AN UNDERGRADUATE
JOURNAL OF ECONOMICS

The University of Texas at Austin

The Developing Economist

An Undergraduate Journal of Economics
The University of Texas at Austin
Volume 3

2016

Two Roads in a Wood: An Econometric Analysis of the Major Choice of First-Generation College Students

Sam Trejo[†]

Abstract

Using data from the National Longitudinal Survey of Youth 1997, I estimate a multinomial logit choice model for the college major decisions of first-generation college students—students who are the first in the families to attend college—and non-first-generation students. The model controls for other factors such as sex, race, ability, and family income to isolate the effect of first-generation status on major choice for two otherwise identical students. I find that first-generation college students do make statistically different college major selections than otherwise identical students. I then examine whether the estimated differences between the major selection of first-generation and non-first-generation students is systematically related to characteristics of the majors. In particular, I use data extracted from the American Community Survey to create these measures of safety and stability. First-generation college students tend to be more risk averse than otherwise identical non-first-generation students whose parents have attended college, as they are more likely to select majors with well-defined career paths, high expected wages, and low unemployment rates.

[†]I cannot thank my advisor Dr. Gerald Oettinger and my second reader Dr. Robert Crosnoe enough for their mentorship and guidance. In addition, special thanks to Dr. Brian Duncan at The University of Colorado at Denver and Dr. Joseph Altonji at Yale University for methodological assistance.

I. Introduction

In the fall of 2015, over 13 million American students attended 4-year universities[1]. Needless to say, not all 13 million students attended similar colleges or received a similar academic experience. An incoming freshman might decide to enroll at Arizona State University with annual tuition of about $10,000, over 65,000 undergraduates, and a mix of small and large lectures classes to receive a B.S. in computer science. Contrastingly, another incoming freshmen might instead enroll at Vassar College, with an annual tuition of almost $50,000, under 2,500 undergraduates, and small, discussion based classes to receive a B.A. in Philosophy. The wide range of educational institutions, degrees, and majors present in the United States illustrates a diversity in both the pedagogical philosophies and methodologies held by higher education providers and the resources, preferences, and educational interests of individual students.

To many students, a college education represents a transformative life experience. As Plato wrote over 2000 years ago, "the object of education is to teach us to love what is beautiful," a sentiment echoed on the University of Texas Plan II Honors program's website. There the interdisciplinary liberal arts program is advertised as "an education without boundaries," one "for a life, not for a living." To other students, a college education simply represents a practical investment in their human capital, affording degree holders higher wages and other rewards in the labor market. In 1967, then Governor of California Ronald Reagan embodied this outlook when he famously criticized the University of California system for offering courses on organizing social movements, arguing there was no value in "subsidizing intellectual curiosity" at universities.

Most American students likely view their college education as a healthy compromise of the positions represented by

[1] "College Enrollment and Work Activity of 2014 High School Graduates." U.S. Bureau of Labor Statistics. U.S. Bureau of Labor Statistics, 16 Apr. 2015. Web. 26 Aug. 2015.

Plato and President Reagan. That is, students pursue education both because it adds value and texture to their lives and because it builds human capital that the labor market rewards. Though normative questions about the proper goals of "good" education are difficult to address concretely, the social and economic repercussions of the existence of a diverse set of higher educational opportunities, like the range of college major options, can be analytically explored.

A student's college major decision is not a trivial one; not only does each college major require a unique set of coursework and talents, but not all college degrees are equally valuable. While a college degree today may be worth more than ever before[2], expected earnings vary tremendously across different majors. According to a 2015 report titled *From Hard Times to Better Times*[3] from the Georgetown University Center on Education and the Workforce (GCEW), there exists a wide range in returns to education by major, with median mid-career full-time earnings at $29,000 for Counseling Psychology majors compared to $120,000 for Petroleum Engineering majors. Thus, studying the tendencies in college major choice across demographic groups can serve to deepen an understanding of economic and educational disparities present in America.

Economic and sociological research has long explored how disadvantaged students make decisions differently in regards to higher education when compared to other students. The different contingent factors in a student's life such as their family's monetary constraints, their information about different degree options, and their unique set of cultural expectations likely influence the type of college degree an individual elects to pursue. In particular, we would expect these factors within a student's life to affect his or her likelihood of selecting dif-

[2]According to the same GCEW Report, the ratio of the median wage between workers with a college degree and workers with only a high school education has steadily increased from the mid-1970s to present day. College graduates now make over two times more than high school graduates.

[3]Carnevale, Anthony, and Ban Cheah. "From Hard Times to Better Times." *Center on Education and the Workforce.* Georgetown University, 05 Feb. 2015.

ferent college majors. While some students may be drawn to the renaissance-style education offered by the Plan II Honors program, it may not be a practical or appealing option for all college-bound youths.

Due to the disparate returns to education across different major fields, research targeted at reducing economic and educational inequality should focus not only on investigating motivations of college attendance but also of the selection of a field of study. In this paper, I investigate how first-generation college students differ from other students in their choice of college major. Using data from the National Longitudinal Survey of Youth 1997 (NLSY97), I estimate a multinomial logit choice model for the college major decisions of first-generation college students and non-first-generation students. The model controls for other factors such as sex, race, ability, and family income to isolate the effect of first-generation status on major choice for two otherwise identical students. I find that first-generation college students—students who are the first in the families to attend college—do make statistically different college major selections than otherwise identical students.

