

**Education Shapes the Mind, but What Shapes Education?**  
**A Comparative Study of First-Generation vs Non-First-Generation**  
**Students**

May 6, 2019

Brianna Lukowicz

Department of Economics and Finance

Rhode Island College, Providence, RI 02908, USA

**Education Shapes the Mind, but What Shapes Education?**

**A Comparative Study of First-Generation vs Non-First-Generation Students**

By

Brianna Lukowicz

An Honors Project Submitted in Partial Fulfillment of the Requirements

for Honors in the Department of Economics and Finance

Rhode Island College

2019

*For my Mother and Father, Lisa and John,*

*my support and encouragement through it all*

## **ACKNOWLEDGMENTS**

To begin, I would like to thank Dr. Sanae Tashiro, who kept me motivated and helped me to create this project. Without her constant encouragement, tireless recommendations, and endless desire to help me succeed, none of this would have been possible. Her efforts were invaluable, and they mean the world to me. I would also like to thank the Dean of the School of Business, Dr. Mello, and the members of the Economics and Finance Departmental Honors Committee, especially Dr. Aydogdu, Dr. Basu, Dr. Blais, Dr. Karim, and Dr. Saatcioglu for their time spent reading my paper, and sitting through the presentations as the research progressed. I value their feedback immensely; the final paper reflects contributions from everyone named herein.

## **Table of Contents**

<b>I. INTRODUCTION</b>	<b>1</b>
<b>II. TRENDS SHAPING EDUCATION</b>	<b>5</b>
<b>III. THEORY AND TESTABLE HYPOTHESES</b>	<b>14</b>
<b>IV. DATA AND DESCRIPTIVE STATISTICS</b>	<b>16</b>
<b>V. EMPIRICAL METHODOLOGY</b>	<b>22</b>
<b>VI. EMPIRICAL RESULTS</b>	<b>24</b>
<b>VII. CONCLUSIONS AND DISCUSSION</b>	<b>34</b>
<b>VIII. FUTURE RESEARCH</b>	<b>37</b>
<b>IX. TABLES</b>	<b>38</b>
<b>X. REFERENCES</b>	<b>52</b>

## ABSTRACT

This research paper examines in depth the relationship between educational attainment in first-generation college students vs. non-first-generation college students. It analyzes how one's educational attainment level is affected by changes in selected demographic and socioeconomic factors in the United States. This study further analyzes if differences in earnings among the two groups post collegiate education persist. Using the 2015 and 2017 data obtained from the Panel Study of Income Dynamics, the proposed hypotheses are tested with the Linear Probability model and the Binomial Logit model to answer which demographic factors impact educational attainment of each group of students, as well as the standard wage equation, estimated using the Ordinary Least Squares regression to investigate wage differentials.

The empirical results on the analysis of socioeconomic impacts on educational attainment find that first-generation college students are 23-27% less likely to obtain a bachelor's degree than non-first-generation college students. Further, an additional higher level of achievement in parental educational attainment in non-first-generation students nearly doubles the likelihood that an individual in that group will obtain a bachelor's degree, as compared to those first-generation students. The empirical results also find that race, gender, and ethnicity are of the most influential variables in this study. The empirical analysis of earnings differential among first-generation and non-first-generation college students finds first-generation bachelor's degree holders earn 14.1%

less than non-first-generation bachelor's degree holders. In order to improve the first-generation student educational attainment rate, several policies could be considered such as financial compensation, further academic support, and involving families throughout the college process more. In terms of improving the wage differences within the two groups, further advanced degree is beneficial, as well as selecting higher-paying majors, improving job search skills via campus career centers, increasing communication skills and professional networks (public speaking, internships), and increase in alumni connection whilst in college (via alumni office on campus).

Key Words: Education, Returns to Education, First-Generation, Earnings

JEL Classification: I21, I23, I24, I26, J31

## **I. INTRODUCTION**

First-generation college students, in general, may face more challenges stemming from various aspects of life. Some students are lucky enough to have a family that encourages them to continue to further their education and remind them how truly important it is to obtain a college education, and to be the first in their family to accomplish the college-graduate milestone. Only a percentage of students continue to persevere and overcome obstacles that are placed in front of them on the road to achieving a degree. Additionally, there are some students who either choose not to further their education upon high school graduation or begin but do not finish their higher education. This becomes the point in which the question becomes what caused that individual to put their education on pause. This research opportunity allows a better understanding as to what factors influence these individuals to halt their education, as well as explore various other aspects of educational attainment, such as the educational generation concept.

The first objective of this paper is to analyze the cause and effect relationship between educational attainment in individuals and their related socioeconomic and demographic factors. In particular, it examines the effect of education on two distinct groups - first generation students and non-first-generation students at the college level, and further analyze the degree to which the socioeconomic factors impact individuals from each group. This study further extends the analysis by gender, and by race and ethnicity.



The second objective of this study is to examine the earnings of individuals from both the first-generation and non-first-generation college students. It investigates if earnings differential exists after earning a bachelor's degree between first and non-first-generation college degree recipients.

In order to conduct this research effectively, selected academic articles were analyzed to formulate testable hypotheses to be empirically tested with both the linear probability model and the binomial logit model using the 2015 and 2017 samples from the Panel Survey of Income Dynamics (PSID) data. Furthermore, this study follows the theory of human capital to examine the earnings of both first-generation college students, and non-first-generation students after completing their education.

The study finds that first-generation individuals were 23-27% less likely to have obtained a bachelor's degree than those non-first-generation students using the 2015 single-year data, as well as the 2015 and 2017 two-year data. The study also finds that mother and fathers education increases by about 0.04 and 0.02, respectively, among first-generation college students in both the single-year sample and the two-year data. While these same variables were 0.06 and 0.07 among non-first-generation students for mothers and fathers of non-first-generation students, respectively, indicating an additional higher level of achievement for a mother or father's education had a positive impact on the likelihood that the individual obtained a bachelor's degree, and that parent's education has a larger impact on non-first-generation students than first-generation students.

Additionally, females were 6-8% more likely to obtain a bachelor's degree than men, based on the single-year and the two-year data samples. Further, whites were about 8% more likely to have obtained a bachelor's degree than those considered "other" race in the single-year and the two-year data samples. Blacks were about 5% less likely to obtain a bachelor's degree than those considered "other" race in the single-year full-sample data, and 7% less likely based on the two-year full-sample data.

The analysis of earnings differential in this study using the two-year-wage data sample further finds that while there is a wage premium when obtaining a bachelor's degree, annual salaries of those who were first-generation degree recipients, earn 14.1% less than those non-first-generation degree recipients. It is possible that an increase in the marginal product of labor of those who are non-first-generation college degree recipients is less than that of those who were non-first-generation college degree recipients, leading to wage differentials, based on the theory of human capital. Additionally, a lack of job searching skills, professional networking, alumni connections, among first-generation college degree recipients may limit their career advancement in the labor market, as well as occupational segregation and/or occupational choice may lead to wage penalty among first-generation college degree recipients.

In order to improve the educational attainment rate of first-generation students, some of the largest challenges that they face need to be overcome. It is noted that some of the largest challenges

that first-generation students face are navigating the college admissions process, not being academically prepared for college, and managing the financial aspects of college (Engle et al., 2004). By creating programs to bridge the gap between first-generation students and some of these issues, educational attainment rates could increase.

In terms of earnings differential, one of the contributing factors could be stemming from where the first-generation students attended university. A bachelor's degree should be equal no matter where it is obtained from, several employers tend to value degrees from certain universities higher than others. The study by Engle et al. (2004) found that first-generation students pay less on average to attend college than their non-first-generation peers due to the fact that first-generation students are more likely to attend lower-cost public universities.<sup>1</sup> One way to reduce this impact of this issue could be to increase the amount of funding provided as aid for first-generation college students. Furthermore, choice of a major in college among first-generation and non-first-generation college students reduces occupational segregation, which improves earnings differential among these two groups.

---

<sup>1</sup> This includes both two-year community colleges, as well as state four-year colleges and universities.

## **II. TRENDS SHAPING EDUCATION**

### **2.1 Educational Attainment in the United States**

Education has quickly become one of the most important attributes that an individual could have in today's labor market. An individual entering this market without a college degree specialized in one particular area will often have difficulty time obtaining any stable career. For this reason, educational attainment at all levels has increased during the years 2000 through 2017 in individuals, according to the National Center for Education Statistics (NCES), individuals who received a high school diploma, or its equivalent, the General Education Diploma, has increased from 88% to 92% (NCES, 2018). Additionally, the percentage of these individuals who received an associate's degree increased from 38% to 46%, bachelor's degrees increased from 29% to 36%, and a master's degree or higher has increased from 5% to 9% (NCES, 2018). In terms of gender, females aged 25-29 have generally had higher levels of attainment than males in this same age range, however this did not hold true for the high school or the master's degree and higher levels (NCES, 2018). The gender gap, however, doubled to 10% points for associate's degrees or higher, as well as a 5% point increase to 7% point gap for those with a bachelor's degree or higher. Finally, in terms of race, those aged from 25-29 who were White, Black, and Hispanic all had increased educational attainment during this time (2% points, 5% points, and 20% points, respectively), according to NCES (2018).

