's characteristics, namely parents education, parents income, and suitable academic preparation in high school. (See Table 2 of the report).

switching decision. Our findings suggest that household income influences both a student's budget constraint during college as well as their relative preferences for graduating as a STEM major versus other majors. If STEM majors are more difficult to complete within a four year period, financially constrained students may be more likely to switch out of STEM in order to graduate on time. We document a higher switching rate out of the STEM majors for students who graduated from any major within four years versus five years. Similarly, we find that the minority graduation gap is smaller across 5 year graduation rates compared to 4 year graduation rates after controlling for ability and first year grades. When we allow for unobserved types in the STEM population in our switching regression we find a substantial fraction of students who appear to be switching in order to graduate within four years. We also find a direct positive relationship between income and the probability of graduating as a non-STEM major. This result points to differences in relative preferences for graduating STEM versus non-STEM by socioeconomic status.

Our paper builds upon the literature regarding college major choice. There are a number of papers that have utilized course level data to examine major switching patterns by gender, by race, or in general (Rask (2010); Arcidiacono, Aucejo, and Spenner (2012); Ost (2009)). Ost (2009) examines the effects of peers and grades on major choice and examines the timing of a student's decision to switch. Arcidiacono, Aucejo, and Spenner (2012) examine changes in the black-white achievement gap over time while controlling for unobserved ability. The authors find that black students leaving STEM majors (which have higher grading standards than other majors) explains most of the GPA convergence. They also argue that black-white differences in switching behavior can be completely accounted for by differences in academic background. Due to the small sample size, the authors are not able to examine switching behavior within the broad category of STEM majors. In our paper we document large variation in ability and switching behavior among students majoring in life sciences, physical sciences, and engineering.

Arcidiacono, Aucejo and Hotz (2013) use data from the University of California (UC) system to examine differences in STEM retention across several public universities. They obtained data from all of the UC campuses (our data comes from only one public university). The data they use is censored in several ways in which our data is not: cohorts are grouped into three-year intervals, there is one broad category of science majors, there is no information on gender, SAT scores are not separated into math and verbal scores, and no course level data. While our data come from only one university, the dataset provide much more detailed information, both about the students and the classes that they take. In

particular, we have the data that appeared on the individual’s application, which includes some basic individual characteristics and information on their high school performance (high school GPA, and a variety of scores on some standardized tests). More importantly, the data set is a census of all students at the university. Therefore we can construct exact measures of performance for each class taken at the university. We thus know the exact composition of each class in term of all the reported characteristics of the individuals.

While Arcidiacono et al. (2012) allow for unobserved student ability through a time varying fixed effect, they do not separate ability into math and verbal components. One would expect the returns to math and verbal abilities to vary across majors. Arcidiacono (2004) estimates a dynamic model of college and major choice where individuals vary in their unobserved major specific ability. Arcidiacono does not have course level data, and only observes changes in major over two broad time periods in a student’s academic career. Our data allows us to precisely identify a student’s performance and switching behavior every quarter the student is enrolled, while controlling for their unobserved math and verbal ability.

The rest of the paper is organized as follows. Section 2 introduces the data and provides some descriptive analysis, Section 3 outlines the estimation strategy, Section 4 presents estimation results, and Section 5 concludes.

2 Data and Descriptive Statistics

To provide context to the subsequent econometric analysis, we present some stylized facts regarding STEM major retention. The data we use consists of the registrar records for all undergraduate students who matriculated at the university between 1980 and 2005. This makes it possible to track students as they progress in college, and in particular to analyze the influence of a particular class and set of classmates on performance on subsequent classes. We classify engineering, life sciences, and physical majors as STEM, and humanities and social Science majors as non-STEM.

For each student, we have the following:

1. A complete record of every course taken, including year, term, meeting times, and grade. On average, each student takes about 40 classes in his or her career, so we have more than 4,000,000 student-course observations.
2. The student’s declared major every quarter, as well as their degree outcomes.
3. Background characteristics such as gender, race, high school grade point average

(GPA), the zip code of last school attended, and performance on standardized exams, including SAT math and verbal scores.

A drawback of our dataset is that it does not include information on the student’s financial aid status or family income. We use zip code level data on median household income from the 2000 Census as a proxy for household income.

We begin by comparing minority and non-minority students who enter the university as STEM majors (Table 1). We use the term minority to refer to under-represented minorities (Black and Hispanic students) and non-minority to refer to White and Asian students. Relative to non-minorities, minority students are 18 percentage points less likely to graduate from STEM fields within four years and are 22 percentage points less likely to graduate within five years. Minorities are also 18 percentage points more likely to switch out of STEM relative to non-minorities. The difference in outcomes could partially be explained by differences in math preparation and performance in their first year. On average, minority students have a math SAT score of 561 and first year STEM GPA of 2.4, whereas non-minority students have an average math SAT score of 667 and first year STEM GPA of 2.9. We also observe smaller differences between minorities and non-minorities in verbal ability (SAT verbal scores and first year non-STEM GPA), which could be pulling minorities out of STEM fields. Taken together, this suggests that students consider their relative and absolute math and verbal skills when deciding whether or not to stay in STEM.

We also observe differences in the specific STEM fields that minority and non-minority students choose as their initial major: minority students have a relatively larger presence in life sciences and smaller presence in engineering.⁶ Table 2 presents tabulations of the transition between a student’s initial major and degree major, separately by minority status. For both minorities and non-minorities, the vast majority of all STEM graduates were students who initially declared a STEM major. Because of this, we focus our attention on STEM retention in subsequent analysis.

It appears that low minority retention is unique to STEM fields. For example, the retention gap is only 3 percent for social science majors versus 20 percent for physical science majors. While retention is typically lower for STEM fields, engineering is a notable exception for its high retention rates and smaller minority gap. The engineering school

⁶At UCLA, life and physical sciences are part of the College of Letters and Sciences, while the engineering school is separate from the college. The College of Letters and Science and the engineering school have separate admissions, so students apply to either the College of Letters and Science or the School of Engineering. As a result, it’s likely that the admissions criteria are more rigorous for students admitted through the engineering school.

at the university retains 83 percent of non-minorities and 73 percent of minority students who initially declare an engineering major. The success of the engineering school is likely due to a more selective admissions policy, which means that students accepted into the major are more uniformly prepared. We also observe very little movement between STEM fields. The overwhelming majority of STEM transitions occur out of STEM versus within STEM, with most life and physical science students switching to a social sciences major. Engineering students are equally likely to transition to another STEM field as they are to social sciences.

2.1 Minority Time Trends

Figure 1 plots the distribution of the admitted major for minority (Panel A) and non-minority students (Panel B) by their admitted major and the year in which they were admitted. It is important to note that beginning in 1998, race-conscious admissions were banned. Previous studies have examined whether affirmative action bans lead to changes in the college choice (Card and Krueger (2005); Antonovics and Backes (2013); Antonovics and Sander (2013)), acceptance rates (Backes (2012)), and completion rates (Hinrichs (2012); Hinrichs (2014))), but it is unclear whether minority students changed their choice of major after the ban. Compared to non-minority students, there is a notable drop in the fraction of minority students entering as an undeclared major in 1998. Over the entire sample period, there appears to be a gradual growth in the fraction of minorities declaring social sciences and life sciences relative to other majors. For non-minorities, the number of life science majors has increased while the number of physical science majors has decreased.

Figure 2 plots the distribution of minority students (Panel A) and non-minority students (Panel B) by their degree major and the year in which they were admitted. For minority students, there does not appear to be any dramatic shift in graduating major after affirmative action was banned. The fraction of minority students graduating as an engineering or physical science major has remained small over the entire sample period. We also observe that there has been an increase in life science graduates for both minority and non-minority students after 1990.

Figure 3 plots how the fraction of minority students relative to non-minority students by admit major (Panel A) has changed over time. The overall fraction of students who were minorities across all of the majors (blue dashed line) had climbed to an overall high of 22 percent by the entering class of 1988, but this fraction dropped to 12 percent with the

entering class of 1998. Minority students are consistently overrepresented in the undeclared and social science majors prior to 1998; after 1998, minorities seem to substitute away from undeclared and into social sciences. A stronger sorting pattern is evident when examining the fraction of minority students by degree major over time (Figure 3 Panel B).

In addition to changing the fraction of minority students admitted into STEM fields, it is possible that the change in admissions policy led to a change in the ability composition of admitted students. Figure 4 shows the average SAT math score of minority students and non-minority students by admitted major. For non-minority students, there is a clear sorting into initial major by math ability; admitted STEM majors have the highest average math SAT scores of all the majors. On the other hand, there is less of a separation in ability for minority students. Engineering minority students consistently have higher math SAT scores, but there is little difference between life sciences, social sciences, and humanities majors. Despite having race-conscious admissions in place, the average SAT math score of minority students admitted as a STEM major had been increasing prior to 1998, and appears to have leveled off post-1998. To see whether students are sorted into their major based on their verbal ability, Figure 5 presents the average SAT verbal scores by admitted major for minority students (panel A) and non-minority students (panel B). Engineering minority students have the highest verbal scores, possibly because math and verbal ability is correlated. Overall, there is less of a separation between STEM and non-STEM majors based on verbal ability, which suggests that students consider their absolute math ability more than their relative abilities when selecting a major.

Figures 6 and 7 show mean SAT math and verbal scores by graduation major, respectively. By the time students graduate, both minorities and non-minorities are more strongly sorted into their major based on math ability (Figure 6). With respect to verbal ability, non-minority humanities majors have the highest verbal ability, but that is the only group that has a clear separation from the other majors (Figure 7).

2.2 Switching Patterns by Race

Because we observe the quarter in which a student switched out of STEM, we next examine the timing of a student's decision to switch out of STEM. Figure 7 plots STEM major retention for students over the course of their career. We consider more narrowly defined fields of study (Life Sciences, Physical Sciences, Engineering), separately by race. We first consider major retention for students who graduated within four years (first column of

Figure 8). Black and Hispanic students are more likely to switch out of STEM relative to White and Asian students across all of the subfields, but there is less of a difference in switching behavior for engineering majors. This is likely due to the engineering school having a more rigorous admissions standard for admitted students.

Interestingly, the differences in major retention appear in the later years of a student's career. Major retention looks similar across all of the groups of students for the first 6 terms and diverges afterwards. Because the transition tends to happen later in a student's career, it is important to consider factors beyond a student's STEM experiences in their first year or course, as is common in the literature. The timing of the transition is likely a function of the institutional setting. At this university, major requirements are divided into "preparation for the major" and "the major." Preparation for the STEM major typically consists of lower division courses across a broad range of departments, so most STEM majors are taking similar courses at this stage. After completing preparation for the major, the students advance to taking upper division courses in their major department. Most departments have minimum GPA requirements in order to take upper division classes, and some require a formal petition to transition from pre-major to major status.

When comparing retention for four year graduates versus five year graduates, we see that Black and Hispanic students appear to leave at higher rates as they approach the on-time graduation (12 terms) deadline. This pattern suggests that minority students face a trade-off between on-time graduation and graduating as a STEM major. With more time to complete their major, there is a smaller gap in major retention between minority and non-minority students. Engineering in particular has very high retention rates for all students.

2.3 Switching Patterns and Ability

We next examine how ability plays a role in the timing of the decision to leave STEM. Figure 9 tracks the Stem GPA of eventual STEM dropouts during their time as a STEM major for minority (panel A) and non-minority (panel B) students. There is a sizable GPA gap between STEM dropouts and non-STEM dropouts, which narrows over time for both minority and non-minority students. This could suggest that the relatively stronger dropouts wait longer to leave STEM because they are closer to the ability threshold for being able to graduate as a STEM major. On the other hand, we showed earlier that minority students are more likely to switch out of STEM later in their career, which could

suggest that relatively weaker students wait longer to switch out of STEM.

It is also possible that students consider their relative strengths when deciding whether or not to switch out of STEM, and when. A weak STEM student may be an even weaker non-STEM student, which would make switching out of STEM less appealing compared to a student with stronger non-STEM ability. Figure 10 plots the mean math and verbal SAT scores for STEM dropouts by their dropout quarter. In absolute terms, students with higher math ability tend to switch out of STEM earlier in their careers. In relative terms, the gap between math and verbal ability is also smaller for early STEM dropouts. Students who wait to switch out of STEM have weaker outside options, but the delay in their eventual switch could come at the cost of decreasing the likelihood of on-time graduation. Because weaker students tend to switch later in their career, we will need further analysis to separate the two effects.

How do students learn about their relative ability? Grades in STEM and non-STEM should serve as a signal of a student's relative strengths, which could help students choose their major based on their comparative advantage. On the other hand, it is well documented that differences in grading standards have distorted the informational content of STEM versus non-STEM grades (Sabot and Wakeman-Linn 1991). Figure 10 shows the changes in average grades by field over time for minority and non-minorities students. Humanities and social science grades have been steadily increasing over time, whereas engineering and life sciences experienced limited grade inflation until the late 1990's.

3 Model

Motivated by the observed data trends, we explore a simplified model which will inform the empirical specification. We assume that students make a nested decision: in the outer nest, they decide either to continue in college that term, to graduate in a major if they meet the graduation requirements, or to drop out. If they decide to continue in college that term, they additionally choose which major to be in. Here we present a simplified model that suggests the underlying dynamics of interest. Period utility for student i of minority/majority group g in term t choosing major m is given by:

$$U_{igtm} = f_{gtm}(\alpha_i, GPA_{it}, P_t | X_i) + \varepsilon_{igtm},$$

where α_i is a vector of ability; in our application, we assume that it is two-dimensional, with STEM and non-STEM ability. GPA_{it} is a vector of the students' GPA by term t in

the different major groupings. P_t is a vector of class characteristics, such as the distribution of students in each major by ethnicity, gender, and ability. X_i is a vector of individual characteristics, including financial situations which may induce lower utility for staying in school instead of graduating. ε_{igtm} are random utility shocks.

