Does Family Income Determine A Children Future Educational Attainment Level?

Diaisha T. Richards
State University of New York College at Buffalo - Buffalo State College, richard01@mail.buffalostate.edu

Advisor
Dr. Frederick Floss

First Reader
Dr. Frederick Floss

Second Reader
Dr. Theodore Byrley

Department Chair
Dr. Frederick Floss

To learn more about the Economics and Finance Department and its educational programs, research, and resources, go to https://economics.buffalostate.edu.

Recommended Citation
https://digitalcommons.buffalostate.edu/economics_theses/38

Follow this and additional works at: https://digitalcommons.buffalostate.edu/economics_theses

Part of the African American Studies Commons, Econometrics Commons, Economic Theory Commons, Ethnic Studies Commons, Finance Commons, Higher Education Commons, Income Distribution Commons, and the Latina/o Studies Commons
Does Family Income Determine A Children Future Educational Attainment Level?

Diaisha T. Richards

An Abstract of a Thesis
In
Applied Economics

Submitted in Partial Fulfillment
Of the Requirements
Of the Degree of

Master of Arts

May 2019

Buffalo State College
State University of New York
Department of Economics and Finance
Abstract

Family income and education have been a major concern in a variety of researches, and as a topic in society. These two components are a major concern because they are known to be key elements in determining future success for an individual. Various studies investigated the significance, correlations and impacts these two factors have on one another. It is common for the amount of family income obtained to determine how much education one will receive in the future. This study focuses on testing the hypothesis that family income determines how much education a child will receive in the future. By exploring the possible relationships between both education and family income, and other factors such as gender and race, this study analyzes and determines the common assumptions. More specifically, the idea that family income significantly influences the amount, or level, of education a child will receive in the future. This analysis is carried out using a linear regression on family income, race, and gender versus the educational attainment received. Findings show that family income continues to play a significant role in a child's future educational attainment level. Findings also show gender and race playing a significant role in a child's future educational attainment if you are a female, and a significant role if your race is white.
Does Family Income Determine A Children Future Educational Level?

A Thesis in
Applied Economics

by

Diaisha T. Richards

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Arts
May 2019

Dates of Approval:

__________________________________________
Frederick Floss, Ph.D.
Professor and Chair of Economics & Finance
Thesis Advisor

__________________________________________
Kevin Miller, Ed.D.
Dean of the Graduate School
Dates of Approval

______________________________
Frederick Floss, Ph.D.
Professor and Chair of the Committee

______________________________
Theodore Byrley, Ph.D.
Associate Professor
I would like to acknowledge Dr. F. Floss and Dr. Byrley for all their help.

I would also like to acknowledge God and my support system: especially my mom Jacqueline, Uncle Kenton and Uncle Mahlon.
**Table of Contents**

Chapter 1: Introduction ........................................................................................................... 2

Chapter 2: Review of Literature ............................................................................................ 3
  2.1 Educational Attainment ................................................................................................. 3
    2.1.1 Intergenerational mobility .................................................................................... 4
  2.2 Family Income ............................................................................................................... 5
  2.3 Gender ......................................................................................................................... 10
  2.4 Race ............................................................................................................................ 12

Chapter 3: Methodology/Theoretical Model ......................................................................... 16
  3.1 Data ............................................................................................................................ 16
  3.2 Hypothesis .................................................................................................................. 17
  3.3 Econometric Regression Analysis ............................................................................... 18
    3.3.1 Regression Analysis .............................................................................................. 18
    3.3.2 Regression Analysis (2000) ................................................................................ 22
    3.3.2 Regression Analysis (2005) ................................................................................ 25
    3.3.3 Regression Analysis (2010) ................................................................................ 29
    3.3.4 Regression Analysis (2016) ................................................................................ 32
    3.3.5 Tables .................................................................................................................... 36
  3.5 Results .......................................................................................................................... 39

Chapter 4: Implications and Predictions ............................................................................. 40

Chapter 5: Conclusions ......................................................................................................... 42

References ............................................................................................................................. 44
Chapter 1: Introduction

Family income and education have been a major concern in a variety of researches, and as a topic in society. These two components are a major concern because they are both known to be key elements in determining the future success for an individual. Various studies have investigated the correlations and impacts of the two factors: family income and education. This study focuses on the idea that the amount of family income can determine the level of educational attainment a child will receive in the future. Throughout the investigation, education is used as a function of family income, race and gender to see whether these variables influence or correlate with the future outcomes of children.

Although there is a plethora of research hypothesizing a positive relationship between family income and education specifically, it has not been elaborated on or specified the exact form of functional relationship between the two. For simplicity, a regression analysis is run for family income, gender and race on educational attainment to cover aspects of the thesis problem. This is the most suitable research method because it can be tested through a statistical analysis to best determine whether the prediction is confirmed, or not. By exploring the possible relationships between education and the three dependent variables family income, gender and race, this study performs an analysis of the common assumptions. This analysis is carried out with a linear regression on the variables. Understanding these factors, the relationship, and the issues surrounding them each variable can ultimately aid in contributing to improvements in policies and also creating upward mobility in the United States.
Chapter 2: Review of Literature

2.1 Educational Attainment

Education is widely accepted to be a fundamental resource, for both individuals and societies. In the United States and other countries, basic education is perceived not only as a right but also a duty. America has always taken pride in being the land of opportunity, and a country in which, if you work hard and sacrifice will lead to a better life for one's children. Since 1940, the nation has made giant strides in educational attainment, however beginning in the 1970s economics changed favoring highly educated workers. Meanwhile, there is also a shift in the single-parent families demographics producing growing income gaps between high- and low-income families.¹

Educational attainment refers to the highest level of education completed (for example, primary and secondary school, a high school diploma or equivalent, an associate degree, a bachelor’s degree, or a master’s degree or higher). Considering the purpose of the study, educational attainment is measured upon a less than 9th grade or higher scale. Achievement gaps in education occur when one group of students grouped by race, or gender, outperforms another group. Education provides a window into the racial inequality in the United States, and potentially the nation’s perception of it. Research on “achievement gaps” has shown large persistent test score differences between white, black and Hispanic students, as well as between students from wealthy and poor families.²

² Jon Valant and Daniel Newark, “Race, class, and Americans’ perspective of achievement gaps,” BROOKINGS (January 16, 2017), accessed May 1, 2018,
Understanding educational attainment in the United States provides an understanding of
the constant cycle being replicated in society. The cycle replicates when a parent does not have a
certain level of educational attainment or a large amount of family income. This makes it
extremely hard for their child to receive a higher educational level. If it is hard for their child to
receive high educational levels, it will then be hard for their child to receive a large amount of
family income in the future. This, in turn, makes it hard for their child's child to receive a higher
educational level as well. It is a never-ending cycle, and if this cycle continues for generation
after generation it will be hard for an individual to break out of it. This cycle, in other words,
makes it an even bigger challenge for society to work to close the educational achievement gap
and makes it even harder for parents to create upward mobility within their families.

2.1.1 Intergenerational mobility

Intergenerational mobility refers to changes in social status between different generations
within the same family. Depending on where children or grandchildren are in economic
circumstances will determine whether they are experiencing upward or downward
intergenerational mobility. An individual can experience upward or downward mobility for a
variety of reasons such as differences in educational attainment levels and family income, due to
gender, race, citizenship, and credit restraints, just to name a few. It is also possible for a child or
grandchild to be in a better economic circumstance than those of their parents or grandparents.