While at a glance these statistics seem positive, looking more in depth at a level regarding first generation college student and second generation, there are major differences in attainment levels. Interestingly, ten years after they were sophomores in high school, 42% of second-generation college students had actually attained a bachelor's degree compared to only 20% of first-generation students. Furthermore, for those individuals that started college but could not complete it, 54% of first-generation students reported that they couldn't afford to continue going to school versus 45% of continuing education students, as a reason for leaving college without some sort of post-secondary education credential (NCES, 2018). Additionally, these degrees provide these individuals an opportunity to climb the social ladder. It is noted that a larger percentage of first-generation college students come from low income households than non-first-generation college students.<sup>2</sup> These students who are able to complete their degree that come from these low-income households, will more than likely earn a higher salary than their parents, especially for those who are first generation students. This will ultimately provide themselves the opportunity to move up a social class, and have a higher salary than what they grew up in.

---

<sup>2</sup> Households earning \$20,000 or less account for 27% of first-generation students, vs. 6% of non-first-generation students from the same income level, and for households earning \$20,001-\$50,000, (50% vs. 23%) for first-generation and non-first-generation college students, respectively (NCES, 2018)

## **2.2 Related Literature**

Many have turned to research to look for answers. In an effort to better understand the existing research in the field, five journals were helpful in assisting this research study. Due to the fact that this study examines multiple relationships between first-generation and non-first-generation students, several academic sources are utilized touched on several topics. In order to organize the findings, the papers can be categorized into the following areas: (i) first-generation college students vs. non-first-generation college students; (ii) family education; and (iii) race and gender.

### **2.2.1 First Generation College Students vs. Non-First-Generation College Students**

The first paper, Gang and Zimmerman (2000), examine the educational aspect of immigration assimilation, and focus on the educational attainment of the children of guest workers that arrived in Germany from the end of the 1950's until early 1970's. The motivation of the study was to analyze the Becker's allocation-of-time model (1965),<sup>3</sup> using data from the German Socioeconomic Panel, including individuals who were aged 17-38 in 1984.

The key variables used in this study include the individual's education, parental capital/educational attainment of both the mother and father, gender, location born, ethnicity, and age. The endogenous variables such as total years of education, (ordinal) level of schooling, and

---

<sup>3</sup> A household seeks to maximize both its utility of consumption, and educational attainment of children.

(binary) vocational training, resulted in the following models being most appropriate for the study: the ordinary least squares (OLS), ordinal probit, and binomial probit models were used. The major finding in regard to education were that parental education had no independent effect on the educational outcomes of immigrants, while for German natives, father's education had a larger impact than mother's education.

### **2.2.2 Family Education**

The second paper by Lindahl, et al. (2014) focused on the intergenerational transmission of human capital in biological families, measuring the relationship in educational attainment levels between great-grandfathers, grandparents, parents, and children. The objective was to measure and identify the relationship, whether positive or negative, between each biological family member from their respective generation, and the child's education. The theory used to formulate hypotheses that a child's education would be influenced by the generations' education before them was based on utility maximization. That is, parents optimize between their own consumption and investment in children's human capital.

The data in the Lindahl et al. (2014) study was obtained from a survey conducted in Malmö, Sweden. From this data 1,542 third graders from the metropolitan area were interviewed. On average, the great grandfathers were born in the late 19<sup>th</sup> century, while the children in the

sample completed their education in the early 21<sup>st</sup> century.<sup>4</sup> The full sample contained 902 families in which each generation was represented by at least one individual. The data was then further restricted. They used the OLS method to quantify this relationship and found that there was a positive and significant relationship between the education of the grandparents, and the education of the children, conditional on parents' education. While they tested their theory on the full sample, they also created a restricted dataset. They did this in order to avoid measurement error, by only including individuals ages 25 and older in 2009. In order to increase the sample size, they included children who were born in 1990 that completed high school, as the fact that they completed their high school education was a good indication that they could continue to pursue a higher education. The descriptive statistics found that on average, each generation's education increased gradually in both the full sample and the restricted sample set.

The primary model utilized in this study was the OLS model. In terms of major findings, it was found that one year of additional parents' education implied 0.07 more years of schooling in the child and was statistically significant at the one percent level.

The third paper, Fessler and Schneebaum (2012), investigates why gender played such a crucial role in the educational attainment of individuals, and how it has shaped these achievement

---

<sup>4</sup> Great Grandfathers were used because there was no information available for women in education during this time period.



levels throughout history. Furthermore, the authors wanted to test the claims made by other studies, such as the OECD which claimed that parental education was the most important variable in determining a child's education (OECD, 2009), and that parents' education explained 68% of cross-sectional variations in schooling (Hansen and Belzil, 2003).

The Fessler and Schneebaum (2012) study hypothesized that same gender relationships will be stronger than cross gender, and both men and women will have had more mobility over time due to legislative changes (at least in part), as well as the Austrian Feminist Movement.<sup>5</sup> Additionally, the authors hypothesized women's mobility to be higher because many incentives were aimed towards households who presumably sacrifice their daughter's education, in addition to the Austrian feminist movement.

The data used in the study was from the Austrian Household Survey of Housing Wealth conducted by the Austrian Central Bank in 2008, which consisted of 1,892 observations in which respondents were 25 and older. They used two different models. The first model of use was the OLS method, in which they found that fathers and mothers education correlates at a statistically significant level with their descendants' education. They also used the multivariate logit analysis, in which it was found that when a father had high educational attainment, it led to a 0.41 probability

---

<sup>5</sup> This was primarily due to a law passed in Austria 1972 which abolished all costs for university education (at least Austrian citizens and citizens from developing countries), and increased scholarship funding heavily.

for son's educational attainment to be higher, versus 0.26 for daughters, and both were statistically significant at the one percent level. In regard to a highly educated mother, the impact on their son was insignificant, while a daughter's educational attainment probability was increased by 0.25 at the one percent level. Ultimately, the pair drew the conclusion that educational expansion for women was of importance.

The fourth piece of literature that assisted with the research in this study was Ermisch and Pronzato (2011). This study considers why there has been a rise in educational achievements across generations in recent years, especially with women's qualifications having increased more than men's in nearly all of the thirty-four countries that are a part of the Organization for Economic Co-Operation and Development (OECD).

The theory behind the relationship between educational attainment in parents and their children in this study was that parents provide investments into their children that impact their educational attainment, using both time and money as inputs. This study is motivated by the Behrman and Rosenzweig (2002) study, which established that a child's education was linearly dependent on the educational attainment of each parent, plus some unobserved pre-education endowments (e.g. earnings endowment of each parent). This research paper examined several different models to try to explain the causes of variations in educational attainment, but only one of which felt relevant to this paper. The Ermisch and Pronzato (2011) study considers education

of child to be a direct correlation of the education of the mother, the education of the father, the pre-education endowments of the mother, the pre-education endowments of the father, the assumed child rearing-skill of the mother, and a child-specific attribute, respectively.

Ermisch and Pronzato (2011) included both data from Norway and the United States, for comparative reasons to see if results would vary across countries. The data for Norway was obtained from the Norwegian Registry, and the U.S. data was the same data from the earlier study in which they followed, by Behrman and Rosenzeig (2012), which was obtained from the Minnesota Twin/Family Registry. The major findings of Ermisch and Pronzato (2011) include that the correlation between the educational attainments of parents and those of their children overstates the causal effect of parents' education on the education of their children, and that there is some evidence that the mother's effect is larger among less-educated parents, while the father's effect is larger among better educated parents. Another conclusion drawn was that the effect of a mother's education has a larger impact on their daughters than their sons.

### **2.2.3 Race and Gender**

The fifth and final research paper which has helped guide me through this study is by Lucia and Baumann (2009) which uses data administered by NCES from the 1988 National Education Longitudinal Study (NELS). Lucia and Baumann (2009) had a goal of understanding the causes for the differences in college enrollment levels across races. While educational attainment in the

United States and other developing countries has increased within the last five decades, these increases have not been equal across all races. For example, both Whites and Blacks have increased in terms of college enrollment over the last thirty years, while Hispanics have fluctuated from 30% to now around 35%. Using a probit model to estimate the effects of selected socio-economic factors amongst the races that would potentially impact educational attainment, they find several interesting relationships.

Overall, the largest differences that impacted the college enrollment decision in blacks versus whites were factors that included parent's education, school quality, and family income. The first of the several major findings is that blacks were more likely to attend college if their mother attended college. At the same time, whites were not as responsible to parental education after considering the rest of the probit model. Additionally, the authors found that black males benefited from an increase in the quality of the surrounding student body, for example, a better-quality school such as a private school. The final major finding of this study was that black females' college enrollment decision is not affected at all by affordability, while black males only show a slight reaction to tuition prices.

While there is a large amount of literature that examines educational attainment of individuals, none of these analyze the effect of education on two distinct groups; first and non-first-generation students, which this study aims to do.

### III. THEORY AND TESTABLE HYPOTHESES

This section discusses the hypotheses formulated for both parts of this study, including educational attainment differential, as well as wage differential analyses, based upon the existing literature. While often times in economics it's considered that individuals act rationally and in their best interest, that isn't always the case in all situations. In terms of education, one would believe that an individual makes decisions based on the idea that they would act on the want to benefit themselves, and further their education in hopes of bettering their chances of a successful career and future.