We include both GPA and ability. We hypothesize that GPA, which varies over time, plays different roles in the utility of the student. First, it serves as an update to the student of their own ability, as they observe α_i with noise. Second, it influences their ability to graduate in a given major; low GPA in a major implies a need to repeat classes and thus longer time until graduation, as well as strict GPA requirements that majors may have for student matriculation. Third, it influences the enjoyment of the student in the major. Finally, it also affects post-graduation utility, as future wages may depend on the GPA in the major from which the student graduates.

Ability has both a direct effect on utility through enjoyment and an indirect effect through GPA and future GPA and thus all of the other effects described above. The grade function is described below.

Utility post-graduation is given by an average utility for graduating in that major with the GPA. Earlier graduation means fewer expected periods to earn the post-graduation utility, which includes wages, but may come at a cost of exiting in a major that yields lower utility.

In our model, we restrict our sample to students that start in STEM majors so as to focus on the STEM retention and matriculation outcomes. Given our model, students may switch out of STEM for several reasons:

1. Ability (both STEM and non-STEM), which conditional on entry into a given quarter affects both future GPA (see below), enjoyment, and differential direct effects on utility in that major.
2. Updated beliefs about their ability from GPA, both in STEM and non-STEM. Students may use this information to adjust their expectations about current and future utility in the majors, post-graduation utility from that major (both from matriculation GPA in the major affecting wages and job placement and overall satisfaction with employment that results from the graduation major),
3. Updated beliefs about their remaining time and ability to graduate in each major. Income may play a role in how students respond to this information. In particular, students who are financial constrained may have a stronger incentive to switch when faced with a

lower likelihood of an on time graduation in a STEM field.

4. Classroom demographics by major changing utility
5. Random utility shocks

We will be interested in several elements of the model. For example, we will examine the multivariate density of STEM and non-STEM ability, and how they differ by those that stay in STEM majors and those that switch out. We will also estimate how the different factors affect retention and graduation probabilities, as explained below. This model encourages the following research questions, which we examine in this paper.

1. How do the STEM and non-STEM abilities of students compare for STEM stayers and quitters, and do those distributions differ by gender, race, and major?
2. How do minorities and women react differently to signals and other factors in deciding transitional outcomes (the timing and incidence of switching out of STEM majors) and terminal outcomes (graduation status and major)?
3. Are switching decisions influenced by socioeconomic background?
4. Is there evidence of different types of students (unobserved), and if so, what can be inferred about these types based on the regression coefficients?
5. Are STEM and non-STEM abilities and GPAs both important (relative ability, etc. etc.)
6. After controlling for unobservables, what portion of the racial and gender gap in retention still remains?
7. Are there additional information and gains to estimating the model at the course-level rather than the terminal outcome, and if so, what?

4 Estimation Strategy

The descriptive statistics suggest that a student's STEM ability relative to non-STEM ability influences both a student's decision to switch out of STEM, as well as the timing of the switch. Therefore, it is important to allow for heterogeneity in the decision-making process to leave STEM, by both unobservables (ability, unobserved type) and observables (race, STEM subfield). We begin our analysis by estimating a student's unobserved ability in both STEM and non-STEM through course level grade regressions. After we recover the fixed effects from the STEM subject area and non-STEM subject area grade regressions, we examine how ability plays a role in various outcomes such as STEM retention and on-time graduation. After controlling for ability, we can see if the retention gap remains

between minorities and non-minorities.

4.1 Grade Regressions

We assume that student i 's grade in term t , and course j , and subject area $k = (STEM, \text{non-STEM})$ can be decomposed as follows:

$$Grade_{ijt}^k = \alpha_i^k + \theta_j^k + \beta^k X_{it} + \sigma^k P_{jt} + \gamma_t^k + u_{ijt}^k$$

By estimating separate grade regressions for STEM and non-STEM subject areas across all undergraduate course offerings, we can recover the student's subject specific fixed effect α_i^k . We interpret α_i^k as the time invariant ability of the student in the subject area. Time varying student controls X_{it} include the student's current major, whether the course is in the same subject area as their major, and the number of classes a student has taken in each subject area (subject specific experience).

Because grading standards and content vary across courses and over time, we include course-year fixed effects θ_j^k . Within a specific course, a student's performance can be affected by time-varying characteristics of a course (P_{jt}) such as peer quality and class size. Controls for class composition include mean SAT math scores, mean SAT verbal scores, mean high school GPA, class size, and the fraction of students in the class who are male, black, white, asian, and hispanic. Quarter fixed effects are included to control for differences in grading practices across terms. For example, summer courses may have lower grading standards relative to classes offered during the regular school year.

A potential threat to identification of the ability measures arises if students select into courses based on ability. While all STEM majors take similar lower division courses, students with higher ability are more likely to take upper division courses where they face stronger peers to compete against. As a result, our estimates of student ability may be biased downwards for relatively higher ability students. To address this selection issue, we can restrict our estimates of the student fixed effects on lower division courses. Estimated fixed effects from these courses would be expected to have less selection bias relative to estimates which include upper division courses. Our estimates of fixed effects using only lower division courses were virtually unchanged compared to when we used all undergraduate courses (available upon request), which suggests that the selection bias is not very strong. Ultimately, we chose to keep the upper level observations to increase the precision of our estimates.

4.2 Student Terminal Outcome

After we recover the STEM and non-STEM fixed effects from the grade regressions, we can analyze how ability plays a role in STEM retention and timely graduation. We expect to find that students with higher STEM ability are more likely to graduate as a STEM major, while students with higher non-STEM ability are more likely to switch out and graduate as a non-STEM major. Due to the small number of students who switch into STEM, we restrict our analysis to students who begin their career as a STEM major.

The probability of outcome $j = \{\text{grad STEM, grad non-STEM, dropout}\}$ for student i is:

$$\pi_{ij} = \frac{\exp\{\eta_{ij}\}}{\sum_{k=1}^3 \exp\{\eta_{ik}\}}$$

where:

$$\eta_{ij} = \hat{\alpha}_i^{STEM} + \hat{\alpha}_i^{nonSTEM} + \beta_1 X_i + \beta_2 F_i + \gamma_t$$

Student level controls include the student's STEM and non-STEM ability (the fixed effects from the STEM and non-STEM regressions), as well as background characteristics X_i such as race, gender, and whether the student was admitted as a life sciences, physical sciences, or engineering major. We also include controls F_i which characterize the student's first year performance (STEM GPA, non STEM GPA) and peers in their first year STEM courses. Time trends are captured by year fixed effects γ_t .

4.3 Student Transitional Outcomes

Because we observe the precise timing of a student's decision to switch out of STEM, we can investigate time varying factors which contribute to the switching decision. In particular, we estimate the likelihood that a student switches out of STEM in a given term is as follows:

$$Pr(d_{it} = 1) = \hat{\alpha}_i^{STEM} + \hat{\alpha}_i^{nonSTEM} + \beta_3 X_{it} + \beta_4 P_{it} + \beta_5 GPA_{it}^{STEM} + \beta_6 GPA_{it}^{nonSTEM} + \gamma_t$$

Where d_{it} is equal to 1 if the student switches out of STEM or drops out of school the following term, 0 otherwise. We predict that students with higher STEM ability are more likely to persist in STEM, but having a higher non-STEM ability is likely to pull students away from STEM. Time varying student controls (X_{it}) include controls for the student's

current course load. We expect that taking any non-STEM courses suggests the student is considering switching. Conversely, students may be less inclined to switch out of STEM if they are currently taking upper division STEM courses because they are closer to completing the major. P_{it} captures peer effects from the student's classmates in term t . By including the composition of a student's STEM classes, we can examine whether or not peers play a role in the decision to remain in STEM, separate from the effect of peers on grades. We also include the student's STEM and non-STEM GPA to examine whether or not student's update their beliefs about their abilities based on their grades. Term fixed effects (γ_t) are included to control for how long the student has persisted in STEM.

The likelihood function for the i th individual is:

$$\mathcal{L}_i(\theta) = \prod_{t=1}^{T_i} \mathcal{L}_{it}(\theta) = \prod_{t=1}^{T_i} \prod_{l=1}^M p_l^s(z_{it}; \theta)^{d_{lit}},$$

where $d_{lit} = 1$ if choice l was taken, and T_i is the number of periods the i th individual is observed in the data. We censor observations after the student leaves STEM.

Alternate specifications of the switching regression are estimated to allow for heterogeneous effects by race. For example, minority students may respond differently to their GPA or peers. Because of this, we estimate the switching regression separately for White, Asian, and minority students. Finally, we estimate a switching regression that assumes there are several unobserved students types in the population. More specifically, we assume that there are J types of individuals in the population, with probabilities of $\pi_1, \dots, \pi_J$ (with $\sum_{j=1}^J \pi_j = 1$). In this specification, we estimate the distribution of types in the population and allow for the switching coefficients to vary by student type.

Then, the likelihood function for the i th individual is given by

$$\mathcal{L}_i(\theta, \tau_1, \dots, \tau_J) = \sum_{j_\tau=1}^J \pi_{j_\tau} \mathcal{L}_i(\theta; \tau_{j_\tau}),$$

where $\tau_1, \dots, \tau_J$, include all the type-specific parameters, including $\pi_1, \dots, \pi_J$.

The overall log likelihood is then

$$l(\theta, \tau_1, \dots, \tau_J) = \sum_{i=1}^N \ln \mathcal{L}_i(\theta, \tau_1, \dots, \tau_J) = \sum_{i=1}^N \ln \left[\sum_{j_\tau=1}^J \pi_{j_\tau} \mathcal{L}_i(\theta; \tau_{j_\tau}) \right].$$

5 Estimation Results

5.1 Grade Regressions

Table 3 reports grade regression results for STEM subject area courses. Comparison of Columns 1 and 2 (no course fixed versus course fixed effects) show that minority students have lower expected grades relative to white students, even after controlling for differences in grading practices across courses. Results with student fixed effects are presented in Columns 3, and results with student and course fixed effects are presented in Column 4. The differences in the estimated coefficients illustrate the importance of including a course-year fixed effect. For example, if we do not include course fixed effects, it appears as though students taking a course in their major perform worse relative to other majors. After course fixed effects are controlled for, it is clear that students earn higher grades in their major courses relative to other majors. This suggests that students choose their major based on their ability in the subject area. It also appears that competition from high quality peers (measured by average SAT math and verbal scores of classmates) tends to have modest negative effect on a student’s expected STEM grade, even after controlling for course year fixed effects.

Table 4 reports grade regression results for non-STEM subject area courses. On average, minority students have lower expected grades in non-STEM subject areas relative to white students (Column 1), although the difference is not as large as in the STEM grade regressions. We also see similar peer quality effects in the non-STEM grade regressions compared to the STEM grade regressions, although now verbal SAT scores have larger effect on expected grades than math SAT scores. Interestingly, we see opposite effects of class size on expected grades in STEM versus non-STEM classes. After controlling for student and course fixed effects, an increase in class size increases expected grades in STEM courses, and lowers expected grades in non-STEM courses.

5.2 STEM and non-STEM ability

We next turn to our estimates of the student fixed effects associated with the STEM courses, denoted $\hat{\alpha}_{ST}$, and for the non-STEM courses, denoted $\hat{\alpha}_{NS}$, for each individual in our sample. The fixed effects are recovered from the student and course-year fixed effects specification. We first examine is how observable characteristics are correlated with the estimated fixed effects. Table 5 presents regression results for the STEM fixed effect

across all students (Column 1). On average, minority students have lower STEM ability than white and asian students (Column 1), even after controlling for factors such as SAT math scores and high school GPA. Columns 2-4 of Table 5 estimate the STEM fixed effect separately by race. Math SAT scores have similar effects across all student groups. On the other hand, high school GPA appears to be a stronger predictor of STEM ability for minority students versus White and Asian students. We also observe a stronger correlation between median household income in the student’s zip code and STEM ability for minority students. This may be due to factors such as larger variance in school quality by income for minorities versus non-minorities. Table 6 presents regression results for the non-STEM fixed effect. Black and Hispanic students have lower non-STEM ability, all else equal. SAT verbal scores are more correlated with non-STEM ability for black and hispanic students, but the overall regression model has less explanatory power relative to white and asian students. The R-squared of the regression models are also smaller relative to the STEM ability regressions, which suggests that there are other unobserved factors to consider when trying to predict who will do well in non-STEM majors.

5.3 The Distribution of Abilities in STEM and Non-STEM Courses

Having estimated the course level fixed-effect model also makes it possible to carefully examine the joint distribution of the estimated fixed effects, denoted $\hat{\alpha} = (\hat{\alpha}_{ST}, \hat{\alpha}_{NS})$, from the STEM courses and from the non-STEM courses, taken by those who were admitted to the university as STEM students. We estimate non-parametrically the joint distribution for $\hat{\alpha}$, denoted $\hat{F}_{ST,NS}$. To do so we use a bivariate normal kernel, and choose the bandwidths with respect to both elements of the bivariate distribution using a (data-dependent) cross-validation method. Summary statistics for selected sub-groups are presented in Table 11. Summary statistics for the various STEM major by sub-groups are provided in Table 12.