I then examine whether the estimated differences in major selection between first-generation and non-first-generation students is systematically related to characteristics of the majors. In particular, I use data from the American Community Survey (ACS) to create measures of the economic safety and stability of different majors. I find that first-generation college students are more risk averse than otherwise identical students who have parents who have attended college and tend to select majors with a well-defined career path, high expected wages, and low expected unemployment.

II. Literature Review

A growing body of economic and sociological research explores how disadvantaged students—namely members of certain minority groups, those from low-income families, and first-generation college students—make decisions differently in regards to higher education when compared to other students.

For example, disadvantaged minorities are less likely than other groups to prepare themselves academically for college. Using a nationally representative sample of the National Educational Longitudinal Study (NELS) data set, Stage, Droogsma-Musoba, and Brown[4] (2002) found that, controlling for ability, Asian American and White students were more likely than Black, Hispanic, and Native American students to take mathematics courses that would prepare them for college. Furthermore, low-income students are significantly less likely to decide to even apply to college. Using the same data set, Cabrera and La Nasa[5] (2002) found that while 76% of high socioeconomic status students submitted applications to four-year colleges, only 21.3% of low socioeconomic status students applied to college.

Not only are students from these disadvantaged groups less likely to apply to college, but when they do apply they tend to under-match, or in other words, apply to colleges below their academic ability level. Using a recent sample of Texas high school graduates, Black, Cortez, and Lincove[6] (2014) investigated race and ethnicity differences in college application decision-making using a sample recent Texas high school graduates. They found that disadvantaged minorities exhibit different college application decisions than Whites and Asians with similar levels of academic achievement. Inequality in college access is magnified through these under-matching tendencies of minority high-performing students. This reinforces the findings of Hoxby and Christopher[7] (2012), who first demon-

[4]Stage, F. K., Droogsma-Musoba, G., & Brown, C. (2002, April). *Mathematics achievement: Racial-ethnicity and course taking patterns.* American Educational Research Association, New Orleans.

[5]Cabrera, A. F., & La Nasa, S. M. (2000a). Overcoming the tasks on the path to college for America's disadvantaged. In A. Cabrera & S. La Nasa (Eds.), *Understanding the college choice of disadvantaged students.* New Directions for Institutional Research, No. 107, pp. 31-44. San Francisco: Jossey-Bass.

[6]Black, Sandra, Kalena Cortes, and Jane Lincove. "You Have to Apply Yourself: Racial and Ethnic Differences in College Application." Working Paper (2014). 1 Jan. 2015.

[7]Hoxby, Caroline, and Christopher Avery. "The Missing "One-Offs":

strated the tendency of high-achieving low-income students to choose not to apply to selective colleges.

While the behavior of disadvantaged students leading up to and navigating through higher education is increasingly the subject of academic research, only minor inquiry has been made regarding how these disadvantaged groups tend to select a field of study. Saks and Shore[8] (2005) estimated the risk associated with different careers and found education, health care, and engineering careers to have relatively safe streams of labor income; on the other hand, business, sales, and entertainment careers are more risky. Controlling for observable measures of ability and demographic background, they showed that students from low-income families tended to select majors that would result in less risky careers than high-income students.

That little economic research on college major selection has been done is surprising, given the strong connection between college major and labor market outcomes. Recent survey information has provided social science researchers with new data linking one's college major to future employment and financial outcomes. Starting in 2009, the ACS began asking for its respondents to report their college field of study. Altonji, Blom, and Meghir[9] (2012) use this data to review literature on the heterogeneous nature of educational specializations and the link between college majors and occupational paths. Amazingly, they find that the "difference in returns across college majors rivals the college wage premium." After adjusting for basic demographics, work experience, and postsecondary degrees, the gap in log wages rates between male electrical engi-

The Hidden Supply of High-Achieving, Low Income Students." National Bureau of Economic Research (2012): Working Paper. Web.

[8]Saks, Raven E., and Stephen H. Shore. "Risk And Career Choice." B.E. Journal Of Economic Analysis & Policy: Advances In Economic Analysis & Policy 5.1 (2005): 1-45. Business Source Complete. Web. 20 Nov. 2015.

[9]Altonji, Joseph, Erica Blom, and Costas Meghir. "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers." Annual Review of Economics 4 (2012): 185-223. Web.

neering and male general education majors is a striking 0.56, nearly as large as the 0.57 difference between college graduates and high school graduates.

In this paper, I expand upon Saks and Shore's work by exploring the additional effect of first-generation status on student college major decisions. In doing so, I am able to separate out the impact of parental education from race and socioeconomic status. By better studying this important educational choice, my research adds to a greater body of work that aims to understand the decisions made by disadvantaged students while preparing for, pursuing, and specializing within higher education.

III. Theory

That low-income students tend to be more risk averse in their college major selection compared to middle and high-income of students is unsurprising. Post-secondary education in America is expensive; there exist substantial explicit academic costs to higher education (namely tuition). These explicit costs are heightened by the implicit opportunity cost of the forgone wages during the time spent in school.

Under perfect credit market conditions, a student might choose to borrow money now for the cost of college only to repay the money later with improved future labor market outcomes. However, human capital cannot be collateralized in the same way that other investments can be, making educational loans riskier than other types of lending. This credit market failure drives up interest rates for college loans, causing an increase in the marginal cost of education for the low-income students who are forced to take out loans. For this reason, attending college may only be a rational investment for low-income students wishing to specialize in a narrower, more lucrative set of college majors.