An opportunity for attaining a college degree is equally available to all students, regardless of first-generation or non-first-generation, but educational attainment levels are not equal. The educational attainment rates among first-generation students are *lower* than non-first-generation college students (NCES, 2018; Gang and Zimmerman, 2000).

**Hypothesis 1: A first-generation student, is less likely to complete their education, and obtain a bachelor's degree than non-first-generation students.**

Parents, both father and mother, have a significant impact on children's educational attainment (Ermisch and Pronzato, 2011). Additionally, it is argued that parents, both father and mother, have a close interaction with their child. Both the father and mother of an individual have a positive effect on children's education (Lindahl et al, 2014 & Fessler and Schneebaum, 2012).

**Hypothesis 2: The more education of the mother and/or father, the more education of their child.**

Although an opportunity for attaining a college degree is equally presented to all students of all genders, races, and ethnicities, educational attainment is predicted to be different across gender, race, and ethnicity. Female students, white students, and non-Hispanic students do better than their counterparts (Ermisch and Pronzato, 2011 and Lucia and Baumann, 2009).

**Hypothesis 3: A female is more likely to have a bachelor's degree than a male.**

**Hypothesis 4: Whites are more likely to earn a bachelor's degree than other races.**

**Hypothesis 5: Hispanics are less likely to earn a bachelor's degree than non-Hispanics.**

The theory of human capital argues that general training (e.g. college education) and specific training (e.g. on the job training) increases a worker's marginal product of labor and increase his or her wages. In other words, once the initial hurdle of completing a collegiate education for a first-generation student is complete, then their earnings should be indifferent than someone who holds the same education credentials but is considered non-first-generation, as long as MPL is equal among the two groups.

**Hypothesis 6: The wages of a first-generation college student and non-first-generation college students are indifferent post bachelor's degree obtainment.**

Table 1 summarizes the proposed hypotheses based on the existing literature and theory of human capital in detail.

## **IV. DATA AND DESCRIPTIVE STATISTICS**

### **4.1 Educational Attainment Differential Data**

This study uses the Panel Study of Income Dynamics (PSID), the longest running longitudinal household study, conducted by the Institute for Social Research at the Survey Research Center at the University of Michigan. The data set represents a linkage of two datasets extracted from the 2015 and the 2017 PSID survey. For the 2015 initial data sample, this study uses two datafiles; the individual file, which consisted of 24,637 observations, and the household sample which consisted of 9,052 households. The two datasets were then linked by using a common variable, the household identification number. Each household surveyed was assigned a value (1-9,052), and each individual surveyed had a household interview number variable, which was the corresponding value indicating which household they were a part of.

Using the 2015 initial data sample, 16,859 observations were dropped due to invalid responses to some of the key variables. The core variables in this data set included education levels of the individual surveyed, mother's education, and father's education, as well as several other key demographic variables including race, and gender. The observation was further dropped from the data set if the respondent answered "Don't Know" or they refused to provide an accurate answer, or if the data was not available for each parent. Additionally, age was restricted to those who were older than 21, based off the assumption that a bachelor's degree could be obtained in a minimum

of 3 years (18-21 years old). Finally, the data was restricted to those who are biologically related to the family household, as education data for parents was provided primarily for head of household and spouses, while education of children, grandchildren, and parents of the head of household were all computed using the data provided by the PSID, leading to the final dataset, with a total of 7,778 observations.

This same process was repeated for the 2017 sample of the PSID, using the same datafiles, both the individual and the household.<sup>6</sup> The additional year of data was observed in order to use two years for analysis rather than one, to ensure the most accurate results and reduce any bias. The original sample size of the 2017 data set was 26,446 individuals stemming from 9,607 households. After performing the same restrictions and deletion of observations that were discussed, leading to the final dataset with 10,186 observations. Finally, the final data samples, both 2015 and 2017 were merged, leading to the final two-year sample size with 17,964 individuals from 9,651 households. The increase in the number of households from the 2015 sample to the 2017 sample indicates that additional households and individuals were added to the cohort.

The key variables in this study include first-generation college student, where 1 indicates this to be true, and 0 indicates that the individual is non-first-generation, and Mother's and Father's Education, which are ordinal variables. Additionally, other key variables are educational

---

<sup>6</sup> 2016 data was not available for the individual and family files in PSID.



attainment of the respondent, defined as a first-generation college student or not (DPE4). This was generated by using the educational attainment levels of the respondents Mother and Father (EduM and EduD). Several variables utilized in this study were created to binary variables based on the existing data.<sup>7</sup>

## **4.2 Earnings Differential Data**

The data sample used to analyze the earnings difference in those who were first-generation students, and those who were non-first generation, after completing their bachelor's degree was also obtained from the PSID. The same two-year combined dataset was used initially, which included the 17,964 observations from 9,607 households. From there, additional variables were added to the dataset and merged accordingly. In addition to the already existing education, salary, and demographic variables, the following variables were also added; employment status in each year (2015 and 2017), and the occupation code for which the individual was employed. From this point, the sample was further trimmed down by controlling for individuals who held strictly a bachelor's degree, as opposed to including those who did not complete college, or individuals that completed more than a bachelor's degree. Additionally, the sample was then controlled to include only those were currently employed at the time of the survey, and those who reported no salary for those years. After these controls were implemented, the remaining two-year data sample included

---

<sup>7</sup> Including the first-generation variable, all race variables, regions, and salary.

2,317 individuals, as shown in table 3.3, which is defined as the two-year-wage data sample in this study.

### **4.3 Educational Attainment Differential Descriptive Statistics**

Table 2 shows the descriptive statistics of the full sample in the data using both the single-year and the two-year data full samples. Table 2.1 and 2.2 show that 71% of respondents are first-generation college students. Table 3.1 shows the descriptive statistics for first-generation and non-first-generation college students separately, using the single-year full sample data (with 7,778 observations). Similarly, table 3.2 also shows the separated data for first-generation and non-first-generation college students, using the two-year full sample data with 17, 964 observations. It is noted that the average education of the individual surveyed from the full sample including both first, and non-first-generation students is 13.989, or just slightly below 2 years of college education (an associate's degree) when observing the 2015 only sample. Similarly, in the two-year full sample data, the average individual's education is increased slightly, to 14, which is considered an associate's degree. While the average levels of educational attainment of the individual's mother and father are 4.779 and 4.650, respectively in the single-year full sample data (i.e. completed high school (4), but did not complete some non-academic training (5)), the average educational attainment levels of the individual's father and mother decreased slightly, to 4.620 and 4.769, respectively, in the two-year full sample data. Additionally, both first and non-first-generation

students included 58.7% of respondents to be female in the single-year full sample data, and 55.3% in the two-year full sample data. Furthermore, 63.7% of respondents were white, 31.8% were black, and 5.1% were other in the single-year full sample data, whereas 63.1% were white, 31.8% were black, and only 5.4% were “other” race in the two-year full sample data..

Table 3.1 reflects the descriptive statistics by two groups using the single-year full sample data, separated into the first and non-first-generation groups. The group of first-generation students contains 5,505 respondents, who’s average educational attainment level is 13.5 (i.e. some college, but no degree). Further, only 24.2% of respondents in this first-generation single-year sample obtained a bachelor’s degree or higher. The average parental education of a first-generation student in this sample was 3.865 (less than high school diploma) for fathers, and 4.061 for mothers. Additionally, 58.2% of the first-generation students in this single-year first-generation sample were white, 37.1% black, and 5.3% other. Finally, it should be noted that the average age of the first-generation individual in this single-year first-generation sample was 44.7 years old.

Table 3.1 also shows that the non-first-generation group from the single-year full sample data contained a sample size of 2,273, with an average educational attainment level of 15.179, or a little over 3 years of college. In this sample, 60.8% of respondents obtained a bachelor’s degree or more, while the average educational attainment of their parents was 6.550 for fathers, and 6.518 for mothers, both just over what’s considered a bachelor’s degree in this dataset. Additionally, 68.5%

of respondents in this sample had a mother who earned her bachelor's degree, and 71.8% had a father with a bachelor's degree. In regard to race 77.0% of respondents in this sample were white, 19% were black, and 4.8% other. Finally, 56.4% of respondents in this sample were female, while the average age in this sub-sample was 39.6 years old.

In reference to the two-year full sample data, table 3.2 reflects the two groups discussed above. The group of first-generation students contains 12,693 observations, whose average educational attainment was equal to the average of the single-year full sample data (13.5 years). In this group, only 28.5% of respondents held a bachelor's degree, which is lower than the single-year first-generation group. The average education of fathers and mothers in this dataset was slightly lower than the single-year first-generation group, as father's average education level was 3.825 (did not complete high school), where the mother's average education level was 4.043 (high school diploma received). 57.6% of the combined data-set of first-generation students were white, 37.2% were black, and 5.5% were other. The average age of the two-year first-generation group was 45.95.

Table 3.2 also shows that the non-first-generation sample had 5,271 respondents in the two-year full sample data, with an average educational attainment of 15.2 years, or one year less than a bachelor's degree. From this group, 62.7% of respondents hold a bachelor's degree, while 68.6% of the individual's mothers held a bachelor's degree, and 71.8% of fathers. The average parental

educational attainment level for both mothers and fathers of non-first-generation students in the two-year data set were both roughly 6.5, slightly below a bachelor's degree (7). In terms of demographics, 53% of this student sample was female, 76.4% were white, 18.7% were black, and 5.3% were other. Finally, the average age of this student sample was 40 years old.