If the United States were to have a high degree of income mobility, they would be
less concerned about inequalities in any given year, but they do not. Inequality continues to
increase year to year and generation to generation causing a decrease in upward economic

https://www.brookings.edu/blog/brown-center-chalkboard/2017/01/16/race-class-and-americans-perspectives-of-achievement-gaps/
mobility for especially poor families. Educational attainment and family income are both key determinants in determining the future success of a child, however, truly understanding these factors can also aid in contributing to improvements in policies and in creating intergenerational upward mobility within more families who are at a disadvantage.

2.2 Family Income

Family income plays a fundamental role in a child’s lifecycle. A family’s income is the amount combined in the gross income of, every resident of that household, who is over the age of 15. This includes wage, salaries and any kind of governmental entitlements. (For example, unemployment insurance, disability payments or child support payments received, any personal business, investment, or other recurring sources of income.) The average household income is used as an indicator of the monetary well-being of a country's citizens. Household income determines not only how an individual will persist, but how their child perseveres, learns, and obtains success in the future. More specifically, in regard to educational attainment levels and earnings.

According to the U.S. Census Bureau, the average household income was $73,298 in 2014. However, household income does not explain the whole story. Depending on the family situation and where they live, the average income can vary drastically. For example, a single person household earning $65,751 could have a completely different financial situation than a family of five with the same income. The average American household income by tax filing status is $117,795 for married filing jointly, $64,819 for married filing separately, $35,874 for the head of household, $57,577 for a widower, and $34,940 for filing single with an adjusted

gross income (AGI).\textsuperscript{4} The distribution of the U.S. household income has been imbalanced since 1980. After falling due to the Great Recession in 2008 and 2009, inequality rose again during the economic recovery. The size of a household is not taken into account in these measures because it may distort the analysis among the household income variables. Household income is very important and is also used in different circumstances such as the government and organizations who observes to determine if a person is eligible for certain programs like FHA, nutrition assistance and even financial aid, among other programs.

Many empirical studies find family income to be an important factor in explaining the school success of children.\textsuperscript{5} The mechanism economist offer to explain this family relation is that children from poor families are restricted in their pursuit of more and higher quality education merely because their parents face credit constraints when financing their children’s education.\textsuperscript{6} Parents also face a variety of other challenges prohibiting them from financing their children education such as no, or insufficient, income. The problem is most studies ignore the strong

\textsuperscript{4} The Internal Revenue Service Statistics of Income (2014), accessed April 29, 2018


https://www.jstor.org/stable/1827730

https://www.jstor.org/stable/2729315

https://www.jstor.org/stable/1602387?seq=1#page_scan_tab_contents

https://academic.oup.com/ej/article/115/506/879/5087767
correlation between both family income and educational attainment. The importance of family income is relevant for understanding the dynamics of educational attainment distribution and designing educational policies.

In the U.S. less privileged children are at a disadvantage when it comes to how far they progress in school and how much they earn as adults. These children are less privileged because of their family backgrounds, which places them at a disadvantage academically in school and for their future incomes. Johnathan Eng did a research questioning whether there have been any improvements within the past decades, and also examined whether income inequality and educational inequality are related in any way. Using longitudinal data from the National Educational Longitudinal Study of 1988 (NELS88), he tested for correlation between family income in eighth grade and education outcomes twelve years later. A nationally represented sample of eighth graders first surveyed in 1988 was used. Johnathan found family income remains an important positive predictor of eventual adult outcomes. The effect persisted even when other characteristics, or variables, are controlled for in a regression framework. (For example, parental education, home environment characteristics, parental involvement, school characteristics, and student ability).

Erik Plug and Wim Vijerberg did a similar study in 2013 investigating family income and whether there is a significant influence explaining the future educational achievement of children. The study used adoptees as their natural experiment because it is believed the evidence in other studies is often tainted by the lack of control for parental ability. Also, because parental

---

ability is often transferred genetically to their child. This study offers a genetically unbiased estimate when determining whether family income remains an important positive factor in a child's future educational attainment levels. Results show family income has a significant effect on educational attainment level. This implies that children with high academic ability, but living with low-income families, will still face inescapable constraints when applying for a school.

(Figure 1)

Duncan and Murnane's study focused mainly on providing an explanation for the rising family income inequalities resulting in inequalities in educational outcomes between children growing up in low and high-income families. Figure 1 above is used to show the average cash income in a particular year (in 2012 dollars) for children at the 20th, 80th and 95th percentile of the nation’s family income distribution in the 1970s, 1990s and in 2010. Compared to 1970, the 2010 cash family at the 20th percentile has fallen by more than 25 percent. The incomes of

---

families at the 80th percentile has grown by 23 percent to $125,000 while the incomes of the richest 5 percent of families rose by even more. The stagnation of the incomes at the lower end of the spectrum is reflected in the nation's child poverty rate which increased by more than six percentage points between 1970 and 2011. The consequence of these changes is that high-income families have a lot more money to spend on their children oppose to lower-income families. These growing income gaps translates into increased gaps in academic achievement and educational attainments levels of children from high and low families. Duncan and Murnane also found that the rise in family income inequality has an influence on the future and financial outcomes for children in the future.

Children from low-income families are at a heightened risk for a number of poor outcomes, including depression, antisocial behavior, poor physical health, and educational failure. Growing up in poverty is generally seen as toxic for children. Candice Odgers did an examination on how both poverty and the growing divide, between low-income children and their peers, may be influencing low-income children life chances. Among wealthy nations, children in countries with higher levels of income inequality consistently face worse challenges when it comes to health, educational attainment, and well-being. It is a double disadvantage when children live and attend school alongside more affluent peers oppose to similarly positioned peers. To understand how the growing gaps are contributing to a rise in educational outcomes, the role of family income and education must first be understood. The gaps between low and high-income families are constantly expanding while the United States is supposed to be

---


a place for better life and opportunities. Inequalities in family income and education mainly refers to unequal distribution of wages and salaries and an expanding achievement gap, which can also be affected by the child’s gender and race.

Family income has an important impact on whether you get into school, and how well you do in the future. There are a number of ways to measure family income such as salaries, parental educational level, neighborhood geographic location, and resources. This study measures family income by wage and salary income versus educational attainment levels. Does the amount of family income determine how much educational attainment a child will receive in the future? By understanding the differences in wage and salary incomes, and the constant cycle being replicated in society, one will also understand how family income influences educational attainment levels. Educational attainment levels influence the future outcomes or salaries of the next generation. Education is supposed to be a way to break the constant cycle and level out the playing field for everyone. Since education is supposed to be a way to break the cycle, the equation tested below questions whether it is true, or whether it is even working? Children from poor families usually cannot get into school because of their family’s constraints. They become even poorer, and then the cycle repeats. Investigating educational attainment, family income, gender and race is important because in order to have a decent middle-class life, and to have a decent amount of family income, today more than ever, you need to do well in school and obtain degrees.