My research, however, focuses on the extent to which being a first-generation college student affects an individual's college major choice. First-generation college students are disproportionately from low-income families and members of disadvan-

taged minority groups in addition to facing a unique set of educational challenges. For these students, going to college is uncharted territory. First-generation college students are likely exposed to less or at least different information about college than those students whose parents have attended college. College, like most things, rewards students with both explicit financial returns and also intangible utility. First-generation college students may lack family member accounts of the non-financial rewards to a college education, leading them to give the expected labor market rewards of a major more relative weight when selecting their field of study.

In addition, because pursuing postsecondary education is atypical for their family and peer group, a first-generation student may be forced to more frequently justify his or her decision or articulate his or her post-graduation plans, thereby incentivizing a more pragmatic major selection. Preference may be given by first-generation college students to majors with higher wages and a more well-defined career path. For example, first-generation students might avoid studying areas such as communications or psychology and instead give preference to degrees in areas such as healthcare or engineering.

Being asked to answer the "Why college?" question might also alter the timeline of the college major decision of first-generation students. Many students enter their first year of college as undeclared students or adopt a major only tentatively, but first-generation students may experience pressure to select a major before they get to college. When a student makes their specialization could easily impact the major he or she selects.

Finally, first-generation students are a group comprised entirely of individuals who are is making a decision that is a departure from their familial and social norms. An individual's propensity to "go against the grain" and make such a choice could be the result of particular character traits or interests, which would therefore be more common in first-generation students than other students. This might result in the selection systematically different majors, though the major characteris-

tics that would be preferred remains unclear.

IV. Data Set

I begin by analyzing individual-level data from the National Longitudinal Study of Youth 1997 Cohort[10] (NLSY97). The NLSY97 consists of a nationally representative sample of approximately 9,000 youths who were between 12 and 16 years old as of December 31, 1996. The initial round of the survey took place in 1997. In that first round, both the eligible youth and one of the youth's parents received hour-long personal interviews. These youths were asked questions on their family background, education, work, and life decisions and were re-interviewed on an annual basis.

My sample consists of all the individuals in the NLSY97 who completed a 4-year college degree by 2010. Though interesting questions exist on a student's initial college major decision and its effect of major switching and drop-out rates, I limit my study to a student's final major selection. A student's final major can be thought of as their ultimate revealed preference and by using this as my metric of analysis I am able to draw connections between a student's college major and their expected labor market outcomes.

First I construct my college major variable COLLEGE_MAJOR. Though the NLSY97 has an already constructed major variable available based off of college transcripts, transcripts for many students were not collected. To increase my number of usable observations, I elect not to use this variable. Instead, I construct my own variable for college major using a self-reported "current college field of study" question that participants were asked on an annual basis. To do so, I order the annual responses from each individual chronologically and use each individual's final non-missing self-reported field of study for their COLLEGE_MAJOR value.

[10]The NLSY97 survey is sponsored and directed by the U.S. Bureau of Labor Statistics and conducted by the National Opinion Research Center at the University of Chicago, with assistance from the Center for Human Resource Research at The Ohio State University.

To obtain major variables each with a sufficient number of observations, I next group together similar COLLEGE_MAJORs to create 11 MAJOR_GROUP variables. In constructing these major groups, I worked to strike a sensible balance between groupings based on both similar expected labor market outcomes and related academic content. The few observations with major variables that did not fit well into any categories were dropped from my dataset. These dropped observations were *Home Economics, Automotive Mechanics, Transportation & Materials Moving, Security & Protective Services*, and *Uncodable*. The final major groupings of my NLSY97 sample set are displayed in Table 1.

Table 1 NLSY97 Major Groupings

MAJOR_GROUP	Number of Observations	NLSY97 Major Code (COLLEGE_MAJOR)
Arts	99	Fine & Applied Arts
Biology & Life Sciences	135	Agriculture & Natural Resources, Biological Sciences
Business	402	Business Management, Hotel & Hospitality Management
Communication	123	Communications
Computer Science & Math	104	Computer & Information Sciences, Mathematics
Education	194	Education
Physical Science & Engineering	128	Architecture & Environmental Design, Engineering, Other Sciences & Applied sciences, Physical sciences
Health	162	Nursing, Nutrition & Dietetics, Other Health Professions, Pre-Dental, Pre-Med, Pre-Vet
Humanities & Liberal Arts	205	Area Studies, English, Ethnic Studies, Foreign Languages, History, Interdisciplinary Studies, Philosophy, Theology & Religious Studies
Psychology & Social Work	153	Human Services, Psychology, Social Work
Social Science	276	Anthropology, Archaeology, Criminology, Economics, Geography, International Relations, Legal Services, Political Science, Pre-Law, Sociology
Total	1,981	

Next, I construct a FG_COLLEGE dummy variable for those individuals who never had a parent that attended college. For the purposes of this paper, I define first-generation student as one whose does not have a parent with greater than 12 years of schooling.