In Figure 12 we present the estimated joint distribution for the entire sample of STEM students, by status, namely *stayers* and *quitters*. Quitters are those that have, at some point switch out of their STEM major. This includes all the individuals who quit the university altogether and those who switch to majors that are not STEM major. The stayers includes all the individuals who graduated from a STEM major. Figure 12a provides the estimated joint distribution of $\hat{\alpha}$ for the stayers, while Figure 12b provides the joint distribution of $\hat{\alpha}$ for the quitters. Finally, Figure 12c provides the difference in the densities for the above two groups. The overall integral of the difference in densities is, of course,

zero, by construction. Depicting this difference in densities figures allow us to examine where in the skill distribution are the quitters and stayers located.

Overall, we see very high correlation between the student's ability in STEM and non-STEM course; the correlation coefficient is 0.67 for the stayers and 0.73 for the quitters. Nevertheless, we have also strong evidence that a one factor model (for talent) would not be suitable. While some individuals are quite talented in both type of courses, there is a significant number that have very clear comparative advantage in only one of the course types. Note that while the general features of the joint distribution of $\hat{\alpha}$ is quite similar for the stayers and quitters, the location of the joint densities are quite different. This is clearly demonstrated in Figure 12c. The concentration of the ability, i.e. of $\hat{\alpha}$, of the quitter is much closer to the origin of the figure than for the stayers.

Those students who are good in STEM courses, and who therefore have large estimated fixed effects, tend to complete their studies at the university as a STEM major. The mean of the STEM fixed effect for the STEM stayers is 0.076 (with a st. dev. of 0.50), while the mean fixed effect for the STEM quitters is -0.44 (with a st. dev. of 0.49). That is, the better students are those who are also more likely to complete their STEM major studies. However, there are two important facts to observe. First, the mean of the non-STEM fixed effect for the STEM stayers is also higher than that of the STEM quitters, 0.18 vs -0.014, respectively. That is, the best STEM students tend also to be better in the non-STEM courses. The concentration of those who do graduate is in the top right hand side corner of Figure 12c, while the concentration of the quitters is at the bottom left hand side corner of the figure. It is worthwhile noting that relative to their ability in the STEM courses, the mean non-STEM ability for the quitters is much better than their STEM ability, indicating that A STEM major might have not been the best choices for them to begin with. It's also notable that there is a wide range of STEM ability among the stayers, which suggests that there must be some element of persistence that drive some students to complete their studies despite having lower STEM ability. To examine this issue further we now turn to examining the minority, and other disadvantage groups.

We provide the results for the minority group, namely the group that includes African American and Hispanics, in Figures 13 through 15. Figure 13 provide the results for the minority, organized in the same fashion as Figure 12 above. Then, Figure 14 and Figure 15 provide the results for the minority group for two sub-periods, respectively: (a) the period up to 1997; and (b) the period from 1998 onward. We do so in order to examine whether or not affirmative action, which was abolished at the end of 1997 had any implication

on the graduation patterns from STEM major of individuals from the minority group. Prior to 1997 only 36% of the minority students who started a STEM completed a degree in a STEM major. In contrast, from 1998 onward over 45% of the minority students completed their STEM degree. This alone seems to indicate that affirmative action was not a useful tool, since it attracted more individuals who had to quit their STEM major. In fact, comparison of the results in Figure 14 with the corresponding results of Figure 15 show that the distribution of the ability among the stayers became more dense, and at a somewhat higher level from 1998 onward. However, It does not seem that the abolishment of affirmative action improved in any way the quality of the minority quitters. In fact, the distribution of the ability among the quitters up to 1997 seems to dominate that from 1998 onward. That is, it is suggestive that the affirmative action was not biased toward getting more bad students. It may be that the quality of the student is only one element of the set of element necessary for one to complete his/her STEM major degree. For example, the results suggest that the higher attrition rate among minority is due other factors, such as liquidity constraint, role models in their communities, etc.

It is worthwhile mentioning that about 65% of the male population who were admitted to a STEM major complete their education in that major, while only 53% of the female population remains in STEM. Overall, it seems that there is less of a clear separation between stayers and quitters for the female population than for the male population. It is apparent (figure not shown) that the joint density of the female population of stayers is more dense along the diagonal than that for the male population for both the quitters and stayers. Overall, it seems that the “right” female and male individuals quit STEM majors, but there is a relatively sizable fraction of graduates that have worse abilities (as measured by their bivariate fixed effects) relative to the population of quitters, and the difference is more pronounced for a male than for a female. That is, a female student is more likely to quit a STEM major than her male counterpart with the same observed ability, holding everything else constant.

The distribution of ability for Asian students is somewhat different from that of any other group, as is seen in Figure 17. First, the attrition rate from STEM majors is the lowest among all groups, at less than 34%. Also, the distribution of stayers and quitters are very distinct, with the stayers concentrated around high levels of fixed effects for both STEM and non-STEM, while the distribution of the quitters is concentrated around values of the fixed effects that are quite low relatively to their stayers counterparts.

5.4 The Distribution of Abilities by STEM Major

In Figures 16 through 21 we present the results by STEM fields, that is the fields in which the students were admitted to the university. The results for the life sciences are presented in Figures 16 and 17 for the minority and non-minority students, respectively. The results for the physical sciences are presented in similar fashion in Figures 18 and 19. Finally, the results for engineering are presented in Figures 20 and 21. Each of these figures depicts as before the results for the stayers, quitters, and the difference in density of the two.

There are a few important things to note for both the minority and non-minority groups. First the joint distribution of the fixed effects for STEM and non-STEM majors are markedly different across the three STEM fields. It is particularly different for the engineering field. The attrition rate for the engineering students is substantially lower than those for the other two STEM subjects. In engineering 80% of the non-minority students and 67% of the minority students complete their studies. In contrast, in the life and physical sciences these percentages are only 30% and 31%, respectively, for the minority students. For the non-minority student there are much higher, at 56% in both the life sciences and Physical sciences fields.

Very clearly, the joint distributions of the STEM and non-STEM fixed effect are quite different for the three fields, especially for the quitters. Higher ability students in the life sciences appear much more likely to continue in their life science major. That is, the level of $\hat{\alpha}$ provides a good predictor as to whether or not they switched out of the life science major. The case for the physical science majors is dramatically different. In fact, for the physical sciences we find the weakest separation between stayers and quitters based on their ability vector $\hat{\alpha}$. The separation based on ability is quite pronounced in the engineering fields for the minority students. The results indicate that there is a clear cutoff for the pair $\hat{\alpha}$ below which one is very likely to switch out of the engineering major and above which an individual is very likely to graduate from that major. This is not the case however for the non-minority engineering students as there is a relatively wide range of values for $\hat{\alpha}$ for each it is difficult to determine the likelihood that a student will graduate from the major. Yet, as we pointed out above, the unconditional probability that a non-minority student will graduate from an engineering major is quite high to begin with, over 80%. This is not to say the students in engineering are better than all other students, but rather that the admission process followed by the engineering school has been consistent in identifying the “suitable” non-minority individuals for the majors.

In sum, examination of the joint distributions of STEM and non-STEM ability suggest that students sort themselves into their terminal major based on their relative skills, rather than absolute skills. Additionally, the distribution of skill varies widely across STEM majors (life sciences, physical sciences, and engineering), which suggests that there are differences in difficulty levels across the majors. While minority students are on average less skilled in STEM and non-STEM relative to non-minorities, we still observe relatively weaker STEM ability minorities switching out of STEM. In the next section we examine whether there are significant differences in switching behavior by race after controlling for STEM and non-STEM ability.

5.5 Student Outcomes

We now turn to the results of the multinomial logit model of the following outcomes: graduating STEM within five years, graduating non-STEM within five years, and not graduating within five years. Table 7a presents the marginal effects of the explanatory variables on the probability of graduating STEM within five years. Column 1 reports the differences in graduation rates by race in a model with only year fixed effects. Without controlling for differences in ability and academic performance, minorities are 21.2 percentage points less likely to graduate within five years relative to white students. In column 2 we see that the graduation gap decreases to 11.6 percent after controlling for their initial STEM major, SAT math and verbal scores, and high school GPA. Interestingly, zip code income does not appear to have a significant effect on STEM graduation rates. The minority gap decreases to 6.3 percentage points after controlling for the student's academic performance their first year (column 3). In particular, STEM GPA is a strong predictor of STEM graduation: A one letter grade increase in first year STEM GPA increases the probability of graduating as a STEM major by 26.8 percentage points. On the other hand, students who take non-STEM courses in their first year are 35.7 percentage points less likely to graduate as a STEM major within five years. These students may have decided to switch out of STEM early in their student career. Even if the student does not switch out of STEM, students who take non-STEM courses their first year may be more likely to fall behind on their STEM courses, which could reduce the likelihood of an on-time graduation.

When the student's STEM and non STEM ability are included as controls (column 4), minorities students are 5 percentage points less likely to graduate as a STEM major within five years. These results illustrate the importance of including a more accurate predictor

of student performance in STEM courses when examining graduation outcomes: much of the differences in graduation rates between minorities and non-minorities can be explained by differences in STEM and non-STEM ability and first year outcomes. The equivalent of a one letter grade increase in STEM ability increases the likelihood of an on-time STEM graduation by 26.7 percentage points. Non-STEM ability also increases the likelihood of STEM graduation, possibly because students with high STEM ability typically also have high non-STEM ability. The graduation gap for minorities continues to decrease after controlling for first year GPA and student preparation (column 5), as well as first year peers (column 6). Despite that fact that having strong peers tends to lower a student's expected STEM grades, students are more likely to graduate on time as a STEM major when they are exposed to strong STEM peers in their first year of college. Students are also more likely to graduate as a STEM major when there is a larger fraction of first year STEM classmates who are their own gender and own race.

Table 7b presents the marginal effects of the explanatory variables on the probability of graduating as a non-STEM major within five years. Without controls, minorities are 2.3 percentage points more likely to graduate non-STEM graduation rates relative to white students (column 1), but the coefficient becomes negative as we add controls for background (column 2), first year performance (column 3), and ability (column 4). Students with higher STEM ability are less likely to graduate as a non-STEM major, whereas students with higher non-STEM ability are more likely to graduate as a non-STEM major (column 4). STEM peers appear to have more modest effects (column 6). Students with a higher fraction of first year STEM peers their own gender are less likely to graduate as a non-STEM major, and there is no effect for the fraction of peers their own race. Whereas zip code income did not play a role in the probability of graduating a STEM major, it does appear to increase the probability the student graduates as a non-STEM major. It's possible that the relative value of a STEM major versus non-STEM major is lower for students from a wealthier background.

Finally, Table 7c presents the marginal effects of the explanatory variables on the probability of not graduating within five years. The raw difference in dropout rates for minorities relative to white students is 18.8 percentage points (column 1), which decreases to 14.6 percentage points after controlling for first year performance (column 3). We also observe that first year STEM GPA and non-STEM GPA both decrease the likelihood of not graduating within five years. Similarly, having higher STEM and non-STEM ability both decrease the likelihood of not graduating within five years (column 4). This suggests

that having higher ability in any field provides at least one pathway to an on-time graduation. After controlling for ability, performance, and peers (column 6), minorities are 6.3 percentage points more likely to not graduate within five years. Students with higher zip code income are less likely to not graduate, possibly because they are more willing to switch to a non-STEM major. This highlights one of the costs of a low ability student persisting in STEM: they may have been able to graduate on time if they had switched to a non-STEM major.

As a robustness check we estimate the same outcomes, but over a four year time period (See Appendix A.1). While the raw STEM graduation gap between minority and white students is larger for the 5 year rate versus 4 year rate, controlling for ability reduces the gap by a larger degree in the 5 year rate. We also observe a stronger relationship between STEM ability and the likelihood of graduating as a STEM major within 5 years versus 4 years. Taken together, this suggests that there are other factors that are playing a role in the switching decision for students who graduate within four years. In particular, students may be weighing the benefits of graduating as a non-STEM major within four years versus the financial cost of taking more time to complete a STEM major. We investigate this issue further in our term level switching regressions.

5.6 Switching Dynamics

So far, we have seen that differences in ability and first year performance can explain a great deal of the differences in STEM graduation rates for minorities and non-minorities, but the gap remains. We now examine time varying factors that contribute to a student deciding whether or not to switch out of STEM in a given term. In particular, we will be able to see if the lower STEM graduation rates for minorities is due to being more likely to switch out of STEM, or dropping out of school entirely. The estimated coefficients from the probit regressions are reported in Table 8, where Column 1 estimates the switching regression across all students who were admitted as a STEM major. Surprisingly, minority students are less likely to switch after controlling for ability and grades. This would suggest that minority students have a stronger preference for staying in STEM, regardless of their ability to complete STEM.

Taking any STEM courses in the current term is associated with a lower likelihood of switching, while taking any non-STEM courses is associated with a higher likelihood of switching. We also observe that students with a higher STEM GPA are less likely to switch

out of STEM, which suggests that students use grades to assess their ability. We also see evidence that students who take upper division STEM courses are less likely to switch out of STEM relative to students who are still taking lower division classes. Students in upper division courses may be less likely to switch because the cost of switching is higher later in their career. The university places a maximum limit on the total number of units a student can take in their career, so it becomes unlikely they will have enough credits left to complete another major at this point. Taking an upper division course also indicates that the student has successfully completed the preparation for the major. Turning to peer effects, we see that students are less likely to leave STEM when they have higher ability peers (measured by SAT math scores). This effect is similar to the effect of first year peer ability on STEM graduation rates. We suspect that this coefficient partially captures the effect of progressing to the more advanced classes within the STEM major because students are more likely to stay in STEM once they've reached upper division courses. On the other hand, the gender composition and racial composition of a student's STEM peers do not have a significant effect on switching behavior.