#### **4.4 Earnings Differential Descriptive Statistics**

Table 3.3 represents the two-year wage data sample used to examine earnings differential between first-generation and non-first-generation students. Given the key variable in this study is the first-generation variable, which accounts for about 49% in the two-year wage data sample. Additionally, approximately 7.2% lived in a metropolitan area at the time of survey. In terms of gender and race, about 54.5% of the 2,317-person sample were female, 78.8% were white, 16.7% were black, and 4.7% were 'Other'. Only 3% of the individuals in the sample were Hispanic, and the average age was around 41 years old. Additionally, the average salary for those in this dataset of individuals who held a bachelor's degree was \$68,271.35 per year.

## **V. EMPIRICAL METHODOLOGY**

### **5.1 Educational Attainment Differential Empirical Methodology**

The first objective of this paper is to examine socioeconomic impacts on education. Because the dependent variable is college degree obtained, which is binary rather than continuous, the empirical method used in this section first considers the Linear Probability model as follows:

$$\Pr(Y=1|X_1, X_2, \dots, X_n) = B_0 + B_1X_1 + B_2X_2 + \dots + B_nX_n \quad \text{--- (1)}$$

Within the LPM there are boundary issues, hence it is more appropriate to use a binomial model.

The independent variables that will be used in this equation include First-Generation, Parent's Education (both mother and father), the gender of the respondent, race of respondent, ethnicity, salary, metropolitan area, and geographical region.

The empirical method thus extends to the binomial logit model, which uses the cumulative distribution standard logistic distribution function that sets boundaries of 0 to 1. The Logit model can be expressed as follows:

$$p(y_{it} = 1) = \frac{e^{\beta'x_{it}}}{1 + e^{\beta'x_{it}}} \quad \text{--- (2)}$$

Post estimation is recommended for the Logit model.<sup>8</sup>

These models will first be utilized on the single-year full sample data, and the two-year full sample data, separately. Then, the respective groups will be separated by first-generation and non-first-generation college students, and the models will be applied again, only to each group individually. The coefficients, in both the LPM and the Logit models, are interpreted by computing the predicted probabilities and differences in predicted probabilities.

---

<sup>8</sup> In order to effectively ensure that no boundary issues are present, the STATA command `dydx(*)` was utilized.

## 5.2 Earnings Differential Empirical Methodology

The second objective of this paper is to analyze wage differential between first and non-first-generation students. The empirical method extends to the Standard wage equation, estimating with the OLS as follows:

$$\ln(w) = B_0 + B_1X_1 + B_nX_n + \varepsilon_i \quad \text{--- (3)}$$

where the dependent variable is the log of annual salary, the independent variables including college degree, age, age squared, gender, race, ethnicity, occupation, metropolitan area, and year,, and  $\varepsilon$  is the standard error.

## VI. EMPIRICAL RESULTS

### 6.1 Full Sample – 2015 Data

Table 4.1 shows the results from both the LPM and the Logit model in the full sample. The adjusted R squared value of 0.2403 in the LPM implies that roughly 24.3% percent of variance in the results can be explained by the model. The coefficient of the first-generation in row (1) of the LPM is -0.273 and is statistically significant at the one percent level. This means that if the individual surveyed was a first-generation student, then the likelihood of having a bachelor's degree is reduced by 27.3%, which is consistent with hypothesis 1. Other notable key coefficients which were positive and statistically significant at the one percent level were female 0.066, White 0.086, metropolitan area 0.102, and all salary variables, as seen in Table 4, except for annual salary below \$50,000. In contrast, notable negative estimates included black individuals, whose

corresponding coefficient was -0.053, and was statistically significant at five percent level, which implies that black individuals were 5.3% less likely to obtain their bachelor's degree than those who were "other" race, which was in line with hypothesis 6 in Table 1. Additionally, as stated previously, individuals earning less than \$50,000 annually were 27% less likely to have a bachelor's degree, which was statistically significant at the one percent level.

In regard to the Logit model in table 4.1, the post estimation was recommended. Therefore, the initial Logit model was run, however from there, a margin command was estimated, then the marginal estimates were calculated to predict the likelihood of  $Y=1$ . When applied to the single-year full sample, the number of observations was 7,377 with a Pseudo R2 value of 0.2019. The coefficients in the Logit model on the full sample overall produced similar results to those of the Limited Probability Model, except for the first-generation variable. For example, the first-generation variable was -0.234 in the logit model, and remained statistically significant at the one percent level. This means that first-generation college students are 23.4% less likely to obtain a bachelor's degree than non-first-generation students, when using the logit model. In terms of similarities, the female coefficient increased by a mere 0.001 and remained significant at the one percent. The only major changes from the LPM model to the Logit model were that the Black coefficient increased by 0.001 and remained statistically significant at the five percent level. Additionally, the variable reflecting salaries between \$200,000 and \$500,000 nearly doubled,



however it became less statistically significant, falling from the one percent level to the five percent level. All of the results mentioned above were in line with the original hypotheses.

## **6.2 Full Sample – Two Year Data Sample**

Table 4.2 shows the empirical results from running both the LPM, and the Logit model, using the two-year full sample data, which includes both first-generation and non-first-generation students. When looking at the results of the linear probability model, first generation students in this data sample were 27.6% percent less likely to obtain a bachelor's degree than those who were non-first generation, which was statistically significant at the one percent level. Females from this sample were 8.4% percent more likely to obtain their bachelor's degree as opposed to men, which was also statistically significant at the one percent level. In terms of race, Whites were 8.5% more likely to obtain a bachelor's degree than those classified as "other", while Blacks were 7.0% less likely, both of which were statistically significant at the one percent level. Further, Hispanics were 4.0% less likely to obtain their bachelor's degree than non-Hispanic individuals, which was significant at the five percent level. This model carried an adjusted R-Squared value of 0.2291, which was slightly lower than the R-squared value for the single-year full sample data mentioned above.

When looking at the logit model as it applies to the two-year full sample data, it is noted that there are no major differences between coefficients, except for the first-generation variable again.

The effects are the same in measurement as compared to the LPM. The first-generation variable with the logit model, states that those individuals that are first-generation were 23.7% less likely to obtain their bachelor's degree, as opposed to non-first-generation students. All key variables were statistically significant at the same levels that they were for the LPM. Finally, the logit model when applied to the two-year full data sample resulted in a Pseudo R2 value of 0.1864. All of the results mentioned above were in line with the hypotheses, as shown in Table 1.

### **6.3 First Generation Sample – 2015 Data**

In terms of the LPM in Table 5.1 which represented the First-Generation group using the single-year full sample data, the adjusted R squared value of 0.1521 was observed. Statistically significant variables from this subsample included the education level of the individual's mother (0.038), education of the individual's father (0.026), female (0.050), White (0.054), Age (0.002), and all salary variables, and were significant at the one percent level. Estimated coefficients were positive on a majority of the variables, except for Black, Hispanic, and salary below \$50,000. In terms of the hypotheses, it was predicted that black and Hispanic would produce that result, whereas surprising results were found with the geographic regions seen in Table 5.1.

Table 5.1 shows that, when the Logit model was applied to the group of those first-generation individuals, a sample size of 5,223 resulted in an R squared value of 0.1268.<sup>9</sup> The results remained similar once again, as all of the signs remained the same as previous. Overall, the coefficients remained relatively similar, as the education of mother variable increased by 0.001 while the education of the father variable decreased by 0.001, the female variable decreased by 0.001, and the variable for white individuals increased by 0.003, all while remaining statistically significant at the one percent level. Additional variables that upheld their statistical significance while remaining nearly the same in value include the metropolitan area variable, and all salary variables (except for the omitted). All of the results mentioned above were in line with hypotheses 1-5, as shown in Table 1. Furthermore, Blacks and Hispanics were not statistically significant in this sample.

#### **6.4 First Generation Sample – Two Year Data Sample**

Table 5.2 reflects the sample of first-generation students from the two-year data set. In terms of the LPM, all key variables were statistically significant at some degree, except for Hispanic. The education of the individual's mother and father produced coefficients of 0.040 and 0.023, respectively, indicating that a one unit increase in their education would produce a 4.0% or

---

<sup>9</sup> It is noted that 9 observations were dropped in this model, due to the fact that the model recognized collinearity between the salary variables and omitted the salary between \$200,000 and \$500,000 variable.

2.3% increase in the probability that the individual obtained a bachelor's degree, as compared to non-first-generation students. Additionally, if the individual was a female, then there was an 6.3% increased chance that the individual had obtained a bachelor's degree, relative to male first-generation students. The three variables discussed above were all statistically significant at the one percent level. In terms of race, White first-generation students were 6.0% more likely to obtain a bachelor's degree than the other race category, which was statistically significant at the one percent level, while Black first-generation students were 4.0% less likely to have obtained their bachelor's degree than other and was statistically significant at the five percent level. All results in Table 5.2 were in line with the hypothesis with the exception of the Hispanic (hypothesis 7 in Table 1), as there was no statistical significance in this model. This model produced an adjusted R squared value of 0.1348, indicating that the model explains roughly 13.5% of the variations in first-generation sample of the two-year full sample data.