Finally, I define several key demographic variables for the individuals in my sample. I begin by creating a FEMALE dummy variable representing a female participant and a BLACK_HISPANIC dummy variable representing an individual who is Black or Hispanic. Next, I create a measure of

family income. Unfortunately, only the initial year of family income was usable for my analysis. In the year following the initial NLSY97 interview in 1997, a fraction of the youths in my sample set become emancipated. Any income reported thereafter is the individual's own personal income. Because comparing an 18 year old's personal income against 17 year old's family income would be unreflective of their actual relative financial situations, I limited my income measure to just family income in the year of 1997 and call this variable INCOME_1997. For a measure of cognitive ability, I use percentile scores from the Armed Forces Qualification Test (AFQT) to create an AFQT_PCT variable.

Unfortunately, a small number of observations are missing values for both AFQT_PCT and INCOME_1997; I drop these observations from my dataset. For any remaining observations, I then estimate any missing values for either AFQT_PCT or INCOME_1997 by forming predicted values based from OLS regressions with the other 4 variables as independent variables. For example, I estimated 168 missing AFQT_PCT values by using an OLS regression with INCOME_1997, FG_COLLEGE, FEMALE, and BLACK_HISPANIC independent variables.

Table 2 contains summary statistics of my NLSY97 sample. There are several substantial differences in characteristics between the first-generation and non-first-generation members of my sample, suggesting that it might be important to control for these differences in order to isolate the true effect of being a first-generation college student on major selection.

Table 2 NLSY97 Dataset Description

	FG_COLLEGE	Not FG_COLLEGE
Percent Female	67.25%	55.77%
Percent Hispanic or Black	47.37%	25.99%
AFQT Percentile Mean	58.58	70.48
AFQT Percentile Standard Deviation	24.88	22.89
1997 Family Income Mean	$45,985	$74,649
1997 Family Income Standard Deviation	$29,804	$54,224
Number of Observations	342	1,639

In addition to the NLSY97, I use data from the American Community Survey (ACS) to create variables measuring

wages, unemployment rates, and occupational concentrations for recent college graduates. The ACS is an ongoing statistical survey conducted by the U.S. Census Bureau, sent to over 3 million households each year. Beginning in 2009, the ACS began asking for its respondents to report their college field of study, providing researchers with unprecedented access to a large data set linking an individual's college major to their employment and financial outcomes.

I begin with an extract of the ACS that contains over 1.5 million observations from 2009, 2010, and 2011. Because I am primarily interested in measuring the labor market rewards for recent college graduates, I keep only observations on individuals between the ages of 22 and 26, leaving me with just over 60,000 observations[11].

Next I have to code the ACS field of study variable into my MAJOR_GROUP variable. I tried to be as consistent with my NLSY97 groupings as possible. The final ACS major groupings are displayed in Table 3.

Next I have to code the ACS field of study variable into my MAJOR_GROUP variable. I tried to be as consistent with my NLSY97 groupings as possible. The final ACS major groupings are displayed in Table 3.

[11] As a check for robustness, I also conducted the same analysis with all observations in the ACS sample. Changing the age did not significantly alter the results.

Table 3 ACS Major Groupings

MAJOR_GROUP	Number of Observations	ACS Major Code
Arts	3,579	Fine Arts
Biology & Life Sciences	6,577	Agriculture, Biology & Life Sciences, Environment & Natural Resources
Business	12,120	Business
Communication	3,918	Communications, Communication Technologies
Computer Science & Math	2,746	Computer & Information Sciences, Mathematics and Statistics
Education	5,879	Education Administration & Teaching
Engineering	5,191	Architecture, Engineering, Engineering Technologies, Physical Sciences, Nuclear & Industrial Radiology
Health	4,012	Medical & Health Sciences
Humanities & Liberal Arts	6,380	Area & Ethnic Studies, English, History, Liberal Arts, Linguistics & Foreign Languages, Philosophy, Theology
Psychology & Social Work	4,613	Psychology
Social Science	5,634	Law & Public Affairs & Social Work, Social Sciences
Total	**60,649**	

With the ACS data, I construct three major group characteristic variables: occupational concentration, mean hourly wage, and unemployment rate. The occupational concentration variable measures the percentage of all graduates from a given major group that work in the 3 most common occupations for that major group. A detailed table showing the underlying most common occupational fields for each major used to construct this occupational concentration value can be found in the data appendix of this paper. For the occupational groupings, I used occupational fields from the ACS[12].

Occupational concentration serves to describe the clarity of the career path of a major group; a major group that has a very high fraction of its graduates employed in the 3 most common occupational groups can be thought of as having a clear career path. In other words, the question of "What can I do after college with this major?" has a relatively more well-defined answer for majors with high occupational concentrations than

[12]The ACS has both broad and narrow occupational concentration groupings available. I chose to use the broad groupings, but as a check for robustness I also conducted the same analyses using the narrow groupings, which yielded very similar results.

for majors with low occupational concentrations.

Next I calculate the average hourly wage for each major group. I adjust[13] all the wage variables to be in 2010 USD and then remove all individuals who work less than 50 weeks a year. The ACS has only interval data on the number of weeks worked for part-year workers, so I was unable include them when constructing my wage variable. To create a measure of hourly wage for each major group, I divide the annual earnings of every individual in a major group by the number of weekly hours typically worked by that individual times 50. I average all the hourly wages in a particular major group to create my wage variable for that major group.

Finally, I calculate the major group unemployment rate. I simply divide the number of people without jobs seeking work in a particular major group by the total number of people in the labor force from that major group. Chart 4 contains summary statistics from the ACS for each major group.