When the switching regression is estimated separately by race (Columns 2-4 of Table 8), it becomes evident that there is heterogeneity in the effect of grades and course choices on the decision to leave STEM. Minority students are more likely to switch out of STEM if they are currently taking a non-STEM course relative to Asian and white students, and also respond more strongly to their non-STEM grades. Minority students are also less likely to switch out of STEM if they are currently taking an upper division STEM course. Lastly, we can see that the effect of having STEM peers of the same race on switching is also largest for minorities, possibly because race may be more representative of peer groups for minority students.

5.6.1 Switching Dynamics with unobserved Types

In order to allow for unobserved heterogeneity, we estimate various specifications of the switching regression. Specifically, we estimate a probit model of the switching decision while allowing for either two unobserved types (Table 9a) or three unobserved types (Table 10a) in the population. Due to the high degree of multicollinearity between ability and GPA, we choose to include the student's cumulative GPA in our preferred specification. Results with ability controls are presented in the Appendix.

In the two type specification, the population of initial STEM majors consists of 56

percent Type I's and 44 percent Type II's. Type I students are much less likely to switch out of STEM when they have a higher STEM GPA. Having a higher non-STEM GPA also lowers the likelihood of switching for Type I students. These results are consistent with the descriptive statistics, which show that STEM stayers have higher STEM and non-STEM ability relative to switchers. It would appear that these students are not selecting their major based on their comparative advantage. Type I students are slightly less likely to switch out of STEM after advancing to upper division STEM courses, but there is no clear relationship between the number of quarters a student has been a STEM major and the switching decision.

Looking with STEM subfields, it appears that engineering majors and physical science majors are less likely to switch out of STEM relative to life science majors. These students have higher STEM ability, which provides further evidence that Type I students are the higher ability students in the population. Minority and Asian Type I's are less likely to switch out relative to White students. Turning to peer effects, we observe that Type I students are less likely to switch out when their STEM peers have higher ability (SAT math scores), but there are no significant effects of having a higher fraction of STEM peers of the same race and gender as the student.

There are a few notable differences between the Type I and Type II coefficients. Type II students are less responsive to their STEM GPA, and are more likely to switch out when they have higher non-STEM GPA. These coefficients suggest that this fraction of the population is considering their comparative advantage when deciding whether or not to switch out of STEM. These student may be placing a higher value on their outside options compared to the Type I students, who do not switch out of STEM despite having relatively high non-STEM ability. One reason for valuing the non-STEM majors more highly could be the fact that non-STEM majors are easier to complete within four years. If students face liquidity constraints, switching majors could be a strategy for ensuring an on time graduation. The coefficients on major progression and duration provide evidence supporting this hypothesis. Type II students are much less likely to switch out of STEM once they have advanced to upper division STEM courses, which is an indication of being on track for graduation. Similarly, we see a large increase in the likelihood of switching between their third and fourth year as a STEM major (terms 9 and 10). Students may be switching during this time if they are concerned about graduating within four years. This is consistent with the descriptive statistics that show higher switching rates for 4 year graduates versus 5 year graduates. Type II engineering and physical science majors are

no more likely to switch relative to life science majors. While these students tend to have higher ability, the major itself may be more challenging to complete on time.

Type II students do not appear to base their switching decision across any dimension of their STEM peer characteristics, which again suggests that liquidity constraints are the most important factor in the switching decision. Minority and Asian Type II students are much less likely to switch out compared to Type I students. This could suggest that minority students have a stronger preference for graduating STEM versus non-STEM despite being financially constrained. It's possible that liquidity constrained minorities are weighing the the higher returns to graduating STEM versus non-STEM major more than the cost of not graduating on time.

We find similar results when allowing for three unobserved types in the population (Table 10a). The Type I coefficients look similar to the Type I's in the two type model, but are estimated to comprise a smaller fraction of the population (41 percent versus 55 percent). There is now a stronger negative relationship between STEM GPA and the likelihood of switching. The relationship between non-STEM GPA and switching also negative, which suggests that these students are choosing their major based on their STEM ability alone. In the three type model, there are fewer Type II students in the population (18%). The type II students appear to be less responsive to their STEM and non-STEM GPA relative to the Type I students. Having a higher STEM GPA makes it more likely that the student will switch out of STEM, but the effect of non-STEM GPA is positive but insignificant. Much like the Type II students in the two type model, progressing to an upper level STEM course appears to have a strong negative effect on the likelihood of leaving STEM. Similarly, the likelihood of leaving STEM peaks around the student's junior year.

The inclusion of a third type in the population appears to have separated students who are liquidity constrained (Type II) from those who choose their major based on their comparative advantage (Type III). Type III students are less likely to switch out of STEM when their STEM GPA is high, and are more likely to switch when their non-STEM GPA is high. These students are less likely to switch out once they take upper division STEM courses, but the effect is not as strong relative to the Type II students. There is no clear relationship between the student's time spent as a STEM major and the likelihood of switching, which suggests that preferences for on-time graduation is not playing a strong role in the switching decision.

6 Conclusion

In this paper we find significant differences in the joint distribution of STEM and non-STEM ability across many types of students. Comparison of STEM switchers vs non-switchers reveal that students with relatively higher non-STEM ability are more likely to switch out of STEM, although there is substantial overlap between the two joint distributions. These results highlight the importance of considering multidimensional ability measures, as some students sort themselves according to their relative strengths. Even after controlling for unobserved STEM and non-STEM ability, we see that race still plays a role in who switches out of STEM, and ultimately graduates on time. We find evidence that Black students are less responsive to their STEM grades, while Hispanic students are less responsive to their STEM ability when deciding whether or not to switch out of STEM in a give term.

The overlap in ability distribution between switchers and non-switchers also raises the question of whether or not a student who switched out could have succeeded as a STEM major if they had remained a STEM major. Similarly, it is unclear whether or not weaker STEM students may have been able to graduate on time in another major. A limitation of our study is that we cannot examine how a student would have fared attending another university. This concept has been referred to as the mismatch hypothesis in the literature. Arcidiacono, Aucejo, and Hotz (2013) find that less prepared minorities would have been more likely to graduate in a STEM field from a less competitive school in the UC system. Similarly, Luppino and Sander (2013) find that minority students are more adversely affected by the competitiveness of a major when they are at a selective university relative to non-minority students.

References

- Antonovics, K.** and B. Backes (2013): “Were Minority Students Discouraged From Applying to University of California Campuses After the Affirmative Action Ban?” *Education Finance and Policy*, 8(2), 208-250.
- Antonovics, K.**, and R. Sander (2013): “Affirmative Action Bans and the “Chilling Effect,” *American Law and Economics Review*, 15(1), 252-299.
- Arcidiacono, P.**, E. M. Aucejo and V. J. Hotz (2013): “University Differences in the Graduation of Minorities in STEM Fields: Evidence from California,” National Bureau of Economic Research, Working Paper No. 18799.
- Arcidiacono, P.**, E. M. Aucejo and K. Spenner (2012): “What Happens After Enrollment? An Analysis of Time Path of Racial Differences in GPA and Major Choice,” Unpublished Manuscript.
- Arcidiacono, P.** (2004): “Ability Sorting and the Returns to College Major,” *Journal of Econometrics*, 121, pp. 343-375.
- Backes, B.** (2012): “Do affirmative action bans lower minority college enrollment and attainment? Evidence from statewide bans,” *Journal of Human Resources*, 47(2), 435-455.
- Card, D.**, & A. B. Krueger (2005): “Would the elimination of affirmative action affect highly qualified minority applicants? Evidence from California and Texas,” *Industrial and Labor Relations Review*, 58(3), 416-434.
- Carnevale, A.P.**, N. Smith and M. Melton (2011): “STEM: Science, Technology, Engineering, Mathematics,” Report, Georgetown University
- Chen, X.** and M. Soldner (2013): “STEM Attrition: College Students’ Paths Into and Out of STEM Fields: Statistical Analysis Report,” U.S. Department of Education, National Center for Education Statistics.
- Gemici, A.** and M. Wiswall (2014): “Evolution of Gender Differences in Post-Secondary Human Capital Investments: College Majors,” *International Economic Review*, 55, pp. 23-56.

- Hinrichs, P.** (2012): “The effects of affirmative action bans on college enrollment, educational attainment, and the demographic composition of universities,” *Review of Economics and Statistics*, 94(3), 712-722.
- Hinrichs, P.** (2014). Affirmative action bans and college graduation rates. *Economics of Education Review*, 42, 43-52.
- Luppino, M. and R. Sander** (2015): “College major peer effects and attrition from the sciences,” *IZA Journal of Labor Economics*, **4**, 1-23.
- Ost, B.** (2010): “The Role of Peers and Grades in Determining Major Persistence in the Sciences,” *Economics of Education Review*, **29**, pp. 923–934
- Rask, K.** (2010): “Attrition in STEM Fields at a Liberal Arts College: The Importance of Grades and Pre-Collegiate Preferences,” *Economics of Education Review*, **29**, pp. 892–900.

Table 1: Summary Statistics of Minority and Non-Minority Students Admitted into a STEM Major

	Minority			Non Minority		
	N	Mean	SD	N	Mean	SD
Graduated within 5 yrs as a STEM major	3,191	0.63	0.48	19,657	0.81	0.39
Graduated within 4 yrs as a STEM major	3,191	0.15	0.36	19,657	0.37	0.48
Switched out of STEM	3,191	0.54	0.49	19,657	0.36	0.48
Male	3,191	0.51	0.50	19,657	0.61	0.49
Transfer Student	3,191	0.08	0.26	19,657	0.06	0.23
Median Household Income (\$1000's)	3,191	54.70	24.63	19,657	67.67	28.14
SAT math score	3,191	561.54	95.20	19,657	667.70	77.25
SAT verbal score	3,191	490.80	96.75	19,657	553.19	104.62
High school GPA	3,191	3.67	0.66	19,657	3.90	0.74
Admitted Physical major	3,191	0.20	0.40	19,657	0.17	0.38
Admitted Engineering major	3,191	0.24	0.43	19,657	0.32	0.47
Admitted Life major	3,191	0.56	0.50	19,657	0.51	0.50
First year STEM GPA	3,191	2.41	0.77	19,657	2.96	0.66
First year non-STEM GPA	3,191	2.82	0.71	19,657	3.10	0.77

Table 2: Summary Statistics of Major Transitions

Panel A. White and Asian Students								
Initial major		Final major						N
		physical	life	engineering	humanities	social	undeclared	
Initial major	physical	55.5%	6.1%	6.6%	3.3%	20.5%	4.1%	5330
	life	6.4%	51.4%	2.1%	4.2%	30.1%	2.3%	14144
	engineering	2.6%	2.1%	83.4%	0.8%	5.5%	3.8%	11077
	humanities	1.0%	2.4%	0.3%	50.3%	34.5%	0.6%	7122
	social	2.2%	3.1%	0.4%	7.9%	72.1%	4.3%	16385
	undeclared	6.9%	11.0%	2.7%	12.4%	57.0%	5.9%	25690
79748								
Panel B. Black and Hispanic Students								
Initial major		Final major						N
		physical	life	engineering	humanities	social	undeclared	
Initial major	physical	35.1%	7.0%	3.0%	6.5%	36.0%	6.6%	833
	life	4.8%	34.0%	0.7%	7.1%	44.2%	6.3%	2844
	engineering	5.3%	2.1%	73.4%	0.8%	9.0%	8.0%	1062
	humanities	0.3%	1.2%	0.1%	46.2%	41.0%	0.5%	1464
	social	1.1%	1.7%	0.2%	7.7%	69.8%	7.5%	4682
	undeclared	4.5%	5.0%	2.1%	12.5%	62.0%	9.8%	6799
17684								

Note: includes all undergraduates who began their career between 1980 and 2002 (graduates and dropouts).

Table 3: STEM Subject Area Grade Regressions

	(1)	(2)	(3)	(4)
Minority	-0.2545*** [0.003]	-0.2492*** [0.003]		
Asian	-0.1179*** [0.002]	-0.1159*** [0.003]		
Male	-0.0497*** [0.002]	-0.0460*** [0.002]		
Transfer student	0.3123*** [0.016]	0.3191*** [0.018]		
Zipcode Income /1000	0.0004*** [0.000]	0.0004*** [0.000]		
SAT Verbal	0.0024*** [0.000]	0.0023*** [0.000]		
SAT Math	0.1194*** [0.001]	0.1297*** [0.002]		
High School GPA	0.0003*** [0.000]	0.0003*** [0.000]		
Winter	0.0329*** [0.002]	0.0199*** [0.005]	0.0574*** [0.002]	0.0229*** [0.002]
Spring	0.0477*** [0.002]	0.0219*** [0.006]	0.0826*** [0.003]	0.0194*** [0.002]
Summer	0.1171*** [0.005]	0.1743*** [0.012]	0.2227*** [0.005]	0.1770*** [0.005]
Major course	-0.0349*** [0.003]	0.0314*** [0.005]	0.0314*** [0.005]	0.0407*** [0.004]
mean SAT math score	-0.0009*** [0.000]	-0.0013*** [0.000]	-0.0038*** [0.000]	-0.0016*** [0.000]
mean SAT verbal score	-0.0011*** [0.000]	-0.0008*** [0.000]	-0.0004*** [0.000]	-0.0006*** [0.000]
mean HS GPA	-0.0952*** [0.002]	-0.0464*** [0.008]	-0.0019 [0.002]	-0.0399*** [0.003]
fraction male	-0.0136 [0.009]	-0.0739* [0.038]	0.0486*** [0.010]	-0.0035 [0.017]
fraction black	0.0877** [0.038]	-0.0517 [0.112]	0.1889*** [0.039]	0.0450 [0.055]
fraction white	0.4266*** [0.022]	-0.1373** [0.059]	0.2327*** [0.022]	-0.2406*** [0.029]
fraction asian	-0.3589*** [0.021]	-0.0554 [0.056]	-0.5614*** [0.021]	-0.1677*** [0.027]
fraction hispanic	-0.1136*** [0.026]	0.0711 [0.077]	-0.1379*** [0.027]	-0.0434 [0.037]
Class size	-0.0003*** [0.000]	0.0000 [0.000]	-0.0002*** [0.000]	-0.0000*** [0.000]
Course-year fixed effects?	N	Y	N	Y
Student fixed effects?	N	N	Y	Y
R-squared	0.125	0.120	0.053	0.019
	1,396,568	1,396,568	1,396,568	1,396,568

Note: Additional controls include the student's subject area level experience (number of previous courses taken) and the student's current course load. Robust standard errors in brackets.