2.3 Gender

For the past fifty years, there has been an unfilled space in the educational achievement gap for males and females in the United States. This is because gender disadvantages have fluctuated over the years. Today, a college education has become increasingly important in the
economy and it is females, not males, who are succeeding in school and higher levels of educational attainment.\textsuperscript{12} Findings also show, across socioeconomic classes, women are increasingly enrolling and completing post-secondary education while rates for men remain stagnant. Meanwhile, the opportunities for people without education are continuing to shrink. For children, of all genders, being raised from poor families this could be the difference between future upward socioeconomic mobility and a lifetime of poverty.

The gender gap in college completion has been a long time in the making. In the early 1900s, when some elite colleges started opening up to women, the women quickly got better grades than men.\textsuperscript{13} In the 1970s, as more women started attending college, they started graduating at higher rates, while men’s enrollment and graduation rates remained relatively flat. It wasn’t until recently, women attending college were mostly from elite families. Now, women from lower-income families are increasing attending college. This is a positive development for women because educational attainment is really important in today’s economy. Out of the 11.6 million jobs created after the recession, 8.4 million of those went to those with at least a bachelor’s degree.\textsuperscript{14} While females across socioeconomic classes are embracing the idea of education being important, and are pursuing post-secondary degrees, the males from lower-income households are not. The problem is males from low-income families appear to struggle more in school than females do. As the gender gap grows, there are wider implications for


society. People are more likely to pair with others who have a similar educational background as them. As more women get more post-secondary degrees than men, women will increasingly find their marriage rate dimming.\textsuperscript{15}

Even though women are on the rise with education, they still face gender inequalities especially when it comes to differences in pay or salaries. Once a woman graduates, and obtains employment, they are more than likely to get paid less than males in their workplace. This is while single-parent households, especially black and Hispanic families, shifts in demographics and are often headed by women. The big question is how is an unequal income distribution reasonable for women or minority races? It is not. The gender gap in pay has narrowed since 1980 but has remained stable over the past 15 years. In 2017 the analysis of median hourly earnings, for both full and part-time workers in the United States, shows women earning 82 percent of what men earned.\textsuperscript{16} Based on this estimate, it would take an extra 47 days of work for women to earn what men did in 2017. The Census Bureau found full time year-round working women earned 80 percent of what males earned in 2016.\textsuperscript{17} A common assumption is women tend to mature and progress at a faster pace than males. The next and upcoming generations will see men succeed, however, they will see more women succeeding when it comes to educational attainment levels.

2.4 Race

Men are known to have less educational attainment than women. Major G. Coleman flipped the script to investigate job skills and black male wage discrimination. He found discrimination towards students in schools is current, however, there is also debates over the causes of wage inequality for black males. Major investigated whether wage inequality has less to do with discrimination and more to do with skill differences. The purpose of the investigation is to examine the impact skill differences have on wage inequalities. It was found that if a white and black men have the same employee’s competitive performance rating, instead of a decrease in racial wage differences, the differences actually increased. Coleman ultimately concluded the wage gap has nothing to do with a gap in skills, simply because of evidence found of racial discrimination in the labor market.18

The United States is one of the most racially and ethnically diverse populations. In regard to education and race, all races do not always have the same opportunities to attend school and do not have the same resources needed to succeed. During the 1990s, the educational attainment for all races increased while the gap between African Americans and non-Hispanic whites decreased. The differences between the races remains the same, especially among those with a bachelor’s degree or higher. The racial achievement gap in the U.S. refers to educational disparities between minority student and Caucasian students. Evidence of the racial achievement gap in the U.S. remains present today because not all groups are advancing at the same rates. The U.S. Census Bureau looked at racial differences in educational attainment and found 92.9 percent of non-Hispanic White Americans, over the age of 18, graduated from high school.19 For

---

19 The United States Census Bureau, “Table 1. Educational Attainment of the Population 18 Years and Over, by Age, Sex, Race, and Hispanic Origin: 2016,” (March 2017), accessed May 1, 2018
African Americans, over the age of 18, the high school graduation rate is 86 percent which is less than non-Hispanic White Americans.

The Census Bureau has a long history of conducting research to improve questions and data on race and ethnicity. It also provided a breakdown by self-identified ethnic groups. For example, as of March 2014, mean household income by ethnicity for Asian is $90,752, white $79,340, Hispanic or Latino $54,644 and black $49,629. The share of non-Hispanic whites who completed four years of high school or more education increased from 86 percent in 2007 to 94 percent in 2017. Over the same period, the percentage of blacks who completed high school or more education increased from 75 percent to 87 percent. Asian Americans have the highest educational attainment of any race, followed by whites who have a higher percentage of high school graduate, but a lower percentage of college graduates. Individuals identifying as Hispanic or Latino had the lowest educational attainment levels. The gap was the largest between foreign-born Asian American, whom 50.1 percent have a bachelor’s degree or higher, and foreign-born Hispanics whom 9.8 percent had the same degree. The racial achievement gap has many individuals and economic implications, however, there have been many efforts in education reform to narrow this gap.

Alicia Brown analyzed how for the first time in American history, the majority of students within the American public system are students of color and how the educational equity promised is still far from reality. After Brown v. Board of Education, some of the nation is able to put "separate but equal" behind them. Although they are no longer segregated racially, there are still many students of color being educated in a system where their skin color, language,

---

20 Alicia Brown, "Educational Equity: The New Institution Revolution" (2016), accessed May 1, 2018
https://www.advanc-ed.org/source/educational-equity-new-institution-revolution
household income, physical/mental ability, and even their zip codes determine the access they have to education. Students who are Black, Latino, American Indian, and Alaska Native are more likely to attend schools with a high concentration of inexperienced teachers.\textsuperscript{21} Only 1/3 of high schools with high numbers of Black and Latino students offer calculus compared to 56\% of high schools with low numbers of these students. It is also found that students of color are more likely than white students to be suspended one or more times.\textsuperscript{22} Educational equity is a civil and human right and it is foundational to exercise these rights. All students are capable of high academic achievement and deserve adequate and equitable resources to help them attain that goal.

Both Erik Plug and Alicia Brown quoted Dr. Martin Luther King Jr. when he stated: ”the job of the school is to teach so well that family background is no longer an issue.”\textsuperscript{23} However, educational opportunities have never been equally available to all students in the United States with regards to their race, ethnicity, home language, family income, gender, or disability. The U.S is unable to maintain the status of the most advanced country in the world especially if there is continued failure to educate a majority of our children. All children in America regardless of their demographic deserve access to quality education, and it is the school's and policymakers’ duty to provide it for them.

\textsuperscript{21} The U.S. Department of Education Office for Civil Rights, Issue No. 4 (March 2014), accessed May 5, 2018 https://www2.ed.gov/about/offices/list/ocr/docs/crdc-teacher-equity-snapshot.pdf
\textsuperscript{22} The U.S. Department of Education Office for Civil Rights, Issue No. 4 (March 2014), accessed May 5, 2018 https://www2.ed.gov/about/offices/list/ocr/docs/crdc-teacher-equity-snapshot.pdf

Chapter 3: Methodology/Theoretical Model

3.1 Data

Data analysis is the process of inspecting, converting and developing data with the purpose of discovering useful information, support in decision making, and informing conclusions. The study uses a multivariate regression analysis method to establish the relationship between education and three variables: family income, gender, and race. The regression model is shown in the equation in the next section, and the periods used is from 2000 to 2016. The years 2000, 2005, 2010 and 2016 are chosen to validate the theory on the relationship between educational attainment and family income, gender, race, and citizenship.