Table 4 ACS MAJOR_GROUP Statistics

MAJOR_GROUP	Occupational Concentration		Mean Hourly Wage	Unemployment Rate
	Narrow	*Broad*		
Arts	21.17%	51.37%	15.62	8.22%
Biology & Life Sciences	16.40%	46.12%	18.06	4.98%
Business	23.64%	55.25%	20.87	5.40%
Communication	13.36%	48.65%	17.59	5.99%
Computer Science & Math	37.66%	66.28%	24.65	5.28%
Education	67.45%	85.64%	16.51	3.77%
Engineering	24.92%	62.18%	25.34	5.37%
Health	51.99%	78.39%	24.93	3.30%
Humanities & Liberal Arts	17.73%	49.44%	16.78	7.57%
Psychology & Social Work	26.00%	46.28%	16.60	6.25%
Social Science	10.65%	36.92%	20.30	7.82%

Notice a wide range of values exists for occupation concentration, mean hourly wage, and unemployment rate among the major groups. For example, *Arts* has an occupational concentration of 51%, a mean hourly wage of $15.6, and an

[13] I use the "U.S. City Averages" Consumer Price Index from the Bureau of Labor and Statistic from the years 2009, 2010, and 2011 to perform these calculations.

unemployment rate over 8% whereas *Health* has an occupational concentration of 78%, a mean hourly wage of almost $25, and an unemployment rate of 3.3%. Importantly, a major group's rank in one descriptive category is not always similar to its ranking in the other categories. For example, *Education* is near the minimum value of mean hourly wage distribution while *Health* is close to the maximum, but both *Education* and *Health* are among the fields with the highest occupational concentrations.

V. Data Analysis

I begin by using my sample from the NLSY97 to construct a naïve comparison of major choice between first-generation students and non-first-generation students. Table 5 contains a simple breakdown of the percentage of first-generation and non-first-generation students in each major group within my sample.

Table 5 NLSY97 MAJOR_GROUP Breakdown

MAJOR_GROUP	FG_COLLEGE	Not FG_COLLEGE	Difference (FG_COLLEGE - Not FG_COLLEGE)
Arts	2.63%	5.49%	-2.86%
Biology & Life Science	5.26%	7.14%	-1.88%
Business	18.13%	20.74%	-2.61%
Communication	3.8%	6.71%	-2.91%
Computer Science & Math	5.56%	5.19%	0.37%
Education	13.45%	9.03%	4.42%
Engineering	6.43%	6.47%	-0.04%
Health	9.65%	7.87%	1.78%
Humanities & Liberal Arts	5.85%	11.29%	-5.44%
Psychology & Social Work	11.11%	7.02%	4.09%
Social Science	18.13%	13%	5.07%

A statistical significance test rejects the hypothesis that these differences are equal to zero at conventional levels. In other words, this table suggests that there is a relationship between first-generation status and college major selection. However, this sort of analysis fails in addressing my true line of inquiry. As seen earlier in Table 2, first-generation students in my sample set are quite characteristically different than non-first-generation students. Specifically, they are more heavily female and minority than non-first-generation students. In

addition, first-generation students have lower average AFQT scores and come from families with lower average income. We would expect to observe differences in first-generation students' college major selection fueled by these demographic differences alone. In order to answer the question of the causal impact of first-generation status on college major selection, I need to account for these confounding variables.

In order to isolate the effect of first-generation status on major choice for two otherwise identical students, I use my NLSY97 sample to estimate a multinomial logit choice model for college major decisions. The model has MAJOR_GROUP as its dependent variable and FG_COLLEGE, FEMALE, BLACK_HISPANIC, AFQT_PCT, and INCOME_1997 as its independent variables. Because FG_COLLEGE is a dummy variable, its average marginal effect measures the average excess likelihood that a first-generation student selects a given major group compared to an otherwise identical non-first-generation student. We can think of this value as the average effect of being a first-generation college student on college major selection that is independent of differences in sex, race, family income, and ability.

I next examine whether these estimated differences between a first-generation and a non-first-generation students' likelihood of choosing different major groups are systematically related to the characteristics of the major groups themselves. I consider three separate major group characteristics: occupational concentration, average wage, and unemployment rate. Each one is illustrative of a different aspect of the relative safety and stability of a major group's labor market rewards. Unemployment rates represent the risk of not being able to find suitable work, average wages represent expected pay conditional on employment, and occupational concentration serves to describe the clarity of the career path of a major group.

I create 3 scatter plots, each one with a different major group characteristic on the x-axis and the average marginal effect of FG_COLLEGE on the y-axis. Here, I link my college

major choice data from the NLSY97 with my college major returns data from the ACS. Unfortunately, with only 11 major groups to use as data points, rigorous statistical tools are unsuited to measure how precisely a major group's characteristics relate to the excess likelihood that first-generation college students select that major relative to otherwise identical non-first-generation students. However, the scatter plots at least allow for a visual inspection of the relationship between the variables.

VI. Results

I find that first-generation status has a statistically and economically significant effect on college major selection. This effect is independent of the compositional differences of first-generation students, including sex, AFQT score, family income, and race (which also all have a statistically significant effect on college major selection). Table 6 contains the results from a joint significance test across all equations for each variable in my multinomial logit model[14].