*** p<0.01, ** p<0.05, * p<0.10.

Table 4: Non-STEM Subject Area Grade Regressions

	(1)	(2)	(3)	(4)
minority	-0.1759*** [0.002]	-0.1720*** [0.002]		
asian	-0.0859*** [0.001]	-0.0891*** [0.001]		
male	-0.1166*** [0.001]	-0.1184*** [0.001]		
transfer student	0.1106*** [0.008]	0.1187*** [0.008]		
Zipcode Income /1000	0.0012*** [0.000]	0.0012*** [0.000]		
SAT Verbal	0.0008*** [0.000]	0.0008*** [0.000]		
SAT Math	0.0715*** [0.001]	0.0701*** [0.001]		
High School GPA	0.0002*** [0.000]	0.0002*** [0.000]		
winter	0.0240*** [0.001]	0.0177*** [0.004]	0.0285*** [0.001]	0.0176*** [0.001]
spring	0.0227*** [0.001]	0.0137*** [0.004]	0.0265*** [0.001]	0.0130*** [0.001]
summer	0.0162*** [0.002]	0.1050*** [0.008]	0.0819*** [0.003]	0.1444*** [0.003]
Major course	0.0716*** [0.001]	0.0733*** [0.002]	0.0305*** [0.002]	0.0589*** [0.002]
mean SAT math score	-0.0014*** [0.000]	-0.0001 [0.000]	-0.0025*** [0.000]	0.0001 [0.000]
mean SAT verbal score	-0.0014*** [0.000]	-0.0008*** [0.000]	-0.0012*** [0.000]	-0.0005*** [0.000]
Mean HS GPA	-0.0345*** [0.001]	-0.0683*** [0.004]	0.0106*** [0.001]	-0.0557*** [0.002]
fraction male	-0.0860*** [0.004]	-0.0340* [0.019]	0.0272*** [0.005]	-0.0041 [0.012]
fraction black	-0.0420*** [0.013]	-0.0154 [0.048]	-0.0223 [0.014]	-0.0152 [0.030]
fraction white	-0.0703*** [0.010]	0.0830*** [0.028]	-0.2149*** [0.011]	0.0008 [0.016]
fraction asian	-0.0991*** [0.010]	0.1384*** [0.031]	-0.1495*** [0.011]	0.0474*** [0.017]
fraction hispanic	-0.1463*** [0.010]	0.1472*** [0.036]	-0.2843*** [0.012]	-0.0041 [0.020]
Class size	-0.0004*** [0.000]	0.0001*** [0.000]	-0.0003*** [0.000]	0.0001*** [0.000]
Course-year fixed effects?	N	Y	N	Y
Student fixed effects?	N	N	Y	Y
R-squared	0.154	0.104	0.065	0.005
Observations	2,706,006	2,706,006	2,706,006	2,706,006

Note: Additional controls include the student's subject area level experience (number of previous courses taken) and the student's current course load. Robust standard errors in brackets.

*** p<0.01, ** p<0.05, * p<0.10.

Table 5: Correlates of STEM Ability (fixed-effect from STEM grade regressions)

Variable	(1) all	(2) white	(3) asian	(4) minority
asian	-0.1125*** [0.004]			
minority	-0.2739*** [0.006]			
male	-0.0832*** [0.004]	-0.0978*** [0.006]	-0.0877*** [0.006]	-0.0380*** [0.010]
sat math	0.0024*** [0.000]	0.0023*** [0.000]	0.0024*** [0.000]	0.0022*** [0.000]
sat verbal	0.0008*** [0.000]	0.0008*** [0.000]	0.0008*** [0.000]	0.0008*** [0.000]
high school gpa	0.0546*** [0.003]	0.0596*** [0.005]	0.0299*** [0.006]	0.1038*** [0.009]
transfer	-0.1472*** [0.009]	-0.0488*** [0.014]	-0.2627*** [0.016]	-0.1929*** [0.015]
Zipcode Income /1000	0.0005*** [0.000]	0.0002** [0.000]	0.0005*** [0.000]	0.0012*** [0.000]
admit life	-0.1771*** [0.005]	-0.1565*** [0.009]	-0.1633*** [0.008]	-0.2476*** [0.013]
admit physical	-0.1369*** [0.009]	-0.1736*** [0.014]	-0.0829*** [0.016]	-0.1665*** [0.023]
admit engineering	-0.3089*** [0.007]	-0.3448*** [0.012]	-0.2734*** [0.010]	-0.3115*** [0.021]
admit social	0.0838*** [0.005]	0.0805*** [0.008]	0.1055*** [0.009]	0.0509*** [0.011]
admit humanities	0.0423*** [0.010]	0.0340** [0.014]	0.0354* [0.019]	0.0432** [0.022]
Observations	95,994	37,500	35,941	18,355
R-squared	0.252	0.157	0.177	0.188

Note: Robust standard errors in brackets. Regressions include year fixed effects.

*** p<0.01, ** p<0.05, * p<0.10.

Table 6: Correlates of Non-STEM Ability (FE from Non-STEM grade reg.)

Variable	(1) all	(2) white	(3) asian	(4) minority
asian	-0.0850*** [0.003]			
minority	-0.1829*** [0.004]			
male	-0.1274*** [0.003]	-0.1264*** [0.004]	-0.1266*** [0.005]	-0.1278*** [0.007]
sat math	0.0009*** [0.000]	0.0010*** [0.000]	0.0009*** [0.000]	0.0007*** [0.000]
sat verbal	0.0011*** [0.000]	0.0010*** [0.000]	0.0011*** [0.000]	0.0012*** [0.000]
high school gpa	0.0226*** [0.002]	0.0320*** [0.003]	0.0034 [0.004]	0.0380*** [0.005]
transfer	-0.0955*** [0.006]	-0.0326*** [0.009]	-0.1854*** [0.012]	-0.1131*** [0.011]
Zipcode Income /1000	0.0003*** [0.000]	-0.0000 [0.000]	0.0003*** [0.000]	0.0012*** [0.000]
admit life	0.0215*** [0.004]	0.0390*** [0.006]	0.0199*** [0.005]	-0.0140 [0.010]
admit physical	-0.0406*** [0.007]	-0.0422*** [0.011]	-0.0283** [0.012]	-0.0753*** [0.018]
admit engineering	0.0450*** [0.005]	0.0160* [0.009]	0.0587*** [0.008]	0.0708*** [0.018]
admit social	0.0212*** [0.004]	0.0280*** [0.005]	0.0144** [0.006]	0.0075 [0.008]
admit humanities	-0.0007 [0.006]	-0.0007 [0.008]	-0.0246** [0.012]	0.0136 [0.015]
Observations	99,693	39,330	36,870	18,978
R-squared	0.212	0.148	0.152	0.131

Note: Robust standard errors in brackets. Regressions include year fixed effects.

*** p<0.01, ** p<0.05, * p<0.1

Table 7a: Multinomial logit regression results. Probability of graduating STEM within 5yrs.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
asian	0.0742*** (0.00726)	0.0585*** (0.00777)	0.0794*** (0.00828)	0.113*** (0.00805)	0.0862*** (0.00836)	0.0502*** (0.0101)
minority	-0.212*** (0.00949)	-0.116*** (0.0118)	-0.0627*** (0.0127)	-0.0504*** (0.0119)	-0.0434*** (0.0128)	-0.00308 (0.0138)
male	0.0876*** (0.00672)	0.000663 (0.00752)	0.0190** (0.00806)	0.0565*** (0.00769)	0.0300*** (0.00813)	0.00607 (0.00920)
transfer		-0.322*** (0.0101)	-0.342*** (0.0118)	-0.210*** (0.0150)	-0.341*** (0.0128)	-0.373*** (0.0120)
Zipcode Income /1000		-0.000107 (0.000131)	-0.000211 (0.000140)	-0.000120 (0.000135)	-0.000240* (0.000140)	-0.000167 (0.000145)
admit physical		-0.0294*** (0.0110)	-0.0217* (0.0117)	-0.00352 (0.0114)	-0.0156 (0.0118)	-0.0809*** (0.0123)
admit engineering		0.105*** (0.00822)	0.156*** (0.00883)	0.175*** (0.00839)	0.147*** (0.00907)	-0.0281** (0.0110)
STEM ability				0.267*** (0.00925)	0.0211 (0.0142)	0.00751 (0.0148)
non-STEM ability				0.0932*** (0.0124)	0.133*** (0.0150)	0.122*** (0.0153)
SAT math score		0.00133*** (5.44e-05)	0.000573*** (5.99e-05)		0.000605*** (6.02e-05)	-0.000218*** (6.65e-05)
SAT verbal score		-0.000196*** (4.33e-05)	-0.000330*** (4.72e-05)		-0.000417*** (4.81e-05)	-0.000389*** (5.00e-05)
High School GPA		-0.0272*** (0.00640)	-0.0642*** (0.00679)		-0.0643*** (0.00690)	-0.0780*** (0.00717)
took non STEM classes			-0.357*** (0.0237)		-0.165*** (0.0380)	-0.0487 (0.0420)
first yr STEM GPA			0.268*** (0.00713)		0.223*** (0.0110)	0.237*** (0.0117)
first yr non STEM GPA			0.0645*** (0.00815)		0.0126 (0.00956)	-0.0177* (0.00987)
first yr peers (SAT Math)						0.00648*** (0.000195)
first yr fraction own race						0.0478 (0.0359)
first yr fraction own gender						0.0390 (0.0373)
Observations	22,848	22,848	22,848	22,848	22,848	22,848

Note: This table reports the marginal effects of variables on the probability of graduating STEM within 5 yrs. Estimates are derived from a multinomial logit regression with three outcomes: graduate STEM within 5 yrs, graduate non-stem within 5yrs, do not graduate within 5 yrs. The regression is restricted to student who started their career as a STEM major. Robust standard errors in brackets. Regressions include year fixed effects (not shown). The median household income variable is derived from the student's zip code and is expressed in thousands of dollars. *** p<0.01, ** p<0.05, * p<0.1

Table 7b: Multinomial logit regression results. Probability of graduating non-STEM within 5yrs.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
asian	-0.0882*** (0.00656)	-0.0733*** (0.00690)	-0.0815*** (0.00716)	-0.101*** (0.00705)	-0.0729*** (0.00728)	-0.0489*** (0.00838)
minority	0.0232*** (0.00893)	-0.0302*** (0.00950)	-0.0490*** (0.00961)	-0.0128 (0.00970)	-0.0336*** (0.0102)	-0.0602*** (0.0103)
male	-0.146*** (0.00615)	-0.0634*** (0.00660)	-0.0657*** (0.00692)	-0.0828*** (0.00672)	-0.0596*** (0.00703)	-0.0362*** (0.00744)
transfer		-0.188*** (0.0100)	-0.169*** (0.0106)	-0.161*** (0.00969)	-0.169*** (0.0115)	-0.153*** (0.0120)
Zipcode Income /1000		0.000490*** (0.000113)	0.000567*** (0.000117)	0.000415*** (0.000115)	0.000511*** (0.000119)	0.000454*** (0.000122)
admit physical		0.00395 (0.00903)	0.00813 (0.00951)	0.00388 (0.00936)	0.00459 (0.00967)	0.0527*** (0.0109)
admit engineering		-0.178*** (0.00665)	-0.197*** (0.00679)	-0.226*** (0.00648)	-0.200*** (0.00699)	-0.0733*** (0.00942)
STEM ability				-0.175*** (0.00806)	0.104*** (0.0124)	0.113*** (0.0127)
non-STEM ability				0.0945*** (0.0106)	0.0898*** (0.0130)	0.108*** (0.0134)
SAT math score		-0.000876*** (4.62e-05)	-0.000467*** (5.00e-05)		-0.000506*** (5.14e-05)	0.000172*** (5.64e-05)
SAT verbal score		0.000322*** (3.77e-05)	0.000346*** (4.00e-05)		0.000296*** (4.14e-05)	0.000265*** (4.33e-05)
High School GPA		-0.0131** (0.00562)	0.0134** (0.00596)		0.00974 (0.00607)	0.0202*** (0.00625)
took non STEM classes			0.196*** (0.0208)		0.238*** (0.0141)	0.191*** (0.0230)
first yr STEM GPA			-0.168*** (0.00595)		-0.249*** (0.00948)	-0.259*** (0.00986)
first yr non STEM GPA			0.0227*** (0.00690)		-0.0346*** (0.00847)	-0.00956 (0.00884)
first yr peers (SAT Math)						-0.00506*** (0.000159)
first yr fraction own race						0.00269 (0.0294)
first yr fraction own gender						-0.118*** (0.0313)
Observations	22,848	22,848	22,848	22,848	22,848	22,848