The Statistical Analysis System (SAS) is used for the regression analysis and descriptive analysis is used to analyzed data. Systematic analysis assists in producing data on various regression coefficients, such as serial correlations, analysis of variance, t-test, R square ($R^2$), f-test, intercepts, the standard error, VIF for multicollinearity, White test for heteroskedasticity, $R$ square, and Durbin Watson. The serial correlation tests for the relationship between the independent variable given and dependent variables given over various time intervals. The t-test is used to measure the significance of each individual coefficient, however, it can only assess one regression coefficient at a time. An f-test compares the fits of different linear models and can assess multiple coefficients simultaneously. $R^2$ is the coefficient of determination which is used to measure the explanatory power of the regression model, and the Durbin Watson test is used for autocorrelation in the residuals. It is usually between the numbers 0 and 4, for example, a value of 2 means there is no autocorrelation in the sample. This studies regression analyses, however, focuses mainly on the following tests: t-test, R square, F-test and the Durbin-Watson test.
Variance Inflation Factor (VIF) is used to detect for multicollinearity in the regression analysis. Multicollinearity is when there is a correlation between predictors in the model, which can affect the regression results. It shows how much the variance of a regression coefficient is inflated due to multicollinearity in the model. The White test is used to test for heteroskedasticity, and it has the ability to establish whether the variance of errors in the regression model is constant. The corrected error methodology is used to improve the confidence level, and the three variables are measured with a 95% confidence level. This is because it provides a range of values, which is likely to contain the population parameter of interest.

The second half of the cycle explores the relationship between education and family income, gender and race from the period 1990 to 2000, and it uses primary sources of data. The data was gathered through the U.S. Census Bureau, Current Population Survey and Annual Social and Economic Supplements. The table(s) below shows the number of educational attainment levels needed to receive a certain amount of income (in 2017 dollars). It also shows the educational attainment levels for different genders and race.

3.2 Hypothesis

First, it is assumed all variables have an impact on Education. The regression model is:

\[ E = \beta_0 + \beta_1 I + \beta_2 G + \beta_3 R + \epsilon \]

In which,
E = Education
I = Family Income
G = Gender
R = Race
\( \epsilon \) = error term

This study looks at education (E) as a function of family income (I), gender (G), and race(R). These variables are used to determine how, and whether, family income, gender, and
race are important factors when determining a child's future educational attainment levels. The single equation model assumes that education is linearly related to the variables. The implications of this single equation model hypothesis that the dependent variable (E) is linearly related to the independent explanatory variables (I) (G), and (R). The constant (E) represents the error term, which provides an explanation for the differences between the results of the model and the actual observed results. The regression looks at family income and the amount of education that is attained because of it. It also looks at how the amount of education depends on gender and race.

Family income is expected to be statistically significant, and the main factor in determining a child's future educational attainment or success of a child. If it is statistically significant it would imply that the higher the family income the more education a child receives in the future, vice versa. Gender is expected to be statistically significant if you are a male when determining a child's future education level. If it is statistically significant it would imply males are more likely to have a higher educational attainment level. Race is expected to be statistically significant if you're white when determining a child's future educational level. If it is statistically significant it would imply whites are more likely to have a higher future educational attainment level.

3.3 Econometric Regression Analysis

Based on the econometric model, there is one dependent variable (education) and three independent variables (family income, gender, and race). The regression analysis is through SAS, the multivariable regression yields the following results for:

\[ E = \beta_0 + \beta_G + \beta_R + \beta_I + \epsilon \]

3.3.1 Regression Analysis
<table>
<thead>
<tr>
<th>Census year</th>
<th>N Obs</th>
<th>Variable</th>
<th>Label</th>
<th>Mean</th>
<th>Median</th>
<th>Std Dev</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>809118</td>
<td>FTOTINC</td>
<td>Total family income</td>
<td>284531.93</td>
<td>49900.00</td>
<td>6616452.68</td>
<td>8091118</td>
</tr>
<tr>
<td></td>
<td></td>
<td>INCWAGE</td>
<td>Wage and salary income</td>
<td>26410.43</td>
<td>19000.00</td>
<td>166406.81</td>
<td>8091118</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EDUC</td>
<td>Educational attainment [general version]</td>
<td>7.0122773</td>
<td>6.000000</td>
<td>10.7653380</td>
<td>8091118</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EDUCD</td>
<td>Educational attainment [detailed version]</td>
<td>71.9908204</td>
<td>65.000000</td>
<td>108.7074298</td>
<td>8091118</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SEX</td>
<td>Sex</td>
<td>1.5048818</td>
<td>2.000000</td>
<td>2.2522382</td>
<td>8091118</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RACWHT</td>
<td>Race: white</td>
<td>1.7760772</td>
<td>2.000000</td>
<td>1.8778749</td>
<td>8091118</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AGE</td>
<td>Age</td>
<td>40.9889050</td>
<td>40.000000</td>
<td>53.5988508</td>
<td>8091118</td>
</tr>
<tr>
<td>2005</td>
<td>1683034</td>
<td>FTOTINC</td>
<td>Total family income</td>
<td>69226.77</td>
<td>53000.00</td>
<td>693010.95</td>
<td>1683034</td>
</tr>
<tr>
<td></td>
<td></td>
<td>INCWAGE</td>
<td>Wage and salary income</td>
<td>30116.20</td>
<td>21000.00</td>
<td>420616.22</td>
<td>1683034</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EDUC</td>
<td>Educational attainment [general version]</td>
<td>7.1922781</td>
<td>7.000000</td>
<td>23.9043350</td>
<td>1683034</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EDUCD</td>
<td>Educational attainment [detailed version]</td>
<td>73.7980867</td>
<td>71.000000</td>
<td>241.6201249</td>
<td>1683034</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SEX</td>
<td>Sex</td>
<td>1.5084257</td>
<td>2.000000</td>
<td>5.0533260</td>
<td>1683034</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RACWHT</td>
<td>Race: white</td>
<td>1.7663356</td>
<td>2.000000</td>
<td>4.2773499</td>
<td>1683034</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AGE</td>
<td>Age</td>
<td>41.8926610</td>
<td>42.000000</td>
<td>123.4699452</td>
<td>1683034</td>
</tr>
<tr>
<td>2010</td>
<td>1790038</td>
<td>FTOTINC</td>
<td>Total family income</td>
<td>284861.30</td>
<td>58100.00</td>
<td>14483579.87</td>
<td>1790038</td>
</tr>
<tr>
<td></td>
<td></td>
<td>INCWAGE</td>
<td>Wage and salary income</td>
<td>31199.93</td>
<td>20000.00</td>
<td>451391.39</td>
<td>1790038</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EDUC</td>
<td>Educational attainment [general version]</td>
<td>7.2490256</td>
<td>7.000000</td>
<td>24.1393672</td>
<td>1790038</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EDUCD</td>
<td>Educational attainment [detailed version]</td>
<td>74.7424040</td>
<td>71.000000</td>
<td>240.3135559</td>
<td>1790038</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SEX</td>
<td>Sex</td>
<td>1.5052437</td>
<td>2.000000</td>
<td>5.0633054</td>
<td>1790038</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RACWHT</td>
<td>Race: white</td>
<td>1.7630339</td>
<td>2.000000</td>
<td>4.3062873</td>
<td>1790038</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AGE</td>
<td>Age</td>
<td>42.4494474</td>
<td>43.000000</td>
<td>128.1370420</td>
<td>1790038</td>
</tr>
</tbody>
</table>
The data shown above is from census years 2000, 2005, 2010 and 2016. The dependent variable is educational attainment and for the independent variables are total family income, gender and race. In the year 2000, there are 8,091,118 observations. The mean is the average of the data, which is the sum of all the observations divided by the number of observations. The median is the midpoint of the data set. The midpoint value is the point at which half the observations are below the value. The standard deviation is used to determine how spread out the data are from the mean. For total family income the mean is $284,531.93, and the median amount is $49,900. The standard deviation is 6,616,452.68, which is about 82 percent of the total observations. For educational attainment the mean is 71.9 percent, and the median is 65 percent. The standard deviation is 108.7074298, which indicates how spread out the distribution is. For gender the mean is 1.5042004 and a median of 2, which indicates women are in the average however the midpoint is at males. The standard deviation is 2.2522382, which is indicating how spread out the distribution is. Lastly, for the race variable there mean is 1.7760772 and the median of 2. The standard deviation is 1.8778749, which indicates how spread out the distribution is. Age is not taken into account because it could have gone either way.
In census year 2005, there are 1,683,034 observations which is a big difference from census year 2000. For the total family income, the mean is $69,226.77 and the median amount is $53,000. The standard deviation is 693010.95, which indicates the distribution spread. From census years 2000 to 2005, there is a drastic decrease observed in the average total family income. The decrease observed is over $200,000, and the median increases by approximately $3,100. The observation amount and average total family income have changed, however, the median only shows a minor change. For educational attainment the mean is 73.7 percent and the median is 71 percent. The standard deviation is 241.6201249. For gender, the mean is 1.5084257 and the median of 2 for male. The standard deviation is 5.053320. Lastly, the race the mean 1.7663356 and the median is 2.0 for white. The standard deviation is 4.2773499.