Table 6 — Significance test results for the hypothesis that the coefficients on a variable are 0 in all equations of the multinomial logit model

Variable	Chi^2 Statistic	P-value
FG_COLLEGE	24.84	0.0057
AFQT_PCT	58.1	0.0000
INCOME_1997	17.15	0.0711
FEMALE	181.59	0.0000
BLACK_HISPANIC	32.29	0.0004

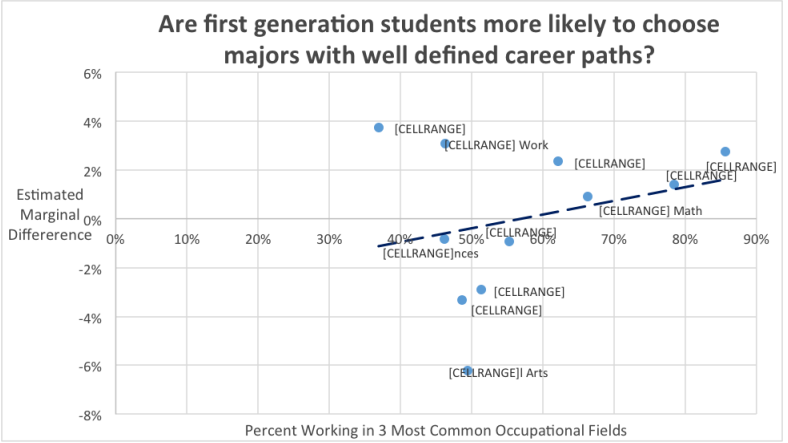
Using the results from my multinomial logit model, I isolate the effect of being a first-generation college student on the probability of selecting a particular of major by calculating the average marginal effect of the FG_COLLEGE variable. These impacts are displayed in table 7.

[14] A full table of results from the multinomial logit model can be found in the data appendix at the end of this paper.

Table 7 Logit Model FG_COLLEGE Difference Results

MAJOR_GROUP	Raw Difference	Adjusted Difference
Arts	-2.86	-2.89%
Biology & Life Sciences	-1.88	-0.85%
Business	-2.61	-0.93%
Communication	-2.91	-3.33%
Computer Science & Math	0.37	0.92%
Education	4.42	2.73%
Engineering	-0.04	2.34%
Health	1.78	1.40%
Humanities & Liberal Arts	-5.44	-6.20%
Psychology & Social Work	4.09	3.08%
Social Science	5.07	3.73%

Next, I construct scatter plots to examine the relationship between the propensity of first-generation students to select a given major group and characteristics of that major group. The vertical axis measures the excess likelihood that a first-generation student selects the major group compared to an otherwise identical student. The horizontal axis measures a particular descriptive statistic of the major group, taken from the ACS. Figures 1 & 2 display scatter plots using the major group's occupational concentration and average wage, respectively. In both cases, simple OLS regressions using the 11 major groups have positive slopes, with the slope in Figure 1 being .26 and the slope in Figure 2 being .40. This suggests that as occupational concentration and expected wages of a major group increase, so does the excess likelihood that first-generation college students select that major relative to otherwise identical non-first generation students.



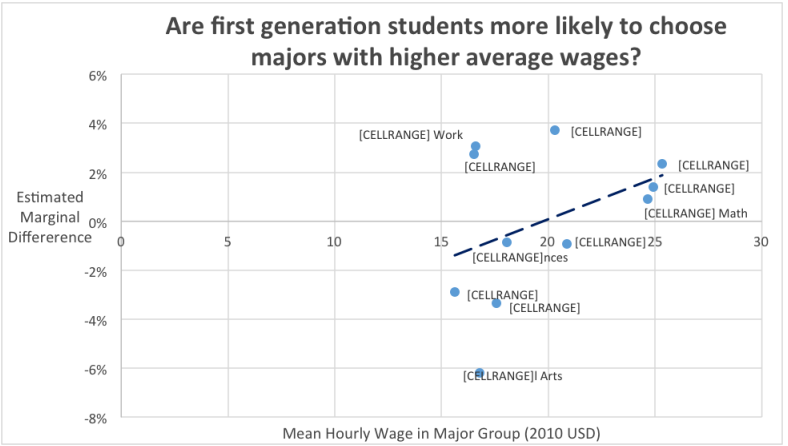
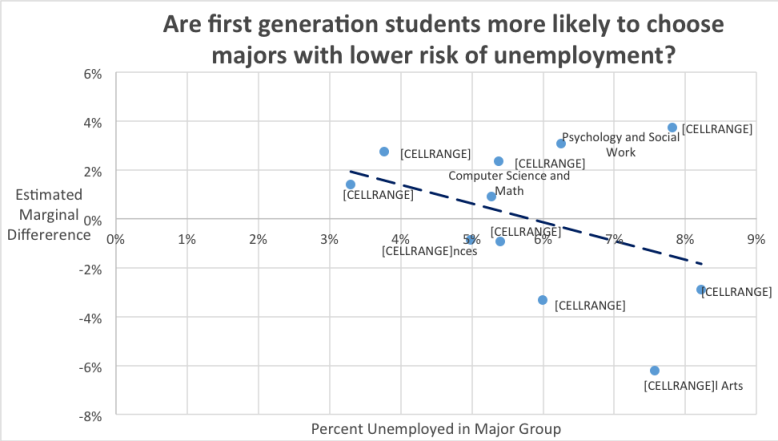


Figure 3 displays a final scatter plot for major group unemployment rates. A simple OLS regression using the 11 major fields has a negative slope with a coefficient of -.39. This suggests that as the unemployment rate for graduates of a major field increases, the likelihood that first-generation college students select that major compared to otherwise identical non-first-generation students decreases.