Note: This table reports the marginal effects of variables on the probability of graduating non-STEM within 5 yrs. Estimates are derived from a multinomial logit regression with three outcomes: graduate STEM within 5yrs, graduate non-stem within 5yrs, do not graduate within 5 yrs. The regression is restricted to student who started their career as a STEM major. Robust standard errors in brackets. Regressions include year fixed effects (not shown). The median household income variable is derived from the student's zip code and is expressed in thousands of dollars. *** p<0.01, ** p<0.05, * p<0.1

Table 7c: Multinomial logit regression results. Probability of not graduating within 5yrs.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
asian	0.0140** (0.00634)	0.0147** (0.00674)	0.00209 (0.00702)	-0.0119* (0.00665)	-0.0133* (0.00697)	-0.00129 (0.00822)
minority	0.188*** (0.0100)	0.146*** (0.0114)	0.112*** (0.0113)	0.0632*** (0.00968)	0.0770*** (0.0108)	0.0632*** (0.0114)
male	0.0582*** (0.00548)	0.0628*** (0.00608)	0.0467*** (0.00640)	0.0263*** (0.00601)	0.0296*** (0.00646)	0.0301*** (0.00707)
transfer		0.510*** (0.0138)	0.511*** (0.0155)	0.371*** (0.0156)	0.510*** (0.0169)	0.526*** (0.0165)
Zipcode Income /1000		-0.000383*** (0.000116)	-0.000356*** (0.000119)	-0.000295*** (0.000110)	-0.000270** (0.000117)	-0.000288** (0.000121)
admit physical		0.0255*** (0.00966)	0.0135 (0.00976)	-0.000357 (0.00908)	0.0110 (0.00970)	0.0282*** (0.0105)
admit engineering		0.0732*** (0.00722)	0.0400*** (0.00748)	0.0515*** (0.00697)	0.0536*** (0.00766)	0.101*** (0.00978)
STEM ability				-0.0924*** (0.00706)	-0.125*** (0.0110)	-0.120*** (0.0115)
non-STEM ability				-0.188*** (0.00895)	-0.223*** (0.0111)	-0.231*** (0.0115)
SAT math score		-0.000450*** (4.41e-05)	-0.000106** (4.75e-05)		-9.84e-05** (4.68e-05)	4.61e-05 (5.14e-05)
SAT verbal score		-0.000126*** (3.64e-05)	-1.56e-05 (3.95e-05)		0.000121*** (3.91e-05)	0.000123*** (4.09e-05)
High School GPA		0.0403*** (0.00466)	0.0508*** (0.00496)		0.0545*** (0.00515)	0.0578*** (0.00539)
took non STEM classes			0.161*** (0.0108)		-0.0728** (0.0366)	-0.143*** (0.0418)
first yr STEM GPA			-0.100*** (0.00548)		0.0264*** (0.00851)	0.0219** (0.00891)
first yr non STEM GPA			-0.0873*** (0.00621)		0.0220*** (0.00730)	0.0272*** (0.00762)
first yr peers (SAT Math)						-0.00142*** (0.000146)
first yr fraction own race						-0.0505* (0.0286)
first yr fraction own gender						0.0795*** (0.0301)
Observations	22,848	22,848	22,848	22,848	22,848	22,848

Note: This table reports the marginal effects of variables on the probability of not graduating within 5yrs.

Estimates are derived from a multinomial logit regression with three outcomes: graduate STEM within 5 yrs, graduate graduate within 5 yrs. The regression is restricted to student who started their career as a STEM major. Robust standard errors non-stem within 5yrs, do not in brackets. Regressions include year fixed effects (not shown).

The median household income variable is derived from the student's zip code and is expressed in thousands of dollars.

*** p<0.01, ** p<0.05, * p<0.1

Table 8: Transition out of STEM Major—with STEM and Non-STEM GPA's.

VARIABLE	(1) All		(2) White		(3) Asian		(4) Minority	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Minority	-0.1179	0.0150						
Asian	-0.1553	0.0117						
Male	-0.1024	0.0108	-0.1429	0.0190	-0.0870	0.0163	-0.0720	0.0239
Transfer Student	-0.0878	0.0208	0.0562	0.0358	-0.0770	0.0338	-0.3010	0.0425
Zipcode Income	0.0003	0.0002	0.0001	0.0003	0.0007	0.0003	0.0006	0.0004
Winter	-0.0172	0.0230	-0.0366	0.0437	0.0327	0.0342	-0.0869	0.0457
Spring	0.1645	0.0224	0.1207	0.0426	0.2489	0.0342	0.0477	0.0437
Any stem classes	-0.4778	0.0149	-0.5795	0.0256	-0.5127	0.0248	-0.3329	0.0300
Any non stem classes	0.3917	0.0601	0.3315	0.1047	0.3351	0.0921	0.6033	0.1554
STEM GPA	-0.2443	0.0085	-0.1998	0.0143	-0.3169	0.0137	-0.1823	0.0188
Non-STEM GPA	-0.0627	0.0115	-0.0346	0.0204	-0.0368	0.0180	-0.1606	0.0254
Admitted to Physical	-0.0249	0.0147	-0.0224	0.0230	-0.0481	0.0258	0.0058	0.0315
Admitted to Engineering	-0.2176	0.0128	-0.1402	0.0224	-0.2653	0.0187	-0.2080	0.0335
any upper stem classes	-0.7137	0.0157	-0.7568	0.0286	-0.6329	0.0229	-0.7937	0.0387
Qtr×SAT math /1000	0.0062	0.0578	0.5428	0.1385	-0.8018	0.1605	0.6852	0.1905
Qtr×gender	-0.0379	0.0545	0.0460	0.0991	-0.0794	0.0818	0.0251	0.1351
Qtr×race	-0.0001	0.0001	-0.0005	0.0001	0.0005	0.0001	-0.0004	0.0001
term2	0.1444	0.0362	0.1470	0.0607	0.0654	0.0572	0.3733	0.0850
term3	0.3214	0.0341	0.3031	0.0578	0.2375	0.0540	0.6476	0.0794
term4	0.4655	0.0271	0.3416	0.0431	0.4872	0.0433	0.7021	0.0688
term5	0.6062	0.0344	0.5429	0.0587	0.5701	0.0538	0.8574	0.0811
term6	0.7025	0.0337	0.5991	0.0580	0.6579	0.0531	1.0714	0.0782
term7	0.7187	0.0289	0.5444	0.0475	0.7416	0.0460	1.0235	0.0716
term8	0.7719	0.0360	0.6024	0.0630	0.7500	0.0557	1.1937	0.0830
term9	0.7092	0.0360	0.5908	0.0625	0.6248	0.0569	1.1741	0.0825
term10	0.6946	0.0330	0.4484	0.0577	0.6825	0.0513	1.1824	0.0795
term11	0.6260	0.0401	0.3860	0.0728	0.5012	0.0628	1.2729	0.0897
term12	0.7706	0.0387	0.5963	0.0687	0.6992	0.0593	1.2838	0.0915
term13	0.7151	0.0391	0.5001	0.0719	0.6844	0.0594	1.1673	0.0918
term14	0.7158	0.0462	0.5232	0.0868	0.5981	0.0704	1.2695	0.1032
term15	0.8882	0.0460	0.6030	0.0888	0.8505	0.0686	1.3775	0.1041
term16	0.9226	0.0641	0.5092	0.1271	0.9821	0.0942	1.4503	0.1325
term17	0.9021	0.0763	0.6286	0.1445	0.8411	0.1168	1.5048	0.1628
term18+	0.9936	0.0659	0.6659	0.1236	0.9314	0.0961	1.6352	0.1573
N	245,746		76,483		125,072		35,142	

Table 9a: Transition out of STEM Major—Two-Type Model with STEM and Non-STEM GPA's

No.	Variable	Type I		Type II	
		Coefficient	St. Error	Coefficient	St. Error
1	Constant	-5.8836	0.1491	-1.0032	0.0968
2	Minority	-0.0810	0.0378	-0.1277	0.0286
3	Asian	-0.0458	0.0312	-0.2052	0.0228
4	Male	-0.0108	0.0294	-0.1640	0.0201
5	Transfer Student	-0.3557	0.0503	-0.0041	0.0401
6	STEM fixed-effect	—	—	—	—
7	Non-STEM fixed effect	—	—	—	—
8	Winter Quarter	-0.5687	0.0350	-0.1303	0.0164
9	Spring Quarter	-0.2646	0.0269	0.1169	0.0237
	STEM GPA	0.2886	0.0548	-0.3543	0.0589
10	Any non-STEM classes	0.1986	0.0887	0.0659	0.0671
	Non-STEM GPA	2.3083	0.0239	-0.0307	0.1031
11	Admitted into Physical	-0.0995	0.0373	-0.0736	0.0263
12	Admitted into Engineering	-0.1418	0.0354	-0.3023	0.0260
13	Any STEM classes	-0.5665	0.0596	-0.4612	0.0522
14	Any upper classes	-0.8262	0.0551	-0.6471	0.0433
15	Quarter $\times$ SAT math /1000	-0.0140	0.1499	-0.1092	0.1085
16	Quarter $\times$ gender	-0.1045	0.1349	-0.0989	0.0981
17	Quarter $\times$ race	-0.0956	0.1501	0.0140	0.1041
18	Zipcode Income /1000	0.0001	0.0005	0.0009	0.0003
19	Term 2	3.4188	0.2366	0.3869	0.0658
20	Term 3	4.1885	0.2223	0.3079	0.0773
21	Term 4	4.6552	0.1856	0.2387	0.0481
22	Term 5	4.3413	0.2122	0.8111	0.0704
23	Term 6	4.6466	0.2255	0.7478	0.0774
24	Term 7	5.1167	0.1759	0.3295	0.0814
25	Term 8	5.0521	0.1864	0.5377	0.1267
26	Term 9	5.2496	0.1836	0.2942	0.1090
27	Term 10	5.3516	0.1671	0.0037	0.1347
28	Term 11	4.6623	0.2016	0.6940	0.0868
29	Term 12	4.7772	0.2280	0.7816	0.0855
30	Term 13	4.9561	0.1966	0.5902	0.0831
31	Term 14	4.1272	0.3428	1.0087	0.0892
32	Term 15	3.0454	2.9418	1.1862	0.0844
33	Term 16	5.4705	0.2118	0.3228	0.3831
34	Term 17	4.9235	0.2926	1.0529	0.1696
35	Term 18 or more	5.7068	0.1830	0.3570	0.2711
	Goup Fraction in population	0.5201	0.0224	0.4799	0.0224

Table 9b: Summary Statistics by Estimated Type From Two-Type GPA Model

Variable	type I	type II
Asian	0.44	0.52
Minority	0.15	0.16
Male	0.58	0.58
Transfer	0.00	0.00
Zipcode Income	66,013	64,826
Admit physical	0.12	0.11
Admit engineering	0.30	0.25
STEM fixed effect	-0.04	-0.23
Non-STEM fixed effect	0.09	-0.02
SAT math score	652.30	639.24
SAT verbal score	551.40	535.62
high school GPA	3.90	3.80
took any non-STEM classes	0.98	0.97
STEM GPA	2.89	2.66
non-STEM GPA	3.14	2.93

Note: Summary statistics are obtained by assigning individuals to their highest likelihood type.

Table 10a: Transition out of STEM Major—Three-Type Model with STEM and Non-STEM GPA's

No.	Variable	Type I		Type II		Type III	
		Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
1	Constant	-2.8212	0.2757	-1.1061	0.1237	-4.7753	1.3099
2	Minority	-0.0982	0.0420	-0.1396	0.0492	-0.1061	0.0467
3	Asian	-0.0513	0.0330	-0.2719	0.0435	-0.1965	0.0396
4	Male	-0.0148	0.0268	-0.1339	0.0386	-0.2124	0.0345
5	Transfer Student	-0.0686	0.0551	-0.3611	0.0740	-0.1879	0.0594
6	STEM fixed-effect	—	—	—	—	—	—
7	Non-STEM fixed effect	—	—	—	—	—	—
8	Winter Quarter	-0.7669	0.0045	-0.1400	0.0249	-0.0181	0.0177
9	Spring Quarter	-0.1510	0.0311	0.0811	0.0398	0.1291	0.0290
10	STEM GPA	0.2408	0.0262	0.3430	0.0909	-0.7772	0.1005
11	Any non-STEM classes	0.2120	0.0842	0.1951	0.3083	-0.0738	0.1321
12	Non-STEM GPA	0.4803	0.1418	0.2410	0.1455	2.1651	1.7587
13	Admitted into Physical	-0.1302	0.0423	0.0311	0.0503	-0.1236	0.0440
14	Admitted into Engineering	-0.2244	0.0399	0.2408	0.0450	-0.6918	0.0112
15	Any STEM classes	-1.1503	0.0286	-0.3304	0.1126	-0.0559	0.0713
16	Any upper classes	0.0461	0.0651	-3.2891	0.6262	-0.9365	0.0870
17	Quarter ✕ SAT math /1000	0.0779	0.1548	0.1351	0.2589	-0.3811	0.1800
18	Quarter ✕ gender	-0.1186	0.1402	-0.2451	0.2403	0.0692	0.1611
19	Quarter ✕ race	-0.1769	0.1468	-0.2484	0.2401	0.2515	0.1680
	Zipcode Income /1000	0.0002	0.0005	0.0000	0.0006	0.0016	0.0005
20	Term 2	2.9812	0.3094	-0.6463	0.1146	1.5886	0.1972
21	Term 3	3.7868	0.2689	-0.3758	0.3811	1.3425	0.1980
22	Term 4	4.2861	0.3052	-0.6934	0.2437	1.2356	0.1860
23	Term 5	4.0663	0.3245	-0.7334	0.1844	2.1927	0.2142
24	Term 6	4.0941	0.2768	-0.1045	0.3370	2.0246	0.2871
25	Term 7	4.0516	0.2873	0.4274	0.0873	1.5947	0.2753
26	Term 8	3.9537	0.2998	0.4062	0.1286	2.0374	0.2416
27	Term 9	4.0535	0.2657	0.7718	0.2362	1.5758	0.1992
28	Term 10	4.0045	0.2801	1.4120	0.0706	-0.0871	0.5360
29	Term 11	3.5931	0.3111	-0.0378	0.2677	2.3704	0.3297
30	Term 12	3.7843	0.2558	-0.0167	0.8486	2.3538	0.3338
31	Term 13	3.8524	0.3020	0.0107	0.6754	2.0840	0.3674
32	Term 14	3.5443	0.3222	-2.0764	0.4399	2.8128	0.3387
33	Term 15	3.8621	0.3234	-1.9587	1.3437	2.8190	0.3380
34	Term 16	4.2170	0.2743	0.6408	0.3128	1.9089	0.4789
35	Term 17	3.7534	0.2644	1.2024	0.3642	2.6553	0.4010
36	Term 18 or more	4.2927	0.2796	0.8048	0.3945	2.2038	0.3672
37	Group Fraction in population	0.4378	0.0101	0.2801	0.0149	0.2821	0.2204