In terms of the Logit model when applied to first-generation students from the two-year full sample data, almost all coefficients were the same as above, changing less than one percent. All variables remained statistically significant at their levels mentioned from the LPM referenced above. The Logit model resulted in a Pseudo R2 value of 0.1175. All of the results mentioned above were in line with the hypotheses 1-6 in Table 1, except for Hispanic (hypothesis 7), as it was again statistically insignificant.

## 6.5 Non-First-Generation Sample – 2015 Data

The final application of the LPM was on non-first-generation students, resulting in a sample size of 2,145 when the regression was run, and an adjusted R squared value of 0.2379, as seen in Table 6.1. As previously stated in the last two model specifications, the coefficients were similar in the sense that most of the signs (positive or negative) remained the same; however, there were some notable differences between this non-first-generation group from the single-year full sample data, and the first-generation student group from the single-year full sample data. The effects of an additional higher level of educational achievement in parental educational attainment in those individuals who were non-first-generation college students were greater than that of first-generation college students. This is represented by the coefficients for mother's and father's education levels for non-first-generation students of 0.060 (mother) and 0.067 (father), which were both statistically significant at the one percent level. The coefficients resulted in a near double of the effect of an additional higher level of achievement in parental educational attainment from first-generation students. Further, this indicates that while a mother's education had a larger impact on first-generation students, the opposite was true of those non-first-generation students. Furthermore, the coefficient for female individuals nearly doubled to 0.105, which was still significant at the one percent level. Other notable or statistically significant results were White

(0.099 at five percent), Black (-0.132 at one percent), Age (0.002 at one percent), and Salary below \$50,000 (-0.267 at one percent).

Finally, the Logit model was applied to those under the non-first-generation category, resulting in a sample size of 2,145, and a pseudo R squared 0.2107. Of notable mention are the coefficients for mother's and father's education, as they remained close in terms of the value (i.e., 0.058 for mother's education, and 0.062 for father's education) to those same coefficients in the LPM and remained statistically significant at the one percent level. The female coefficient decreased slightly to 0.100, or a 10% chance that a female is a non-first-generation college student, which also remained statistically significant at the one percent level. In terms of race, the coefficients for all whites and blacks decreased slightly in magnitude to 0.087 for White, and -0.116 for Black, with the White coefficient being statistically significant at the five percent level, and the Black coefficient at the one percent. Further, the coefficient for those of Hispanic ethnicity decreased in magnitude to -0.059; however, it remained statistically insignificant in both models. Finally, a statistically significant variable at the one percent in this group is that for individuals who have a salary up to \$50,000. The coefficient received from the Logit model was -0.257, or in other words, an individual who earned a salary below \$50,000 was 25.7% less likely to obtain a bachelor's degree than a first-generation non-Hispanic college student. All of the results mentioned

above were in line with the hypotheses (1-6) as seen in Table 1, with the exception of the Hispanic variable, and hypothesis 7, as there was no statistical significance.

## **6.6 Non-First-Generation Sample – Two Year Data Sample**

Table 6.2 represents those who were non-first-generation students from the two-year full sample data. When applying the LPM to this sample, all key variables except for Hispanic, which was statistically insignificant, and White, which was significant at the five percent, were statistically significant at the one percent level. The coefficients of mother's education and father's education are 0.060 and 0.071, respectively, indicating that a one unit increase in their education would result in a 6% or 7.1% increase in likelihood that a non-first-generation student obtained a bachelor's degree. It should be noted that these coefficients are 1.5 and 3 times greater the causal effect of parental education on those first-generation individuals. Additionally, being a female non-first-generation college student, she was 12.7% more likely to have a bachelor's degree than a non-first-generation male. In terms of race, Whites and Blacks had coefficients of 0.070, and -0.137, respectively. This model resulted in an adjusted r squared of 0.2180, meaning that the model explains 21.80% of the variations in the dataset.

In terms of the Logit model, the coefficients were again similar, however the causal effect was reduced (i.e. coefficients decreased), which could indicate that the LPM model overstated the causal effect. As mentioned previously, all key variables remained statistically significant, except

for Hispanic. This model resulted in a pseudo R squared of 0.1837. The relationships represented were in line with the original hypotheses.

## **6.7 Earnings Differential Between First-Generation and Non-First-Generation**

Table 7.1 shows the empirical results from using the OLS regression to determine if there was an earnings differential between first-generation and non-first-generation college students who both hold a bachelor's degree. The coefficient of first-generation students is -0.132 and is statistically significant at the one percent level. It implies that first-generation college recipients earn 14.1% lower than non-first-generation bachelor's degree recipients, which is inconsistent with hypothesis 8 in Table 1. In terms of other key variables, it is noted that being a female negative coefficient, which was statistically significant at the one percent level, of -0.315, meaning that if the college degree recipient was a female, they earn 37% less than male respondents with a bachelor's degree, on average.

The unexpected result regarding the wage differentials between first and non-first-generation college bachelor's degree recipients in table 7.1 poses the following potential reasons. First, it is possible that an increase in marginal product of labor due to general skill from college degree training among first generation college degree recipients, is less than that among their counterparts leading to lower annual salary, based on the theory of Human Capital (Becker, 1964). Additionally, skill mismatching and job sorting could be another factor, which means that



individuals are obtaining their bachelor's degree in one area of study, however when they enter the workforce, they enter a different field than what they were previously trained for. Third, it is possible that differences in selection of major, communication skill, professional network, and alumni network. could potentially impact job selection. Often times, individuals who were non-first-generation students have the advantage of an early- established professional network via their parents. First-generation students are disadvantaged in that they must establish a professional network on their own. Finally, the sample is limited given the small size of 2,317 observations, so the sample selection in the data sample may lead to this empirical finding. Additionally, omitted variable bias does exist, such as specific occupation and industry code, as these provide a more accurate measure of an individual's experience in the field that needs to be controlled in order to obtain a more reliable result. Finally, there is a selection bias present, as the sample contains only head of households, as well as spouses, due to the limitations of the salary variable. Additional years of data would need to be added to the sample in order to control for experience. In order to obtain more reliable results, further analysis is warranted.

## **VII. CONCLUSIONS AND DISCUSSION**

This paper examines the differences in what socio-economic factors impact individuals and their educational attainment levels. To do so, I estimate these effects using data from the 2015 and the 2017 Panel Survey of Income Dynamics sample. The key factors in this study include first-

generation or non-first-generation, mother's education, father's education, the education of the individual, and other demographic variables such as race, gender, and ethnicity.

This study finds that while selected socio-economic factors all impact the individual's educational attainment, those with the largest impact are parental education levels, both mothers and fathers, race, and gender. The differences were prominently recognized when comparing two groups separately, as the impact of an additional higher level of achievement in parental educational attainment (i.e., from high school diploma to associates degree, or associate's degree to bachelor's degree) nearly doubles the likelihood that the individual respondent obtained a bachelor's degree in the LPM. For example, the increased likelihood that an individual would obtain a bachelor's degree if they're first-generation based on their mother's and father's educational attainment are 4% and 2.3%, respectively, while non-first-generation students had increased likelihoods of 6%, and 7%, respectively. This study also examined two groups, both first-generation, and non-first-generation students to be further examined to determine which socio-economic factors have the most impact on the likelihood of obtaining a college education.

This study further finds that first-generation college degree recipients earn 14.1% less than those who were non-first-generation degree recipients. The theory of human capital suggests that an increase in marginal product of labor due to general skill and training from college degree

training among first generation college degree recipients, is less than that among their counterparts leading to lower annual salary.

The findings of this study offer the following policy recommendations regarding educational attainment among first-generation college students. By creating groups or programs at the high school and college level to bridge the gap between first-generation students, such as navigating the college admissions process, not being academically prepared for college, and managing the financial aspects of college, is useful. Establishing organization at a high school or even a university, to specifically help to inform first-generation students about the admissions process, as well as helping to prepare them for the college environment would help to improve first-generation educational attainment rates. Additionally, including families in the college process more could help to encourage the support that first-generation students often need.

In terms of earnings differential, it was found that first-generation students receive about the same amount of financial aid as non-first-generation students, regardless of their greater financial need (Pell Institute, 2004). This possibly becomes a problem due to the fact that in most circumstances, first-generation students come from lower-income families, who cannot afford to send their children to more expensive universities. One way to reduce the differences in earnings among first-generation bachelor's degree recipients, and non-first-generation bachelor's degree recipients could be to allow first-generation students to attend more competitive universities by

providing additional financial compensation to first-generation students. Further, selection of major toward higher-paying fields could be one way to reduce occupational segregation. Other ways to reduce the earnings differential would be to enhance communication skills through training, increase professional networks while in college (i.e., internships), increase the activity via an alumni association to provide labor opportunities, and to continue education and obtain an advanced degree.

## **VIII. FUTURE RESEARCH**

This study will be further expanded to include additional years of data, in an effort to expand the data periods for obtaining more reliable results. Furthermore, additional variables will be added to the data and then the models, to see if there are any other major findings within each group. In terms of earnings differential, additional variables will be obtained from the PSID, as well as additional years of data, in an effort to eliminate any biases that exist in the current wage equation. Further, the future of this study will find a way to take single parent household effects into account. By doing so, the objective is to minimize measurement errors that could arise from not being able to previously differentiate between single parent, and two parent households when analyzing educational attainment differences within first and non-first-generation students.