VII. Discussion

My results highlight how a simple comparison between the college majors selected by first-generation students compared to non-first-generation students can be misleading. For example, a naïve inspection of the raw data would suggest that first-generation students are actually slightly less likely than non-first-generation's students to become engineering majors. However, when I control for the systematic differences in race, ability, family income, and gender between the two groups, it becomes clear that engineering is actually among the specializations which first-generation students prefer most relative to their non-first-generation counterparts.

The multinomial logit model that I estimate not only confirms Saks and Shore's (2005) findings on the effect of family income on a student's college major decision, but also shows that not having parents who have attended college has a significant effect on a student's college major selection. Though an individual's lifetime socioeconomic status is likely partially captured in any variable measuring parental education levels, that first-generation status had an effect even alongside the family income variable suggests that first-generation status may represent an independent effect and mechanism[15].

[15]Variables of FG_COLLEGE interacted with the other 4 explanatory

Perhaps unsurprisingly, in addition to first-generation status, family income, ability, race, and gender all appear to be significantly related to an individual's college major selection.

According to my multinomial logit model, compared to non-first-generation students, first-generation students prefer the following majors: *Computer Science & Math, Education, Engineering, Health, Psychology & Social Work*, and *Social Science*. These same students are less drawn towards the following majors: *Arts, Biology & Life Sciences, Business, Communication*, and *Humanities & Liberal Arts*. The majors groups that first-generation students prefer tended to have low unemployment, high average wages, and a high occupational concentration. The existence of a preference among first-generation college students towards majors groups with these characteristics is consistent with my theoretical framework. Lacking information on intangible benefits to education, these students emphasize labor market rewards when selecting their field of study.

That occupational concentration appears to be related to first-generation student major preferences is particularly interesting. Unlike unemployment rates and average wages, occupational concentration is not directly linked to any economic returns to a college major. In fact, many major groups on both the high and low ends of average wages have similar occupational concentration scores (for example, education and health). Instead, occupational concentration is linked to the clarity of career path post-graduation. That first-generation students, who are likely forced to answer the "Why college?" question more frequently than non-first-generations students, tend to select fields with clear career paths is an important finding.

Unfortunately, data limitations prevented me from attempting to model the joint impact of these major group characteristics. Because the many qualities and expectations

variables failed statistical significance tests, suggesting that this first generation effect may not vary systematically with to income, race, sex, and ability.

of a major simultaneously contribute to its desirability, further research should focus on better understanding the combined effects of these characteristics on the major selection decisions of different groups of students. Additionally, future research should investigate the consequences of having certain disadvantaged and underrepresented groups concentrated in particular major fields. For example, liberal arts programs like Plan II Honors that are interested in having a diverse student body might struggle to seem attractive to low-income and first-generation students. Additionally, universities seeking a diverse faculty may find their supply limited by the practical, more job market oriented focus of the specializations preferred by these underrepresented students.

VIII. Conclusion

The educational decisions an individual makes can have a large impact on many aspects of their life. For example, the wide range of college majors an individual can select from have a correspondingly wide range of economic outcomes. The selection of a college major is a nuanced decision significantly influenced by numerous factors, including ability, sex, race, income, and parental education.

The empirical analysis in this paper suggests that first-generation students, compared to otherwise identical students, are more likely to select major groups with strong labor market rewards and a clear career path. Importantly, these differences exist even after controlling for sex, race, ability, and family income. Given that first-generation students are disproportionately from low-income families, this behavior is likely to contribute to the reduction of economic inequality over the long run. However, that these students are inclined towards economically safer majors with clear career paths suggests that they might be more constrained in their decisions relative to other students. If students have idiosyncratic, major-specific abilities, working to reduce these "constraints" could allow more suitable specialization among certain low-income and first-generation students and improve economic efficiency.

What exactly causes these differences in major selection is unclear. It may be that first-generation students receive less exposure to information on the potential non-monetary rewards to a college education; or perhaps being the first in one's family to attend college places certain pressures on a student; or it could simply be that a characteristically-distinct subset of the population of those whose parents did not attend college decide to pursue higher education. The underlying mechanisms driving the findings of this paper present interesting questions for future research focusing on educational inequality.

References

[1] Altonji, Joseph, Erica Blom, and Costas Meghir. "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers." *Annual Review of Economics 4* (2012): 185-223. Web.

[2] Black, Sandra, Kalena Cortes, and Jane Lincove. "You Have to Apply Yourself: Racial and Ethnic Differences in College Application." Working Paper (2014): Web. 1 Jan. 2015.

[3] Bureau of Labor Statistics, U.S. Department of Labor. National Longitudinal Survey of Youth 1997 cohort, 1997-2011 (rounds 1-15). Produced by the National Opinion Research Center, the University of Chicago and distributed by the Center for Human Resource Research, The Ohio State University. Columbus, OH. (8 February 2015).

[4] Carnevale, Anthony, and Ban Cheah. "From Hard Times to Better Times." *Center on Education and the Workforce.* Georgetown University, 05 Feb. 2015. Web. 06 Apr. 2015.

[5] Cabrera, A. F., & La Nasa, S. M. (2000). Overcoming the tasks on the path to college for America's disadvantaged. In A. Cabrera & S. La Nasa (Eds.), *Understanding the college choice of disadvantaged students*. New Directions for Institutional Research, No. 107, pp. 31-44. San Francisco: Jossey-Bass.