Table 10b: Summary Statistics by Estimated Type From Three-Type GPA Model

Variable	Type I	Type II	Type III
Asian	0.49	0.52	0.42
Minority	0.12	0.19	0.20
Male	0.61	0.59	0.50
Transfer	0.00	0.01	0.00
Zipcode Income	66279.94	63984.35	65048.90
Admit physical	0.12	0.12	0.10
Admit engineering	0.36	0.21	0.15
STEM fixed effect	0.02	-0.36	-0.23
Non-STEM fixed effect	0.14	-0.10	-0.03
SAT math score	662.48	629.56	626.93
SAT verbal score	552.69	526.71	540.90
high school GPA	3.94	3.69	3.82
took any non-STEM classes	0.98	0.96	0.99
STEM GPA	2.99	2.51	2.61
non-STEM GPA	3.15	2.81	3.00

Note: Summary statistics are obtained by assigning individuals to their highest likelihood type.

|

**Table 11: Statistics for Joint Distribution of STEM and Non-STEM Fixed-Effect
for Selected Sub-Groups**

Group	Mean		Median		Standard Dev.		Corr.
	STEM	n-STEM	STEM	n-STEM	STEM	n-STEM	
Entire population							
All	-0.1294	0.0534	-0.1037	0.1019	0.6151	0.4557	0.7412
Stayers	0.0758	0.1814	0.0888	0.2028	0.4988	0.3795	0.6711
Quitters	-0.4395	-0.1401	-0.4712	-0.1058	0.6438	0.4916	0.7272
White							
All	0.0089	0.1564	0.0435	0.2009	0.5731	0.4316	0.7077
Stayers	0.1659	0.2533	0.1952	0.2699	0.4801	0.3700	0.6534
Quitters	-0.2394	0.0032	-0.2409	0.0514	0.6187	0.4756	0.7011
Asian							
All	-0.0920	0.0607	-0.0657	0.1046	0.5854	0.4385	0.7325
Stayers	0.0694	0.1624	0.0832	0.1884	0.4917	0.3768	0.6673
Quitters	-0.4101	-0.1395	-0.4356	-0.1077	0.6238	0.4806	0.7304
Minority							
All	-0.5662	-0.2057	-0.5940	-0.1753	0.6070	0.4633	0.7098
Stayers	-0.2219	0.0232	-0.2523	0.0204	0.4877	0.3690	0.6407
Quitters	-0.7768	-0.3457	-0.8017	-0.3272	0.5759	0.4595	0.6559
Minority up to 1997							
All	-0.5757	-0.2138	-0.6093	-0.1910	0.5998	0.4615	0.7036
Stayers	-0.2350	0.0113	-0.2672	0.0098	0.4768	0.3785	0.6270
Quitters	-0.7711	-0.3429	-0.7994	-0.3274	0.5753	0.4553	0.6581
Minority from 1998							
All	-0.5211	-0.1673	-0.5256	-0.1102	0.6388	0.4703	0.7351
Stayers	-0.1721	0.0684	-0.2050	0.0793	0.5252	0.3270	0.7047
Quitters	-0.8077	-0.3610	-0.8027	-0.3251	0.5788	0.4819	0.6451
Female population							
All	-0.1700	0.0700	-0.1430	0.1155	0.6199	0.4209	0.7647
Stayers	0.0615	0.2077	0.0750	0.2252	0.5009	0.3440	0.6989
Quitters	-0.4276	-0.0832	-0.4505	-0.0477	0.6378	0.4452	0.7492
Male population							
All	-0.1013	0.0420	-0.0776	0.0906	0.6102	0.4780	0.7346
Stayers	0.0837	0.1668	0.0954	0.1891	0.4975	0.3973	0.6632
Quitters	-0.4508	-0.1940	-0.4909	-0.1657	0.6492	0.5263	0.7183

**Table 12: Statistics for Joint Distribution of STEM and Non-STEM Fixed-Effect
by STEM Fields**

Group	Mean		Median		Standard Dev.		Corr.
	STEM	n-STEM	STEM	n-STEM	STEM	n-STEM	
Life Sciences							
Minority							
All	-0.6214	-0.2305	-0.6518	-0.1973	0.6057	0.4335	0.7307
Stayers	-0.2397	0.0072	-0.2824	0.0152	0.4934	0.3216	0.6866
Quitters	-0.7814	-0.3302	-0.7935	-0.3163	0.5763	0.4357	0.6879
Non-Minority							
All	-0.0516	0.0878	-0.0165	0.1445	0.5892	0.4025	0.7677
Stayers	0.1390	0.1879	0.1765	0.2160	0.4962	0.3279	0.7433
Quitters	-0.2973	-0.0412	-0.3008	-0.0015	0.6085	0.4503	0.7469
Engineering							
Minority							
All	-0.4490	-0.0746	-0.4341	-0.0195	0.6000	0.5297	0.6828
Stayers	-0.2076	0.1015	-0.2055	0.0885	0.4682	0.3894	0.6052
Quitters	-0.9438	-0.4357	-0.9983	-0.3802	0.5339	0.5932	0.5372
Non-Minority							
All	-0.0667	0.1491	-0.0372	0.1949	0.5802	0.4811	0.6933
Stayers	0.0621	0.2310	0.0668	0.2467	0.4802	0.4132	0.6209
Quitters	-0.5753	-0.1743	-0.6396	-0.1314	0.6569	0.5834	0.6921
Physical Sciences							
Minority							
All	-0.5061	-0.2932	-0.5699	-0.2788	0.5939	0.4426	0.7023
Stayers	-0.1958	-0.1642	-0.2486	-0.1700	0.5259	0.4063	0.7443
Quitters	-0.6421	-0.3497	-0.6832	-0.3368	0.5710	0.4466	0.6707
Non-Minority							
All	-0.0001	0.0254	0.0107	0.0568	0.5572	0.4464	0.7255
Stayers	0.1488	0.1203	0.1796	0.1426	0.4816	0.4074	0.7064
Quitters	-0.1891	-0.0949	-0.1832	-0.0573	0.5886	0.4645	0.7059

Figure 1: Admit Major for Minority and Non-Minority Students by Division

Figure 2: Graduating Major for Minority and Non-Minority Students by Division

Figure 3: Fraction of Minority by Admit and Graduating Major

Figure 4: Average Math SAT for Admitted Students

Figure 5: Average Verbal SAT for Admitted Students

Figure 6: Average Math SAT for Graduating Minority and Non-Minority Students

Figure 7: Average Verbal SAT for Graduating Minority and Non-Minority Students

Figure 8: Fraction Retained in Major for 4-Year and 5-Year Graduates

Figure 9: STEM GPA for Stayers and Quitters STEM Students

Figure 10: Mean SAT for Minority and Non-Minority STEM Students, by Quitting Term

Figure 11: Average Course Grade For Minority and Non-Minority Students

Figure 12: Joint Density of STEM and Non-STEM Fixed-Effects for the Entire Population

Figure 13: Joint Density of STEM and Non-STEM Fixed-Effects for Minority Students

Figure 14: Joint Density of STEM and Non-STEM Fixed-Effects Minority Students, up to 1997

Figure 15: Joint Density of STEM and Non-STEM Fixed-Effects

Figure 16: Joint Density of STEM and Non-STEM Fixed-Effects

Life Sciences Minority Students

Figure 17: Joint Density of STEM and Non-STEM Fixed-Effects

Figure 18: Joint Density of STEM and Non-STEM Fixed-Effects

Figure 19: Joint Density of STEM and Non-STEM Fixed-Effects

Figure 20: Joint Density of STEM and Non-STEM Fixed-Effects

Figure 21: Joint Density of STEM and Non-STEM Fixed-Effects

Appendix 1a: Multinomial logit regression results. Probability of graduating STEM within 4yrs.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
asian	0.0112** (0.00549)	0.00313 (0.00547)	0.0124** (0.00507)	0.0290*** (0.00516)	0.0182*** (0.00516)	0.00630 (0.00616)
minority	-0.180*** (0.00522)	-0.112*** (0.00750)	-0.0803*** (0.00712)	-0.0943*** (0.00671)	-0.0729*** (0.00746)	-0.0578*** (0.00763)
male	-0.0186*** (0.00526)	-0.0547*** (0.00571)	-0.0422*** (0.00526)	-0.0216*** (0.00510)	-0.0404*** (0.00531)	-0.0542*** (0.00613)
transfer		-0.135*** (0.00573)	-0.128*** (0.00510)	-0.121*** (0.00643)	-0.130*** (0.00522)	-0.127*** (0.00459)
Median HH income		9.57e-05 (9.10e-05)	5.70e-05 (8.36e-05)	0.000224*** (8.46e-05)	4.11e-05 (8.42e-05)	7.32e-05 (8.01e-05)
admit physical		-0.0209*** (0.00792)	-0.0139* (0.00730)	-0.0151** (0.00740)	-0.0171** (0.00724)	-0.0447*** (0.00608)
admit engineering		-0.0509*** (0.00543)	-0.0239*** (0.00523)	-0.00831 (0.00543)	-0.0257*** (0.00538)	-0.0919*** (0.00507)
STEM ability				0.201*** (0.00648)	0.0920*** (0.00932)	0.0840*** (0.00893)
non-STEM ability				0.0354*** (0.00897)	0.0216** (0.00965)	0.0157* (0.00907)
SAT math score		0.00105*** (4.24e-05)	0.000493*** (3.94e-05)		0.000477*** (4.01e-05)	0.000135*** (4.09e-05)
SAT verbal score		0.000108*** (3.22e-05)	2.27e-05 (2.89e-05)		3.30e-06 (2.97e-05)	1.69e-05 (2.76e-05)
High School GPA		0.0274*** (0.00866)	-0.0139*** (0.00485)		-0.0161*** (0.00481)	-0.0215*** (0.00450)
took non STEM classes			-0.341*** (0.0438)		-0.190*** (0.0432)	-0.101*** (0.0353)
first yr STEM GPA			0.151*** (0.00467)		0.0832*** (0.00720)	0.0805*** (0.00696)
first yr non STEM GPA			0.0486*** (0.00569)		0.0296*** (0.00640)	0.0170*** (0.00598)
first yr peers (SAT math)						0.00277*** (0.000115)
first yr fraction own race						0.00219 (0.0222)
first yr fraction own gender						0.0355* (0.0214)
Observations	22,848	22,848	22,848	22,848	22,848	22,848

Note: This table reports the marginal effects of variables on the probability of graduating STEM within 4 yrs. Estimates are derived from a multinomial logit regression with three outcomes: graduate STEM within 4 yrs, graduate non-stem within 4yrs, do not graduate within 4 yrs. The regression is restricted to student who started their career as a STEM major. Robust standard errors in brackets. Regressions include year fixed effects (not shown). The median household income variable is derived from the student's zip code and is expressed in thousands of dollars. *** p<0.01, ** p<0.05, * p<0.1