## IX. TABLES

**Table 1**

<b>Hypothesis</b>	<b>Independent Variable</b>	<b><math>\beta</math></b>	<b>Expected Sign</b>	<b>Theory</b>
<b>H1</b>	First Generation Student (0=No, 1=Yes)	B1	$B1 < 0$	If the individual is a first-generation college student, the less likely they are to obtain a bachelor's degree.
<b>H2</b>	Father's College	B2	$B2 > 0$	The more education of the father, the more likely the child is to continue their education.
<b>H3</b>	Mother's College	B3	$B3 > 0$	The more education of the mother, the more likely the child is to continue their education.
<b>H4</b>	Being Female (0=Male, 1=Female)	B4	$B4 > 0$	If the individual respondent is a female, they are more likely obtain a bachelor's degree.
<b>H5</b>	White (0=No, 1=Yes)	B5	$B5 > 0$	If the individual interviewed is White, they will be more likely to have a bachelor's degree.
<b>H6</b>	Black (0=No, 1=Yes)	B6	$B6 < 0$	If the individual interviewed is Black, they will be less likely to have a bachelor's degree.
<b>H7</b>	Hispanic (0=No, 1=Yes)	B7	$B7 < 0$	If the individual interviewed is Hispanic, they will be less likely to have a bachelor's degree.
<b>H8</b>	Current Salary	B8	$B8 = 0$	The current salary of the educated first-generation individual (Bachelor's Degree or more), will be indifferent than the salary of the non-first-generation bachelor's holder.

**TABLE 2.1**  
**Full Sample Descriptive Statistics – 2015**

**Educational Attainment First Generation vs. Non-First Generation**

<b>Variable</b>	<b>n</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Min</b>	<b>Max</b>
First Generation (1=Yes, 0=No)	7,778	0.708	0.455	0	1
Mother has Bachelor's (1=Yes, 0=No)	7,778	0.200	0.400	0	1
Father has Bachelor's (1=Yes, 0=No)	7,778	0.210	0.407	0	1
Years of Ed. – Individual	7,778	13.989	2.127	0	17
Degree Attained (1=Yes, 0=No)	7,778	0.424	0.494	0	1
Education – Father	7,778	4.650	1.783	1	8
Education – Mother	7,778	4.779	1.653	1	8
Number of Educated Parents	7,778	0.410	0.691	0	2
Female (1=Yes, 0=No)	7,778	0.587	0.492	0	1
White	7,778	0.637	0.481	0	1
Black	7,778	0.318	0.466	0	1
Other	7,778	0.051	0.220	0	1
Hispanic	7,778	0.034	0.181	0	1
Metropolitan Area	7,748	0.816	0.388	0	1
Salary	7,272	\$ 22,814.93	\$ 47,863.29	0	\$ 2,000,000.00
Age	7,778	43.243	14.757	21	102
Parents Divorced during Childhood	7,407	0.178	0.383	0	1
Moved during Childhood	7,407	0.097	0.296	0	1
Grew Up In North East	7,778	0.143	0.350	0	1
Grew Up In Mid-West	7,778	0.283	0.450	0	1
Grew Up In south	7,778	0.382	0.486	0	1
Grew Up In West	7,778	0.134	0.341	0	1
Grew Up in Alaska / Hawaii	7,778	0.001	0.038	0	1
Grew Up In Foreign Country	7,778	0.003	0.058	0	1

**TABLE 2.2**  
**Full Sample Descriptive Statistics – 2015 & 2017**

**Educational Attainment First Generation vs. Non-First Generation**

<b>Variable</b>	<b>n</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Min</b>	<b>Max</b>
First Generation (1=Yes, 0=No)	17,964	0.707	0.455	0	1
Mother has Bachelor's (1=Yes, 0=No)	17,964	0.201	0.401	0	1
Father has Bachelor's (1=Yes, 0=No)	17,964	0.211	0.408	0	1
Years of Ed. – Individual	17,964	14.001	2.138	0	17
Degree Attained (1=Yes, 0=No)	17,964	0.352	0.478	0	1
Education – Father	17,964	4.620	1.807	1	8
Education – Mother	17,964	4.769	1.669	1	8
Number of Educated Parents	17,964	0.412	0.692	0	2
Female (1=Yes, 0=No)	17,964	0.553	0.497	0	1
White	17,964	0.631	0.483	0	1
Black	17,964	0.318	0.466	0	1
Other	17,964	0.054	0.226	0	1
Hispanic	17,964	0.033	0.179	0	1
Metropolitan Area	16,490	0.115	0.319	0	1
Salary	17,565	\$ 31,966.90	\$ 54,003.21	0	\$ 2,120,000.00
Age	17,964	44.202	15.495	21	102
Parents Divorced during Childhood	17,700	0.182	0.386	0	1
Moved during Childhood	17,700	0.099	0.299	0	1
Grew Up In North East	17,964	0.148	0.356	0	1
Grew Up In Mid-West	17,964	0.294	0.456	0	1
Grew Up In South	17,964	0.393	0.488	0	1
Grew Up In West	17,964	0.138	0.345	0	1
Grew Up in Alaska / Hawaii	17,964	0.001	0.037	0	1
Grew Up In Foreign Country	17,964	0.003	0.058	0	1

**TABLE 3.1**

**Descriptive Statistics: First Generation vs. Non-First Generation – 2015**

	<b>First Generation</b>					<b>Non-First Generation</b>				
<b>Variable</b>	<b>Obs</b>	<b>Mean</b>	<b>Std. Dev</b>	<b>Min</b>	<b>Max</b>	<b>Obs</b>	<b>Mean</b>	<b>Std. Dev</b>	<b>Min</b>	<b>Max</b>
<b>First Generation (1=Yes, 0=No)</b>	5,505	1	0	1	1	2,273	0.000	0.000	0	0
<b>Years of Ed. – Individual</b>	5,505	13.498	2.045	0	17	2,273	15.179	1.831	0	17
<b>Degree Obtained (1=Yes, 0=No)</b>	5,505	0.242	0.428	0	1	2,273	0.608	0.488	0	1
<b>Education – Father</b>	5,505	3.865	1.216	1	6	2,273	6.550	1.481	1	8
<b>Education – Mother</b>	5,505	4.061	1.164	1	6	2,273	6.518	1.339	2	8
<b>Mother has Bachelor’s (1=Yes, 0=No)</b>	5,505	0	0	0	0	2,273	0.685	0.465	0	1
<b>Father has Bachelor’s (1=Yes, 0=No)</b>	5,505	0	0	0	0	2,273	0.718	0.450	0	1
<b>Number of Educated Parents</b>	5,505	0	0	0	0	2,273	1.403	0.491	1	2
<b>Female (1=Yes, 0=No)</b>	5,505	0.597	0.491	0	1	2,273	0.564	0.496	0	1
<b>White</b>	5,505	0.582	0.493	0	1	2,273	0.770	0.421	0	1
<b>Black</b>	5,505	0.371	0.483	0	1	2,273	0.190	0.393	0	1
<b>Other</b>	5,505	0.053	0.223	0	1	2,273	0.048	0.213	0	1
<b>Hispanic</b>	5,505	0.037	0.188	0	1	2,273	0.028	0.165	0	1
<b>Age</b>	5,505	44.743	15.235	21	102	2,273	39.610	12.824	21	94
<b>Salary</b>	5,147	\$ 17,836.54	\$ 43,980.47	\$ -	\$ 2,000,000.00	2,125	\$ 34,873.19	\$ 54,317.95	\$ -	\$ 815,000.00
<b>Grew Up In North East</b>	5,505	0.132	0.339	0	1	2,273	0.170	0.376	0	1
<b>Grew Up In Mid-West</b>	5,505	0.279	0.449	0	1	2,273	0.290	0.454	0	1
<b>Grew Up In South</b>	5,505	0.412	0.492	0	1	2,273	0.309	0.462	0	1
<b>Grew Up In West</b>	5,505	0.121	0.326	0	1	2,273	0.167	0.373	0	1
<b>Grew Up In Alaska / Hawaii</b>	5,505	0.002	0.040	0	1	2,273	0.001	0.030	0	1
<b>Grew Up in Foreign Country</b>	5,505	0.001	0.033	0	1	2,273	0.009	0.093	0	1
<b>Current Household Region</b>	5,505	2.649	0.882	1	6	2,273	2.701	1.009	1	6
<b>Metropolitan Area</b>	5,497	0.787	0.409	0	1	2,251	0.886	0.318	0	1
<b>Parents Divorced during Childhood</b>	5,240	0.162	0.368	0	1	2,167	0.219	0.413	0	1