In census year 2010 there are 1,790,038 observations, which is a slight increase from the year 2005. For total family income the mean is $284,861.30, the median amount is $58,100 and the standard deviation is 14483579.87. From census years 2005 to 2010, an increase can be observed for the average total family income by about $215,000. The median shows an increase by approximately $5,100. This is partly because of the increase in observations by 100,000 people. For educational attainment the mean is 74.74 percent and the median is 71 percent, which is the same as census year 2005. The standard deviation for 2010 is 240.3135559. For gender, the mean is 1.5052437 and the median of 2 for male. The standard deviation is 5.0633054, which is a minor change from the year 2005. Lastly, for race the mean is 1.76630339, the median is 2.0 for white, the standard deviation is 4.3062873.

For the last observed census year 2016, there are 1,816,878 observations. For the total family income, the mean is $298,372.79 and the median amount is $67,000. From the year 2010 to 2016 another slight increase can be observed for the average family income by $10,000. The
median family income shows an increase for about $9,000. The standard deviation is 14625714.23. For educational attainment, the mean is 76.32 percent and the median is 71 percent, which is the same median for both census years 2005 and 2010. The standard deviation is 244.4525107. For gender, the mean is 1.5042004 and the median is 2. The standard deviation is 5.1198470. For race variable the mean is 1.7455209 mean, the median is 2.0 and the standard deviation is 4.4602402.

3.3.2 Regression Analysis (2000)

**Year: 2000**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>DF</th>
<th>Parameter Estimate</th>
<th>Standard Error</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>Variance Inflation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Intercept</td>
<td>1</td>
<td>5.08333</td>
<td>0.00493</td>
<td>1031.78</td>
<td>&lt;.0001</td>
<td>0</td>
</tr>
<tr>
<td>FTOTINC</td>
<td>Total family income</td>
<td>1</td>
<td>-6.9295E-8</td>
<td>5.42384E-10</td>
<td>-127.76</td>
<td>&lt;.0001</td>
<td>1.01974</td>
</tr>
<tr>
<td>INCWAGE</td>
<td>Wage and salary income</td>
<td>1</td>
<td>0.00002019</td>
<td>2.202767E-8</td>
<td>916.55</td>
<td>&lt;.0001</td>
<td>1.06391</td>
</tr>
<tr>
<td>AGE</td>
<td>Age</td>
<td>1</td>
<td>-0.00981</td>
<td>0.00006691</td>
<td>-146.60</td>
<td>&lt;.0001</td>
<td>1.01832</td>
</tr>
<tr>
<td>SEX</td>
<td>Sex</td>
<td>1</td>
<td>0.36708</td>
<td>0.00162</td>
<td>226.20</td>
<td>&lt;.0001</td>
<td>1.05780</td>
</tr>
<tr>
<td>RACWHT</td>
<td>Race: white</td>
<td>1</td>
<td>0.71229</td>
<td>0.00191</td>
<td>372.53</td>
<td>&lt;.0001</td>
<td>1.02082</td>
</tr>
</tbody>
</table>

**Analysis of Variance**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>5</td>
<td>111157188</td>
<td>22231438</td>
<td>217563</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>8.09E6</td>
<td>826542610</td>
<td>102.18412</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>8.09E6</td>
<td>937699799</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the above data, from year 2000, the parameter estimates of the model are as followed: $\beta_0 = 5.08333$, $\beta_1 = -6.9295E^{-8}$, $\beta_2 = 0.36708$, $\beta_3 = 0.71229$. The results from the
regression year 2000 are what was expected. This is multivariable regression analysis, therefore, the method of ordinary least squares (OLS) is used. The sample regression function is:

\[ E = 5.08333 + (-6.9295E^{-8}) I + (0.36708) G + (0.71229) R + \varepsilon \]

An interpretation of the coefficients: The coefficient \(-6.9295E^{-8}\) is the partial regression coefficient of total family income. With the influence of gender and race ratios are held constant. As education increases one-unit, total family income goes up \(-6.9295E^{-8}\) percent. The coefficient 0.36708 and 0.71229 tells us the influence of gender and race are held constant.

(1) \textit{t-test}

T-test compares means between two samples and identifies if they are significantly or statically different. There are three coefficients estimated using t-tests. The hypothesized true coefficient is \(\beta_i = 0\). The estimated value for is \(\beta_i = -6.9295E^{-8}\) and the standard error of this estimate is \(se(\beta_i) = 5.42384E^{-10}\). The degrees of freedom is 5. If we assume \(\alpha = 5\%\) and \(t_{\alpha} = 2.57\),

\[ H_0: \beta_i = 0 \text{ and } H_1: \beta_i \neq 0. \]

\[ t = (-6.9295E^{-8} - 0)/ 5.42384E^{-10} = -127.76. \]

Absolute value of \(t\) is less than \(t_{\alpha} = 2.57\), so the null hypothesis is rejected.

The hypothesized true coefficient \(\beta_i = 0\). The estimated value for \(\beta_i = 0.36708\) and the standard error of this estimate is \(se(\beta_i) = 0.00162\) and the degrees of freedom is 5. If we assume \(\alpha = 5\%\) and \(t_{\alpha} = 2.57\), so \(H_0: \beta_i = 0 \text{ and } H_1: \beta_i \neq 0. \)

\[ t = (0.36708 - 0)/ 0.00162 = 226.20. \]

Absolute value of \(t\) is 226.20. larger than \(t_{\alpha} = 2.57\), so the null hypothesis is not rejected.