[6] "College Enrollment and Work Activity of 2014 High School Graduates." U.S. Bureau of Labor Statistics. U.S. Bureau of Labor Statistics, 16 Apr. 2015. Web. 26 Aug. 2015.

[7] Hoxby, Caroline, and Christopher Avery. "The Missing "One-Offs": The Hidden Supply of High-Achieving, Low Income Students." National Bureau of Economic Research (2012): Working Paper. Web.

[8] Montmarquette, Claude, Kathy Cannings, and Sophie Mahseredjian. "How Do Young People Choose College Majors?" *Economics of Education Review* 21.6 (2002): 543-56. Web.

[9] Saks, Raven E., and Stephen H. Shore. "Risk And Career Choice." B.E. Journal Of Economic Analysis & Policy: Advances In Economic Analysis & Policy 5.1 (2005): 1-45. Business Source Complete. Web. 20 Nov. 2015.

[10] Stage, F. K., Droogsma-Musoba, G., & Brown, C. (2002, April). *Mathematics achievement: Racial-ethnicity and course taking patterns*. American Educational Research Association, New Orleans.

[11] U.S. Census Bureau, U.S. Department of Commerce. American Community Survey, 2009-2011. <www.census.gov>; (7 September 2015).

Appendix

Top 3 Occupational Fields by Major Group of Recent College Graduates (Wide Groupings)

MAJOR_GROUP	Top Occupational Field Share	Second Occupational Field Share	Third Occupational Field Share	Total Top 3 Share
Arts	Arts, Design, Entertainment, & Media 22.67%	Sales 14.69%	Office and Administrative Support 14.01%	**51.37%**
Biology and Life Sciences	Healthcare Practitioners 18.25%	Education 16.06%	Scientific Research 11.80%	**46.12%**
Business	Financial Specialists 21.52%	Sales 17.86%	Office and Administrative Support 15.87%	**55.25%**
Communication	Office and Administrative Support 17.20%	Arts, Design, Entertainment, & Media 16.27%	Sales 15.18%	**48.65%**
Computer Science and Math	Computation, Programming, and Math 46.79%	Education 13.41%	Sales 6.08%	**66.28%**
Education	Education 75.70%	Sales 5.21%	Office and Administrative Support 4.73%	**85.64%**
Engineering	Architecture and Engineering 41.46%	Computation, Programming, and Math 13.27%	Education 7.45%	**62.18%**
Health	Healthcare Practitioners 69.91%	Office and Administrative Support 4.41%	Sales 4.07%	**78.39%**
Humanities and Liberal Arts	Education 23.13%	Office and Administrative Support 15.25%	Sales 11.07%	**49.44%**
Psychology and Social Work	Education 16.63%	Community and Social Services 16.29%	Office and Administrative Support 13.36%	**46.28%**
Social Science	Office and Administrative Support 14.40%	Sales 11.71%	Community and Social Services 10.81%	**36.92%**

NLSY97 MAJOR_GROUP Multinomial Logit Model Coefficients and Z-statistics

	Arts	*Biology & Life Sciences*	*Business*	*Communication*	*Computer Science & Math*	*Education*	*Engineering*	*Health*	*Humanities & Liberal Arts*	*Psychology & Social Work*	*Social Science*
INCOME_1997	-0.001 (-0.24)	-0.003 (-1.08)	(Omitted)	0.003 (-1.37)	-0.006 (-1.95)	-0.002 (-0.96)	-0.004 (-1.71)	-0.003 (-1.31)	-0.005 (2.45)*	0 (-0.01)	-0.001 (-0.8)
AFQT_PCT	0.009 (2.01)*	0.015 (3.56)**	(Omitted)	0.007 (-1.65)	0.009 (-1.92)	-0.006 (-1.53)	0.024 (5.55)**	0.011 (2.73)**	0.015 (4.01)**	0.004 (-1)	0.005 (-1.36)
FEMALE	0.467 (2.03)*	0.185 (-0.92)	(Omitted)	0.597 (2.77)**	-1.271 (4.97)**	1.282 (6.26)**	-1.231 (5.20)**	(1.22) (5.67)**	0.466 (2.64)**	1.08 (5.02)**	0.174 (-1.09)
FG_COLLEGE	-0.545 (-1.41)	-0.087 (-0.29)	(Omitted)	-0.502 (-1.49)	0.229 (-0.75)	0.334 (-1.44)	0.411 (-1.43)	0.216 (-0.85)	-0.566 (2.00)*	0.452 (-1.83)	0.321 (-1.53)
BLACK_HISPANIC	-0.012 (-0.04)	0.04 (-0.16)	(Omitted)	0.86 (3.62)**	0.397 (-1.48)	-0.274 (-1.24)	0.179 (-0.68)	0.268 (-1.18)	0.148 (-0.69)	0.385 (-1.69)	0.626 (3.37)**
_cons	-2.194 (5.25)** *	-2.094 (5.68)**	(Omitted)	-2.498 (6.59)**	-1.288 (2.99)**	-1.035 (2.89)**	-2.198 (5.83)**	-2.336 (6.11)**	-1.592 (4.77)**	-2.131 (5.44)**	-0.956 (3.16)**

N= 1,981
* $p<0.05$; ** $p<0.01$