**TABLE 3.2**

**Descriptive Statistics: First Generation vs. Non-First Generation – 2015 & 2017**

Variable	First Generation					Non-First Generation				
	Obs	Mean	Std. Dev	Min	Max	Obs	Mean	Std. Dev	Min	Max
<b>First Generation (1=Yes, 0=No)</b>	12,693	1	0	1	1	5,271	0.000	0.000	0	0
<b>Years of Ed. – Individual</b>	12,693	13.503	2.063	0	17	5,271	15.200	1.816	0	17
<b>Degree Obtained (1=Yes, 0=No)</b>	12,693	0.245	0.430	0	1	5,271	0.610	0.488	0	1
<b>Education – Father</b>	12,693	3.825	1.234	1	6	5,271	6.535	1.509	1	8
<b>Education – Mother</b>	12,693	4.043	1.176	1	6	5,271	6.517	1.356	1	8
<b>Mother has Bachelor’s (1=Yes, 0=No)</b>	12,693	0	0	0	0	5,271	0.686	0.464	0	1
<b>Father has Bachelor’s (1=Yes, 0=No)</b>	12,693	0	0	0	0	5,271	0.718	0.450	0	1
<b>Number of Educated Parents</b>	12,693	0	0	0	0	5,271	1.403	0.491	1	2
<b>Female (1=Yes, 0=No)</b>	12,693	0.563	0.496	0	1	5,271	0.530	0.499	0	1
<b>White</b>	12,693	0.576	0.494	0	1	5,271	0.764	0.425	0	1
<b>Black</b>	12,693	0.372	0.483	0	1	5,271	0.187	0.390	0	1
<b>Other</b>	12,693	0.055	0.228	0	1	5,271	0.053	0.223	0	1
<b>Hispanic</b>	12,693	0.035	0.184	0	1	5,271	0.029	0.167	0	1
<b>Age</b>	12,693	45.948	16.001	21	102	5,271	39.999	13.294	21	94
<b>Salary</b>	12,425	\$ 25,969.67	\$ 44,104.22	\$ -	\$ 2,000,000.00	5,140	\$ 46,464.08	\$ 70,481.36	\$ -	\$ 2,120,000.00
<b>Grew Up In North East</b>	12,693	0.136	0.343	0	1	5,271	0.179	0.383	0	1
<b>Grew Up In Mid West</b>	12,693	0.291	0.454	0	1	5,271	0.302	0.459	0	1
<b>Grew Up In South</b>	12,693	0.427	0.495	0	1	5,271	0.312	0.463	0	1
<b>Grew Up In West</b>	12,693	0.124	0.329	0	1	5,271	0.173	0.379	0	1
<b>Grew Up In Alaska / Hawaii</b>	12,693	0.001	0.038	0	1	5,271	0.001	0.036	0	1
<b>Grew Up in Foreign Country</b>	12,693	0.001	0.035	0	1	5,271	0.008	0.091	0	1
<b>Current Household Region</b>	12,693	2.651	0.877	1	6	5,271	2.697	1.011	1	6
<b>Metropolitan Area</b>	11,520	0.135	0.341	0	1	4,970	0.070	0.255	0	1
<b>Parents Divorced during Childhood</b>	12,518	0.166	0.372	0	1	5,182	0.219	0.413	0	1

**TABLE 3.3**  
**Descriptive Statistics: Earnings Analysis**  
**Bachelor's Degree Holders**

<b>Variable</b>	<b>n</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Min</b>	<b>Max</b>
Year	2,317	2016.233	0.973	2015	2017
Bachelor's Degree Received	2,317	1.000	0.000	1	1
First-Generation Student	2,317	0.489	0.500	0	1
Metropolitan Area	2,317	0.072	0.258	0	1
Female	2,317	0.546	0.498	0	1
White	2,317	0.788	0.409	0	1
Black	2,317	0.167	0.373	0	1
Other	2,317	0.047	0.218	0	1
Hispanic	2,317	0.029	0.168	0	1
Salary	2,317	\$ 68,271.35	\$ 79,658.25	\$ 1	\$ 2,120,000.00
Age	2,317	40.659	11.833	21	84

**TABLE 4.1**  
**Empirical Results: Full Sample 2015**

Variable	Linear Probability Model		Logit Model	
	Coef.	Std. Error	Marginal Effect	Std. Error
First Generation (1=Yes, 0=No)	-0.273***	0.011	-0.234***	0.009
Female (1=Yes, 0=No)	0.066***	0.010	0.067***	0.010
White (1=Yes, 0=No)	0.086***	0.024	0.083***	0.024
Black (1=Yes, 0=No)	-0.053**	0.025	-0.054**	0.025
Hispanic (1=Yes, 0=No)	-0.033	0.027	-0.032	0.028
Age	0.000	0.000	0.000	0.000
Grew Up In the North East	-0.036	0.050	-0.036	0.048
Grew Up In the Mid-West	-0.048	0.050	-0.048	0.048
Grew Up In the South	-0.043	0.050	-0.044	0.048
Grew Up In the West	-0.095*	0.050	-0.093*	0.049
Metropolitan Area	0.102***	0.013	0.102***	0.013
Parent's Divorced	0.024*	0.013	0.022*	0.012
Salary up to \$50K	-0.270***	0.016	-0.220***	0.014
Salary b/w \$50K & \$100K	0.140***	0.028	0.137***	0.028
Salary b/w \$100K & \$200K	0.213***	0.029	0.237***	0.035
Salary b/w \$200K & \$500K	0.257***	0.079	0.448**	0.178
Constant	0.625***	0.058	0.475	0.327
	n= 7,377 Adj. R-Sq = 0.2447		n= 7,377 Pseudo R2= 0.2007	

**TABLE 4.2**

**Empirical Results: Full Sample 2015 & 2017**

Linear Probability Model			Logit Model	
Variable	Coef.	Std. Error	Marginal Effect	Std. Error
First Generation (1=Yes, 0=No)	-0.276***	0.008	-0.237***	0.006
Female (1=Yes, 0=No)	0.084***	0.007	0.084***	0.007
White (1=Yes, 0=No)	0.085***	0.015	0.078***	0.015
Black (1=Yes, 0=No)	-0.070***	0.016	-0.076***	0.016
Hispanic (1=Yes, 0=No)	-0.040**	0.019	-0.042**	0.019
Age	0.001***	0.000	0.001***	0.000
Grew Up In the North East	-0.040	0.033	-0.039	0.032
Grew Up In the Mid-West	-0.050	0.033	-0.048	0.032
Grew Up In the South	-0.045	0.033	-0.043	0.032
Grew Up In the West	-0.103***	0.033	-0.098***	0.032
Metropolitan Area	-0.117***	0.011	-0.119***	0.011
Parent's Divorced	0.010	0.009	0.009	0.009
Salary up to \$50K	-0.153***	0.037	-0.138***	0.033
Salary b/w \$50K & \$100K	0.091**	0.037	0.072**	0.034
Salary b/w \$100K & \$200K	0.246***	0.039	0.222***	0.037
Salary b/w \$200K & \$500K	0.385***	0.053	0.500***	0.082
Constant	0.597***	0.051	0.472*	0.273
	n= 16,272 Adj R2= 0.2291		n= 16,272 Pseudo R2= 0.1864	

**TABLE 5.1: First-Generation  
First-Generation Students 2015**

Variable	Linear Probability Model		Logit Model	
	Coef.	Std. Error	Marginal Effect	Std. Error
Education of Mother	0.038***	0.005	0.038***	0.005
Education of Father	0.026***	0.005	0.026***	0.005
Female (1=Yes, 0=No)	0.050***	0.011	0.049***	0.012
White (1=Yes, 0=No)	0.054**	0.027	0.057**	0.029
Black (1=Yes, 0=No)	-0.022	0.028	-0.024	0.030
Hispanic (1=Yes, 0=No)	-0.018	0.031	-0.020	0.032
Age	0.002***	0.000	0.002***	0.000
Grew Up In the North East	-0.042	0.064	-0.042	0.060
Grew Up In the Mid-West	-0.049	0.063	-0.045	0.059
Grew Up In the South	-0.057	0.063	-0.055	0.059
Grew Up In the West	-0.104	0.064	-0.099	0.060
Metropolitan Area	0.086***	0.014	0.092***	0.015
Parent's Divorced	0.008	0.015	0.007	0.015
Salary up to \$50K	-0.238***	0.019	-0.180***	0.016
Salary b/w \$50K & \$100K	0.173***	0.037	0.119***	0.031
Salary b/w \$100K & \$200K	0.339***	0.041	0.262***	0.041
Salary b/w \$200K & \$500K	0.493***	0.134	-	-
Constant	0.044	0.078	-2.856***	0.476
n=5,232		n=5,223		
Adj. R-Sq.= .1500		Pseudo R2.= .1255		

**TABLE 5.2: First-Generation**  
**First-Generation Students 2015 & 2017**

Variable	Linear Probability Model		Logit Model	
	Coef.	Std. Error	Marginal Effect	Std. Error
Education of Mother	0.040***	0.004	0.041***	0.004
Education of Father	0.023***	0.004	0.023***	0.004
Female (1=Yes, 0=No)	0.063***	0.008	0.062***	0.008
White (1=Yes, 0=No)	0.060***	0.018	0.059***	0.019
Black (1=Yes, 0=No)	-0.040**	0.018	-0.042**	0.019
Hispanic (1=Yes, 0=No)	-0.029	0.021	-0.036	0.023
Age	0.002***	0.000	0.002***	0.000
Grew Up In the North East	-0.021	0.041	-0.023	0.039
Grew Up In the Mid-West	-0.034	0.040	-0.030	0.039
Grew Up In the South	-0.037	0.040	-0.035	0.039
Grew Up In the West	-0.104**	0.041	-0.099**	0.040
Metropolitan Area	-0.101***	0.011	-0.109***	0.013
Parent's Divorced	-0.005	0.010	-0.006	0.010
Salary up to \$50K	-0.088**	0.045	-0.083**	0.040
Salary b/w \$50K & \$100K	0.128***	0.046	0.093**	0.041
Salary b/w \$100K & \$200K	0.378***	0.049	0.271***	0.044
Salary b/w \$200K & \$500K	0.542***	0.076	0.467***	0.095
Constant	-0.050	0.067	-3.039***	0.387