The hypothesized true coefficient \(\beta_i = 0\). The estimated value for \(\beta_i = 0.71229\), the standard error of this estimate is \(se(\beta_i) = 0.00191\) and the degrees of freedom is 7. If we assume \(\alpha = 5\%\) and \(t_{\alpha} = 2.37\), so \(H_0: \beta_i = 0 \text{ and } H_1: \beta_i \neq 0. \)
t = (0.71229 - 0)/ 0.00191 = 372.53.

Absolute value of t is 372.53 which is larger than $t_{c} = 2.37$, so the null hypothesis is not rejected.

(2) **R square**

From the regression model, R square provides an estimate of how significant the relationship between the model and the response variable is. From the 2000 regression model, R square shows that 11.85% of the plots fit along the line of regression but since the variables were more than one, the adjusted R square provides a better overall explanation. The adjusted R square is the same as R square which implies that 11.85% of the changes in the response variables are explained by the predictor variables.

(3) **F test**

From the analysis of variance table, the F value = 217,563, Pr > F is <.0001. Due to the F value being larger, obtaining an insignificant probability of <.0001 indicates that the null hypothesis is rejected. This confirms the relevance of the modeled equation. The above F-test confirms the results are significant. The significance F value obtained from the T test is larger than the required level of 5% which shows that model is suitable in explaining the relationship between the variables under study.

(4) **Durbin-Watson d Test**

The Durbin-Watson statistic is used to detect autocorrelation.

$H_{0}: \rho \leq 0$

$H_{1}: \rho > 0$

$H_{0}$ (No positive serial correlation) $H_{1}$ (Positive serial correlation)

In our regression model, the numbers used were:

$K = 5$ $n = 1,816,878$ $\alpha = 0.05$

Where:
K is the number of independent variables

n is the sample size

α is the level of significance

For the critical values of Durbin Watson from the Durbin Watson critical table, \( d_L \) represents the lower critical value, and \( d_U \) represents the upper critical value. Test D is compared to \( d_L \) and \( d_U \):

If D is lower than \( d_L \), there is evidence of positive autocorrelation among the residuals

If D is lower than \( d_U \), there is evidence of positive autocorrelation among the residuals

If D is between \( d_L \) and \( d_U \), there is inconclusive.

From Durbin Watson tables, we could know \( d_L = 1.486 \) and \( d_U = 1.311 \).

From the regression results, the Durbin Watson statics \( D = 1.486 \) > \( d_L \) shows positive autocorrelations.

3.3.2 Regression Analysis (2005)

Year: 2005
From the above data, from census year 2005, the parameter estimates of the model are as followed: \( \beta_0 = 5.30357, \beta_1 = 0.00000705, \beta_2 = 0.33620, \beta_3 = 0.44561 \). The results from regression year 2005 is also as expected.

The sample regression function is:

\[
E = 5.30357 + (0.00000705) I + (0.33620) G + (0.44561) R + \varepsilon
\]

An interpretation of the coefficients: The coefficient 0.00000705 is the partial regression coefficient of total family income. With the influence of gender and race ratios are held constant. As education one-unit, total family income goes up 0.00000705 percent. The coefficient 0.33620 and 0.44561 tells us the influence of gender and race are held constant.

(1) \textit{t-test}

There are three coefficients being estimated using t-tests. The hypothesized true coefficient \( \beta_i = 0 \). The estimated value for \( \beta_i = 0.00000705 \) and the standard error of this estimate
is \( se(\beta_i) = 3.090411 \times 10^{-8} \). The degrees of freedom is 5. If we assume \( \alpha = 5\% \) and \( t_\alpha = 3.0124 \), \( H_0: \beta_i = 0 \) and \( H_1: \beta_i \neq 0 \). \( t = (0.00000705 - 0) / 3.090411 \times 10^{-8} = 228.24 \). Absolute value of \( t \) is less than \( t_\alpha = 3.0124 \), so we do not reject the null hypothesis.

The hypothesized true coefficient \( \beta_i = 0 \). The estimated value for \( \beta_i = 0.33620 \) and the standard error of this estimate is \( se(\beta_i) = 0.00347 \) and the degrees of freedom is 5. If we assume \( \alpha = 5\% \) and \( t_\alpha = 3.0124 \), so \( H_0: \beta_i = 0 \) and \( H_1: \beta_i \neq 0 \). \( t = (0.33620 - 0) / 0.00347 = 96.90 \). Absolute value of \( t \) is 96.90 larger than \( t_\alpha = 3.0124 \), so the null hypothesis is rejected.

The hypothesized true coefficient \( \beta_i = 0 \). The estimated value for \( \beta_i = 0.44561 \), the standard error of this estimate is \( se(\beta_i) = 0.00402 \) and the degrees of freedom is 5. If we assume \( \alpha = 5\% \) and \( t_\alpha = 3.0124 \), so \( H_0: \beta_i = 0 \) and \( H_1: \beta_i \neq 0 \). \( t = (0.44561 - 0) / 0.00402 = 110.83 \). Absolute value of \( t \) is 110.83 which is larger than \( t_\alpha = 3.0124 \), so the null hypothesis is not rejected.

(2) **R square**

From the regression model, for the year 2005, R square shows that 14.57% of the plots fit along the line of regression but since the variables were more than one, the adjusted R square provides a better overall explanation. The adjusted R square is the same as R square which implies that 14.57% of the changes in the response variables are explained by the predictor variables.

(3) **F test**

From the analysis of variance table, the F value = 57,430.3, \( Pr > F \) is < .0001. Due to the F value being smaller in value, obtaining an insignificant probability of < .0001 indicates that the null hypothesis is rejected. This confirms the relevance of the modeled equation. The above F-
test also confirms that the results are significant. The significance of the F value shows that the model is suitable in explaining the relationship between the variables.

(4) Durbin-Watson d Test

The Durbin-Watson statistic is used to detect autocorrelation.

H₀: ρ = 0
H₁: ρ > 0

H₀ (No positive serial correlation)     H₁ (Positive serial correlation)

In our regression model, the numbers used were:

K = 5  n = 1,683,034  α = 0.05

Where:

K is the number of independent variables
n is the sample size
α is the level of significance

For the critical values of Durbin Watson from the Durbin Watson critical table, d_L represents the lower critical value, and d_U represents the upper critical value. Test D is compared to d_L and d_U:

If D is lower than d_L, there is evidence of positive autocorrelation among the residuals
If D is lower than d_U, there is evidence of positive autocorrelation among the residuals
If D is between d_L and d_U, the is inconclusive.

From Durbin Watson tables, we could know d_L = 1.548 and d_U = 1.232.

The REG Procedure
Model: MODEL1
Dependent Variable: EDUC Educational attainment [general version]
Census year=2005

| Durbin-Watson D | 1.548 |
From the regression results, the Durbin Watson statics $D = 1.548 > d_L$ shows positive autocorrelations.