Appendix 1b: Multinomial logit regression results. Probability of graduating non-STEM within 4yrs.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
asian	-0.0687*** (0.00443)	-0.0552*** (0.00434)	-0.0556*** (0.00444)	-0.0593*** (0.00446)	-0.0462*** (0.00432)	-0.0380*** (0.00489)
minority	-0.0532*** (0.00442)	-0.0442*** (0.00493)	-0.0464*** (0.00497)	-0.0413*** (0.00497)	-0.0370*** (0.00513)	-0.0411*** (0.00487)
male	-0.0941*** (0.00458)	-0.0612*** (0.00458)	-0.0595*** (0.00469)	-0.0594*** (0.00458)	-0.0530*** (0.00456)	-0.0427*** (0.00437)
transfer		-0.0982*** (0.00423)	-0.0979*** (0.00446)	-0.1000*** (0.00410)	-0.0941*** (0.00449)	-0.0870*** (0.00437)
median HH income		0.000314*** (6.88e-05)	0.000335*** (6.97e-05)	0.000347*** (6.94e-05)	0.000296*** (6.82e-05)	0.000258*** (6.66e-05)
admit physical		0.00481 (0.00590)	0.00848 (0.00613)	0.00577 (0.00603)	0.00416 (0.00584)	0.0186*** (0.00636)
admit engineering		-0.0845*** (0.00426)	-0.0835*** (0.00443)	-0.0945*** (0.00417)	-0.0840*** (0.00437)	-0.0405*** (0.00536)
STEM ability				-0.0233*** (0.00547)	0.0712*** (0.00759)	0.0679*** (0.00737)
non-STEM ability				0.0812*** (0.00721)	0.0751*** (0.00820)	0.0806*** (0.00812)
SAT math score		-0.000129*** (3.03e-05)	-5.27e-05 (3.21e-05)		-7.18e-05** (3.16e-05)	0.000170*** (3.34e-05)
SAT verbal score		0.000226*** (2.51e-05)	0.000211*** (2.59e-05)		0.000160*** (2.60e-05)	0.000148*** (2.60e-05)
High School GPA		-0.00257 (0.00438)	0.00367 (0.00485)		-0.00164 (0.00452)	0.00512 (0.00472)
took non STEM classes			0.0781*** (0.0139)		0.0984*** (0.00494)	0.0868*** (0.00727)
first yr STEM GPA			-0.0315*** (0.00369)		-0.0842*** (0.00539)	-0.0813*** (0.00526)
first yr non STEM GPA			0.0271*** (0.00448)		-0.0151*** (0.00529)	-0.00554 (0.00533)
first yr peers (SAT math)						-0.00167*** (8.70e-05)
first yr fraction own race						0.0138 (0.0164)
first yr fraction own gender						-0.0416** (0.0168)
Observations	22,848	22,848	22,848	22,848	22,848	22,848

Note: This table reports the marginal effects of variables on the probability of graduating non-STEM within 4 yrs. Estimates are derived from a multinomial logit regression with three outcomes: graduate STEM within 4 yrs, graduate non-stem within 4yrs, do not graduate within 4 yrs. The regression is restricted to student who started their career as a STEM major. Robust standard errors in brackets. Regressions include year fixed effects (not shown). The median household income variable is derived from the student's zip code and is expressed in thousands of dollars. *** p<0.01, ** p<0.05, * p<0.1

Appendix 1c: Multinomial logit regression results. Probability of not graduating within 4yrs.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
asian	0.0575*** (0.00664)	0.0521*** (0.00675)	0.0432*** (0.00658)	0.0303*** (0.00670)	0.0280*** (0.00664)	0.0317*** (0.00759)
minority	0.234*** (0.00663)	0.156*** (0.00862)	0.127*** (0.00840)	0.136*** (0.00812)	0.110*** (0.00880)	0.0989*** (0.00885)
male	0.113*** (0.00635)	0.116*** (0.00683)	0.102*** (0.00667)	0.0810*** (0.00657)	0.0933*** (0.00672)	0.0969*** (0.00726)
transfer		0.233*** (0.00700)	0.226*** (0.00670)	0.221*** (0.00758)	0.224*** (0.00685)	0.214*** (0.00626)
median HH income		-0.000410*** (0.000111)	-0.000392*** (0.000107)	-0.000571*** (0.000108)	-0.000337*** (0.000107)	-0.000331*** (0.000103)
admit physical		0.0161* (0.00956)	0.00538 (0.00929)	0.00932 (0.00935)	0.0130 (0.00919)	0.0261*** (0.00871)
admit engineering		0.135*** (0.00664)	0.107*** (0.00665)	0.103*** (0.00668)	0.110*** (0.00677)	0.132*** (0.00724)
STEM ability				-0.177*** (0.00803)	-0.163*** (0.0116)	-0.152*** (0.0112)
non-STEM ability				-0.117*** (0.0109)	-0.0967*** (0.0122)	-0.0963*** (0.0117)
SAT math score		-0.000924*** (4.91e-05)	-0.000440*** (4.83e-05)		-0.000405*** (4.90e-05)	-0.000305*** (5.10e-05)
SAT verbal score		-0.000334*** (3.88e-05)	-0.000234*** (3.75e-05)		-0.000164*** (3.84e-05)	-0.000165*** (3.71e-05)
High School GPA		-0.0249*** (0.00886)	0.0102 (0.00627)		0.0177*** (0.00617)	0.0163*** (0.00609)
took non STEM classes			0.263*** (0.0426)		0.0917** (0.0433)	0.0141 (0.0360)
first yr STEM GPA			-0.119*** (0.00561)		0.00102 (0.00866)	0.000862 (0.00844)
first yr non STEM GPA			-0.0757*** (0.00685)		-0.0145* (0.00798)	-0.0115 (0.00772)
first yr peers (SAT math)						-0.00110*** (0.000138)
first yr fraction own race						-0.0160 (0.0267)
first yr fraction own gender						0.00605 (0.0269)
Observations	22,848	22,848	22,848	22,848	22,848	22,848

Note: This table reports the marginal effects of variables on the probability of not graduating within 4 yrs. Estimates are derived from a multinomial logit regression with three outcomes: graduate STEM within 4 yrs, graduate graduate within 4 yrs. The regression is restricted to student who started their career as a STEM major. Robust standard errors non-stem within 4yrs, do not in brackets. Regressions include year fixed effects (not shown). The median household income variable is derived from the student's zip code and is expressed in thousands of dollars. *** p<0.01, ** p<0.05, * p<0.1

Appendix 2a: Transition out of STEM Major—Two-Type Model with STEM and Non-STEM Fixed-Effects

No.	Variable	Type I		Type II	
		Coefficient	St. Error	Coefficient	St. Error
1	Constant	-7.1900	0.4371	-1.3222	0.2076
2	Minority	-0.0897	0.0317	-0.1155	0.0355
3	Asian	-0.0822	0.0257	-0.2411	0.0306
4	Male	-0.0603	0.0265	-0.2073	0.0245
5	Transfer Student	0.0001	0.1632	-0.0756	0.2073
6	STEM fixed-effect	-0.5045	0.0687	0.1814	0.0409
7	Non-STEM fixed effect	-0.1512	0.0356	-0.1874	0.0447
8	Winter Quarter	0.3090	0.0761	-0.5013	0.0811
9	Spring Quarter	0.1529	0.1256	0.0506	0.1352
	STEM GPA	—	—	—	—
10	Any non-STEM classes	1.1251	0.1609	0.3795	0.0781
	Non-STEM GPA	—	—	—	—
11	Admitted into Physical	-0.0468	0.0301	-0.1227	0.0328
12	Admitted into Engineering	-0.1434	0.0356	-0.3083	0.0346
13	Any STEM classes	-0.7488	0.1429	-0.3064	0.1450
14	Any upper classes	-0.7298	0.1222	-0.6867	0.1466
15	Quarter ✖ SAT math /1000	-0.1199	0.1604	-0.1679	0.1344
16	Quarter ✖ gender	0.0016	0.1450	-0.0390	0.1201
17	Quarter ✖ race	-0.1251	0.1420	-0.0551	0.1255
18	Zipcode Income /1000	0.0001	0.0004	0.0008	0.0004
19	Term 2	3.9623	0.4871	0.4852	0.0883
20	Term 3	4.6416	0.5950	0.2911	0.1491
21	Term 4	5.0002	0.5258	0.1731	0.0695
22	Term 5	4.7251	0.5027	0.8904	0.1014
23	Term 6	5.1092	0.5219	0.7055	0.1066
24	Term 7	5.3400	0.4997	0.2776	0.1811
25	Term 8	5.1427	0.4475	0.7474	0.2308
26	Term 9	5.4831	0.5186	0.1055	0.3147
27	Term 10	5.4573	0.4757	-0.0834	0.4522
28	Term 11	4.7409	0.4589	0.8638	0.1243
29	Term 12	4.9850	0.5675	0.8554	0.1097
30	Term 13	5.0875	0.4530	0.6405	0.2098
31	Term 14	4.1352	0.4711	1.2546	0.1199
32	Term 15	4.3217	0.5720	1.3499	0.1030
33	Term 16	5.3554	0.4796	0.7481	0.3593
34	Term 17	5.0473	0.4861	1.2041	0.2109
35	Term 18 or more	5.2484	0.5599	1.2903	0.2505
	Group Fraction in population	0.6289	0.0513	0.3711	0.0513

Appendix 2b: Summary Statistics by Estimated Type From Two-type Fixed Effects Model

Variable	type I	type II
Asian	0.46	0.51
Minority	0.15	0.17
Male	0.58	0.57
Transfer	0.00	0.01
Zip code Income	65,842	64,793
Admit physical	0.13	0.11
Admit engineering	0.31	0.22
STEM fixed effect	-0.04	-0.27
Non-STEM fixed effect	0.08	-0.03
SAT math score	652.52	635.95
SAT verbal score	547.19	537.87
high school GPA	3.90	3.77
took any non-STEM classes	0.98	0.97
STEM GPA	2.88	2.63
non-STEM GPA	3.09	2.94

Note: Summary statistics are obtained by assigning individuals to their highest likelihood type.

Appendix 3a: Transition out of STEM Major—Three-Type Model with STEM and Non-STEM Fixed-Effects

No.	Variable	Type I		Type II		Type III	
		Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
1	Constant	-4.3532	2.3356	-1.4362	0.2214	-4.4659	2.5241
2	Minority	-0.1254	0.0369	-0.1110	0.0682	-0.0614	0.0408
3	Asian	-0.1082	0.0297	-0.3891	0.0551	-0.1683	0.0366
4	Male	-0.0423	0.0218	-0.2352	0.0602	-0.2145	0.0323
5	Transfer Student	-0.0697	0.0488	-0.2487	0.1280	-0.2025	0.0491
6	STEM fixed-effect	-0.6263	0.0297	0.5130	0.0642	0.0323	0.0337
7	Non-STEM fixed effect	-0.1958	0.0363	-0.4645	0.0753	0.0190	0.0472
8	Winter Quarter	0.3631	0.0396	0.2346	0.0359	-0.7130	0.0558
9	Spring Quarter	0.1401	0.0814	0.1937	0.0833	-0.0233	0.0689
	STEM GPA	—	—	—	—	—	—
10	Any non-STEM classes	0.1884	0.1148	0.6240	0.0198	1.8918	2.5016
	Non-STEM GPA	—	—	—	—	—	—
11	Admitted into Physical	-0.0498	0.0360	-0.0888	0.0658	-0.1190	0.0329
12	Admitted into Engineering	-0.1069	0.0342	0.2332	0.0604	-0.6299	0.0857
13	Any STEM classes	-1.2234	0.0756	-0.0343	0.1460	-0.0201	0.0393
14	Any upper classes	-0.1897	0.0349	-4.4390	0.1938	-0.8413	0.0511
15	Quarter $\times$ SAT math /1000	0.0241	0.1532	0.0879	0.0754	-0.3930	0.1562
16	Quarter $\times$ gender	-0.0869	0.1399	-0.1255	0.1757	0.0404	0.1425
17	Quarter $\times$ race	-0.1743	0.1490	-0.2865	0.3202	0.2889	0.1491
18	Zipcode Income /1000	-0.0001	0.0004	0.0000	0.0008	0.0013	0.0005
19	Term 2	2.3810	2.2870	-0.5429	0.1039	1.7142	0.2742
20	Term 3	3.1519	2.3941	-0.3694	0.1442	1.5139	0.2846
21	Term 4	3.4827	2.3975	-0.9196	0.3353	1.5416	0.2725
22	Term 5	3.1677	2.3130	-1.2803	0.4173	2.4655	0.2669
23	Term 6	3.4923	2.3907	-0.3096	0.2465	2.1876	0.2769
24	Term 7	3.4526	2.3912	0.4767	0.1264	1.7958	0.2818
25	Term 8	3.1761	2.3209	0.6216	0.1212	2.3402	0.2830
26	Term 9	3.5089	2.4051	1.0605	0.1759	1.7128	0.3024
27	Term 10	3.3791	2.4271	3.0432	0.2126	-0.7295	2.6529
28	Term 11	2.8366	2.2759	-1.2547	0.1832	2.5529	0.2514
29	Term 12	3.2273	2.3364	-1.1851	0.2798	2.4305	0.2648
30	Term 13	3.1458	2.3208	-1.1254	0.2789	2.2905	0.2672
31	Term 14	2.4108	2.1182	-1.3198	1.3984	2.9811	0.2197
32	Term 15	2.8433	2.1690	-1.2958	0.7706	2.9749	0.2395
33	Term 16	3.3937	2.3045	-1.1214	0.7562	2.5438	0.3270
34	Term 17	3.0792	2.2713	-1.1372	0.3573	3.0345	0.2819
35	Term 18 or more	3.2693	2.2961	-1.0188	1.1482	3.1010	0.2804
	Group Fraction in population	0.5170	0.1493	0.1709	0.0333	0.3121	0.3742

Appendix 3b: Summary Statistics by Estimated Type From Three-Type Fixed Effects Model

Variable	Type I	Type II	Type III
Asian	0.49	0.51	0.43
Minority	0.13	0.18	0.20
Male	0.61	0.58	0.51
Transfer	0.00	0.01	0.00
Zip code Income	66,156	64,189	64,899
Admit physical	0.13	0.11	0.10
Admit engineering	0.36	0.19	0.17
STEM fixed effect	0.02	-0.36	-0.26
Non-STEM fixed effect	0.13	-0.08	-0.06
SAT math score	660.97	629.38	627.11
SAT verbal score	549.92	531.87	539.55
high school GPA	3.93	3.68	3.83
took any non-STEM classes	0.98	0.96	0.98
STEM GPA	2.96	2.52	2.60
non-STEM GPA	3.12	2.88	2.99

Note: Summary statistics are obtained by assigning individuals to their highest likelihood type.