  

n=11,383	n=11,383
Adj. R-Sq.= .1348	Pseudo R2= .1175

**TABLE 6.1: Non- First-Generation  
Non-First-Generation Students 2015**

Variable	Linear Probability Model		Logit Model	
	Coef.	Std. Error	Marginal Effect	Std. Error
Education of Mother	0.060***	0.007	0.058***	0.007
Education of Father	0.067***	0.007	0.063***	0.006
Female (1=Yes, 0=No)	0.105***	0.019	0.101***	0.018
White (1=Yes, 0=No)	0.099**	0.046	0.086**	0.042
Black (1=Yes, 0=No)	-0.132***	0.049	-0.120***	0.046
Hispanic (1=Yes, 0=No)	-0.061	0.055	-0.059	0.053
Age	0.002***	0.001	0.002***	0.001
Grew Up In the North East	0.013	0.079	0.016	0.078
Grew Up In the Mid-West	-0.006	0.077	-0.006	0.076
Grew Up In the South	0.045	0.077	0.043	0.076
Grew Up In the West	-0.044	0.079	-0.047	0.077
Metropolitan Area	0.062**	0.029	0.055**	0.027
Parent's Divorced	0.032	0.023	0.036*	0.022
Salary up to \$50K	-0.267***	0.026	-0.257***	0.027
Salary b/w \$50K & \$100K	0.085**	0.042	0.182***	0.063
Salary b/w \$100K & \$200K	0.065	0.043	0.129**	0.059
Salary b/w \$200K & \$500K	0.076	0.098	0.206	0.187
Constant	-0.311**	0.122	-4.239***	0.675

n= 2,145 Adj. R-Sq.= .2376
-------------------------------

n= 2,145 Pseudo R2= .2097
------------------------------

**TABLE 6.2: Non- First-Generation  
Non-First-Generation Students 2015 & 2017**

Variable	Linear Probability Model		Logit Model	
	Coef.	Std. Error	Marginal Effect	Std. Error
Education of Mother	0.060***	0.005	0.058***	0.004
Education of Father	0.071***	0.004	0.066***	0.004
Female (1=Yes, 0=No)	0.127***	0.012	0.122***	0.012
White (1=Yes, 0=No)	0.070**	0.029	0.063**	0.027
Black (1=Yes, 0=No)	-0.137***	0.031	-0.120***	0.029
Hispanic (1=Yes, 0=No)	-0.026	0.037	-0.025	0.036
Age	0.003***	0.000	0.002***	0.000
Grew Up In the North East	-0.010	0.054	-0.011	0.055
Grew Up In the Mid-West	-0.020	0.053	-0.022	0.054
Grew Up In the South	0.026	0.054	0.025	0.054
Grew Up In the West	-0.049	0.054	-0.052	0.054
Metropolitan Area	-0.082***	0.024	-0.075***	0.023
Parent's Divorced	0.019	0.015	0.022	0.015
Salary up to \$50K	-0.216***	0.061	-0.195***	0.067
Salary b/w \$50K & \$100K	0.022	0.061	0.033	0.063
Salary b/w \$100K & \$200K	0.077	0.063	0.109*	0.063
Salary b/w \$200K & \$500K	0.183**	0.078	0.440**	0.146
Constant	-0.308***	0.099	-4.23***	0.548

<p>n= 4,889 Adj. R-Sq.= .2180</p>
---------------------------------------

<p>n= 4,889 Pseudo R2= .1837</p>
--------------------------------------



**TABLE 7.1: Wage Premium**  
**Full Sample – 2015 & 2017**

Variable	Coef.	Std. Error
Degree	0.400***	0.027
Age	0.108***	0.006
Age Squared	-0.001***	0.000
Female	-0.257***	0.027
White	0.121**	0.050
Black	-0.063	0.052
Hispanic	0.001	0.062
Metropolitan Area	-0.196***	0.034
Management	.599***	0.095
Business/Financial Services	0.604***	0.100
Computers & Math	0.729***	0.113
Architectural/Engineering	0.645***	0.124
Science Occupations	0.445***	0.148
Communication Social	0.036	0.122
Legal	0.712***	0.168
Education/ Training/ Library	-0.224**	0.104
Arts, Design, Entertainment, and Sports Media	0.003	0.117
Health Practitioners & Technicians	0.500***	0.102
Healthcare Support	0.137	0.106
Protective Services	0.307***	0.107
Food Prep & Serving	-0.203**	0.102
Building & Grounds Cleaning & Maintenance	-0.502***	0.110
Personal Care & Services	-0.421***	0.104
Sales	0.158*	0.095
Office Administrators	0.222**	0.093
Farming, Fishing, & Agriculture	0.106	0.173
Construction / Extraction	0.239**	0.103
Installations, Maintenance, and Repairs	0.200*	0.105
Production	0.247**	0.099
Transportation Materials	0.103	0.098
Year (1= 2015)	0.326***	0.027
Constant	-0.013	0.013

n= 7,070

R Squared= 0.2306

Adj. R Squared= 0.2271

**TABLE 8.1: Earnings Analysis**  
**Bachelor's Degree Holders 2015 & 2017**

Variable	Coef.	Std. Error
First Generation Student	-0.132***	0.038
Age	0.114***	0.011
Age Squared	-0.001***	0.000
Female	-0.315***	0.041
White	-0.010	0.087
Black	-0.081	0.096
Hispanic	-0.070	0.108
Metropolitan Area	-0.225***	0.072
Management	0.600***	0.182
Business/Financial Services	0.523***	0.184
Computers & Math	0.603***	0.194
Architectural/Engineering	0.547***	0.199
Science Occupations	0.280	0.220
Communication Social	-0.110	0.202
Legal	0.839***	0.261
Education/ Training/ Library	-0.240	0.188
Arts, Design, Entertainment, and Sports Media	0.112	0.199
Health Practitioners & Technicians	0.381**	0.190
Healthcare Support	-0.307	0.240
Protective Services	0.175	0.212
Food Prep & Serving	-0.250	0.231
Building & Grounds Cleaning & Maintenance	-0.303	0.265
Personal Care & Services	-0.527**	0.210
Sales	0.179	0.184
Office Administrators	0.015	0.185
Farming, Fishing, & Agriculture	-0.305	0.619
Construction / Extraction	-0.002	0.230
Installations, Maintenance, and Repairs	0.016	0.234
Production	0.108	0.220
Transportation Materials	0.345	0.243
Year (1= 2015)	0.150***	0.039
Constant	-0.036*	0.019

<p>n= 2,177</p> <p>R Squared= 0.2447</p> <p>Adj. R Squared= 0.2335</p>
--

## X. REFERENCES

Engle, Jennifer, Adolfo Bermeo, and Colleen O'Brien. "Straight From the Source - What Works for First-Generation College Students." The Pell Institute for the Study of Opportunity in Higher Education □ Publications. December 2006. Accessed May 5, 2019.  
[http://www.pellinstitute.org/publications-Straight\\_from\\_the\\_Source.shtml](http://www.pellinstitute.org/publications-Straight_from_the_Source.shtml).

Ermisch, John, and Chiara Pronzato. "Chapter 8, Causal Effects of Parents' Education on Children's Education." In *Persistence, Privilege, and Parenting: The Comparative Study of Intergenerational Mobility*, by Timothy M. Smeeding, 237-60. New York: Russell Sage Foundation, 2011.

Fessler, Pirmin, and Alyssa Schneebaum. 2012. "Gender and Educational Attainment across Generations in Austria." *Feminist Economics* 18 (1): 161–88. doi:<http://www-tandfonline-com.ric.idm.oclc.org/loi/rfec20>.

Gang, Ira N., and Klaus F. Zimmermann. "Is Child like Parent? Educational Attainment and Ethnic Origin." *The Journal of Human Resources* 35, no. 3 (2000): 550-69. doi:10.2307/146392.

"Human Capital Theory." Simple Economist. [Accessed Online: May 5, 2019].  
<http://simpleeconomist.com/human-capital-theory>]

Lindahl, Mikael, Mårten Palme, Sofia Sandgren-Massih, and Anna Sjögren. "A Test of the Becker-Tomes Model of Human Capital Transmission Using Microdata on Four Generations." *Journal of Human Capital* 8, no. 1 (2014): 80-96. doi:10.1086/674104.

Lucia, Kathlyn E., and Robert W. Baumann. "DIFFERENCES IN THE COLLEGE ENROLLMENT DECISION ACROSS RACE." *The American Economist* 53, no. 1 (2009): 60-74. <http://www.jstor.org/stable> [Accessed Online May 5, 2019]

"The NCES Fast Facts Tool Provides Quick Answers to Many Education Questions (National Center for Education Statistics)." National Center for Education Statistics (NCES) Home Page, a Part of the U.S. Department of Education. 2018 <https://nces.ed.gov/fastfacts/display.asp?id=27>. [Accessed online May 5, 2019]

Redford, Jeremy, and Kathleen Mulvaney Hoyer. *National Center for Education Statistics. Stats in Brief*

Stock, James H., and Mark W. Watson. *Introduction to Econometrics*. Boston: Pearson, 2014.