### 3.3.3 Regression Analysis (2010)

#### Year: 2010

| Variable | Label                      | DF | Parameter Estimate | Standard Error | t Value | Pr > |t| | Heteroscedasticity Consistent | Variance Inflation |
|----------|----------------------------|----|--------------------|----------------|---------|-------|--------|-------------------------------|-------------------|
| Intercept| Intercept                  | 1  | 5.43119            | 0.01015        | 534.84  | <.0001|        | 0.01286                       | 422.33            | <.0001  | 0                       |
| FTOTINC  | Total family income        | 1  | -7.42512E-8        | 1.167404E-9    | -63.60  | <.0001|        | 1.315399E-9                   | -56.45            | <.0001  | 1.01720                 |
| INCWAGE  | Wage and salary income    | 1  | 0.00001887         | 3.795514E-8    | 497.24  | <.0001|        | 5.955451E-8                   | 316.90            | <.0001  | 1.04438                 |
| AGE      | Age                       | 1  | -0.00587           | 0.00013192     | -44.51  | <.0001|        | 0.00015741                    | -37.30            | <.0001  | 1.01663                 |
| SEX      | Sex                       | 1  | 0.48658            | 0.00337        | 144.28  | <.0001|        | 0.00410                       | 118.74            | <.0001  | 1.03741                 |
| RACWHT   | Race: white               | 1  | 0.43504            | 0.00392        | 110.96  | <.0001|        | 0.00512                       | 84.94             | <.0001  | 1.01414                 |

#### Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>5</td>
<td>142513776</td>
<td>28502755</td>
<td>56654.8</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>1.7966</td>
<td>900556983</td>
<td>503.09547</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>1.7966</td>
<td>104307059</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the above data, from year 2010, the parameter estimates of the model are as followed: $\beta_0 = 5.43119$, $\beta_1 = -7.42512E^{-8}$, $\beta_2 = 0.48658$, $\beta_3 = 0.43504$. The results from the regression year 2010 is as expected.
The sample regression function is:

\[
E = 5.43119 + (-7.42512E^{-8}) I + (0.48658) G + (0.43504) R + \varepsilon
\]

An interpretation of the coefficients: The coefficient \(-7.42512E^{-8}\) is the partial regression coefficient of total family income. With the influence of gender and race ratios are held constant. As education increases one-unit, total family income goes up \(-7.42512E^{-8}\) percent. The coefficient 0.48658 and 0.43504 tells us the influence of gender and race are held constant.

(1) \(t\)-test

There are three coefficients being estimated using \(t\)-tests. The hypothesized true coefficient \(\beta_i = 0\). The estimated value for \(\beta_i = -7.42512E^{-8}\) and the standard error of this estimate is \(se(\beta_i) = 1.167404E^{-9}\). The degrees of freedom are 5. If we assume \(\alpha = 5\%\) and \(t_\alpha = 3.0124\), \(H_0: \beta_i = 0\) and \(H_1: \beta_i \neq 0\). \(t = (-7.42512E^{-8} - 0)/1.167404E^{-9} = -63.60\). Absolute value of \(t\) is less than \(t_\alpha = 3.0124\), so we do not reject the null hypothesis.

The hypothesized true coefficient \(\beta_i = 0\). The estimated value for \(\beta_i = 0.48658\) and the standard error of this estimate is \(se(\beta_i) = 0.00337\) and the degrees of freedom is 3.0124. If we assume \(\alpha = 5\%\) and \(t_\alpha = 3.0124\), so \(H_0: \beta_i = 0\) and \(H_1: \beta_i \neq 0\). \(t = (0.48658 - 0)/0.00337 = 144.28\). Absolute value of \(t\) is 144.28 larger than \(t_\alpha = 3.0124\), so the null hypothesis is rejected.

The hypothesized true coefficient \(\beta_i = 0\). The estimated value for \(\beta_i = 0.43504\), the standard error of this estimate is \(se(\beta_i) = 0.00392\) and the degrees of freedom is 3.0124. If we assume \(\alpha = 5\%\) and \(t_\alpha = 3.0124\), so \(H_0: \beta_i = 0\) and \(H_1: \beta_i \neq 0\). \(t = (0.43504 - 0)/0.00392 = 110.96\). Absolute value of \(t\) is 110.96 which is larger than \(t_\alpha = 3.0124\), so the null hypothesis is not rejected.

(2) \(R\) square
From the regression model, for the year 2010, R square shows that 13.66% of the plots fit along the line of regression but since the variables were more than one, the adjusted R square provides a better overall explanation. The adjusted R square is the same as R square which implies that 13.66% of the changes in the response variables are explained by the predictor variables.

(3) F test

From the analysis of variance table, the F value = 56,654.8, Pr > F is < .0001. Due to the F value being smaller in value, obtaining an insignificant probability of < .0001 indicates that the null hypothesis is rejected. This confirms the relevance of the modeled equation. The above F-test also confirms that the results are significant. The significance of the F value shows that the model is suitable in explaining the relationship between the variables.

(4) Durbin-Watson d Test

The Durbin-Watson statistic is used to detect autocorrelation.

H₀: ρ = 0
H₁: ρ > 0

H₀ (No positive serial correlation)  H₁ (Positive serial correlation)

In our regression model, the numbers used were:

K = 5  n = 1,790,038  α = 0.05

Where:

K is the number of independent variables
n is the sample size
α is the level of significance
For the critical values of Durbin Watson from the Durbin Watson critical table, \( d_L \) represents the lower critical value, and \( d_U \) represents the upper critical value. Test D is compared to \( d_L \) and \( d_U \):

If D is lower than \( d_L \), there is evidence of positive autocorrelation among the residuals.

If D is lower than \( d_U \), there is evidence of positive autocorrelation among the residuals.

If D is between \( d_L \) and \( d_U \), the is inconclusive.

From Durbin Watson tables, we could know \( d_L = 1.568 \) and \( d_U = 1.324 \).

From the regression results, the Durbin Watson statistics \( D = 1.568 > d_L \) shows positive autocorrelations.

3.3.4 Regression Analysis (2016)

Year: 2016
## Parameter Estimates

| Variable | Label            | DF | Parameter Estimate | Standard Error | t Value | Pr > |t|                  | Heteroscedasticity Consistent | Variance Inflation |
|----------|------------------|----|--------------------|----------------|---------|-------|------------------|-------------------------------|---------------------|
|          |                  |    |                    |                |         |       |                  | Standard Error | t Value | Pr > |t|                   |                      |                     |
| AGE      | Age              | 1  | -0.01122           | 0.00012872     | -87.14  | <.0001| 0.00015722       | -71.34          | <.0001  | 1.01549           |                      |                     |
| SEX      | Sex              | 1  | 0.55249            | 0.00338        | 163.57  | <.0001| 0.00433          | 127.69          | <.0001  | 1.03601           |                      |                     |
| RACWHT   | Race: white      | 1  | 0.39309            | 0.00383        | 102.61  | <.0001| 0.00529          | 74.26           | <.0001  | 1.01138           |                      |                     |

<table>
<thead>
<tr>
<th>Analysis of Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
</tr>
<tr>
<td>Model</td>
</tr>
<tr>
<td>Error</td>
</tr>
<tr>
<td>Corrected Total</td>
</tr>
</tbody>
</table>

From the above data, from census year 2016, the parameter estimates of the model are as followed: $\beta_0 = 5.83708$, $\beta_1 = -7.43432E^{-8}$, $\beta_2 = 0.55249$, $\beta_3 = 0.39309$. The results from the regression year 2016 are contrary to the expected.

The sample regression function is:

$$E = 5.83708 + (-7.43432E^{-8}) I + (0.55249) G + (0.39309) R + \epsilon$$

An interpretation of the coefficients: The coefficient $-7.43432E^{-8}$ is the partial regression coefficient of total family income. With the influence of gender and race ratios are held constant. As education one-unit, total family income goes up $-7.43432E^{-8}$ percent. The coefficient 0.55249 and 0.39309 tells us the influence of gender and race are held constant.

(1) **t-test**

There are three coefficients being estimated using t-tests. The hypothesized true coefficient $\beta_i = 0$. The estimated value for $\beta_i = -7.43432E^{-8}$ and the standard error of this estimate is $se(\beta_i) = 1.169997E^{-9}$. The degrees of freedom are 5. If we assume $\alpha = 5\%$ and $t_{\alpha} = 3.0124$, $H_0$:
\( \beta_i = 0 \) and \( H_0: \beta_i \neq 0 \). \( t = (-7.43432E^{-8} - 0)/ 1.169997E^{-9} = -63.54 \). Absolute value of \( t \) is less than \( t_\alpha = 3.0124 \).

The hypothesized true coefficient \( \beta_i = 0 \). The estimated value for \( \beta_i = 0.55249 \) and the standard error of this estimate is \( se(\beta_i) = 0.00338 \) and the degrees of freedom is 5. If we assume \( \alpha = 5\% \) and \( t_\alpha = 3.0124 \), so \( H_0: \beta_i = 0 \) and \( H_1: \beta_i \neq 0 \). \( t = (0.55249 - 0)/ 0.00338 = 163.57 \). Absolute value of \( t \) is 163.57 larger than \( t_\alpha = 3.0124 \), so the null hypothesis is rejected.

The hypothesized true coefficient \( \beta_i = 0 \). The estimated value for \( \beta_i = 0.39309 \), the standard error of this estimate is \( se(\beta_i) = 0.00383 \) and the degrees of freedom is 5. If we assume \( \alpha = 5\% \) and \( t_\alpha = 3.0124 \), so \( H_0: \beta_i = 0 \) and \( H_1: \beta_i \neq 0 \). \( t = (0.39309 - 0)/ 0.00383 = 102.61 \). Absolute value of \( t \) is 102.61 which is larger than \( t_\alpha = 3.0124 \), so the null hypothesis is not rejected.

(2) **R square**

From the regression model, for the year 2016, R square shows that 12.88\% of the plots fit along the line of regression but since the variables were more than one, the adjusted R square provides a better overall explanation. The adjusted R square is the same as R square which implies that 12.88\% of the changes in the response variables are explained by the predictor variables.

(3) **F test**

From the analysis of variance table, the F value = 53,730.0, Pr > F is < .0001. Due to the F value being smaller in value, obtaining an insignificant probability of < .0001 indicates that the null hypothesis is rejected. This confirms the relevance of the modeled equation. The above F-test also confirms that the results are significant. The significance of the F value shows that the model is suitable in explaining the relationship between the variables.
(4) Durbin-Watson d Test

The Durbin-Watson statistic is used to detect autocorrelation.

\[ H_0: \rho = 0 \]
\[ H_1: \rho > 0 \]

\( H_0 \) (No positive serial correlation) \hspace{1cm} \( H_1 \) (Positive serial correlation)

In our regression model, the numbers used were:

\[ K = 5 \quad n = 1,816,878 \quad \alpha = 0.05 \]

Where:

\( K \) is the number of independent variables
\( n \) is the sample size
\( \alpha \) is the level of significance

For the critical values of Durbin Watson from the Durbin Watson critical table, \( d_L \) represents the lower critical value, and \( d_U \) represents the upper critical value. Test D is compared to \( d_L \) and \( d_U \):

If D is lower than \( d_L \), there is evidence of positive autocorrelation among the residuals
If D is lower than \( d_U \), there is evidence of positive autocorrelation among the residuals
If D is between \( d_L \) and \( d_U \), the is inconclusive.

From Durbin Watson tables, we could know \( d_L = 1.593 \) and \( d_U = 1.356 \).

<table>
<thead>
<tr>
<th>The REG Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model: MODEL1</td>
</tr>
<tr>
<td>Dependent Variable: EDUC Educational attainment [general version]</td>
</tr>
<tr>
<td>Census year=2016</td>
</tr>
<tr>
<td>Durbin-Watson D</td>
</tr>
<tr>
<td>Number of Observations</td>
</tr>
</tbody>
</table>
From the regression results, the Durbin Watson statics $D = 1.593 > d_L$ shows positive autocorrelations.

3.3.5 Tables

Table 1: Median Income by Education Level by Year (in real dollars)

Table 1 relates to the hypothesis because it is a perfect example of how educational attainment can be a determinant on the amount of money one will make in their future. It has been a determinant since 1991 that as the educational attainment levels increase so does the income, or salary, obtained. The data in Table 1 shows, for the year 2000, the more educational attainment there is, the more money one will make in the future. In 2000, high school graduates
made $52,139 which is about $20,000 more than those who had no high school diploma at all. No diploma made about $32,363. However, the high school graduates made about $20,000 less than those who have associate degrees. The gap extends further and further as the educational attainment increases. For example, those with a master’s degree, or higher, automatically makes about $100,000 or more. That is approximately $75,000 more than those who have less than 9th grade, or some high school but no diploma.

For the year 2005, there are consistent patterns for educational levels and more money being obtained in the future. In 2005, high school graduates made $48,055 which is about $17,000 more than those who have no diploma, they also made about $20,000 less than those who have associate degrees. No diploma made $31,048. Educational attainment does not seem to hold as much value as it did in 2000, however it is still holding value for those who have associate degrees or higher. Starting in 2004 professional or doctorate degrees can be observed as having a change in value, but still remains consistent around $100,000. Another consistency can be found between 2000 and 2005, whereas having no high school diploma automatically puts earnings at around $30,000. That is about $70,000 less than those with higher educational levels.

For the year 2010, high school graduates made $43,810 which is about $16,000 more than those who have no diploma. No diploma made $27,896. The high school graduates also made about $20,000 less than those who have associate degrees, which is consistent with the years 2000 and 2005. The value of educational attainment starts to pick back up in 2009 for professional and doctorate degrees. For the year 2016, high school graduates made $44,263 which is about $16,000 more than those who have no diploma. This is consistent with the year 2010, no diploma made $28,982. The high school graduates made about $20,000 less than those who have associate degrees, which is also consistent with the years 2000, 2005, and 2010.
From Table 1 it can be concluded that the more education you have the more income you will receive in the future. This pattern goes as far back as 1990 and is still seen to be true in recent years. The difference between a bachelor’s degree and not having a diploma is about $61,000, and the difference between a high school diploma and a bachelor’s degree is about $48,000, in median salary. These differences alone show the importance of educational attainment towards one’s future. The same patterns are shown for master’s degrees and higher. Once an individual begins to obtain a bachelor’s degree or higher, that is when the incomes begin to show major increases.

Table 2: Difference Income by Education Level by Year (in real dollars)

Table 2 further explains the differences between income and educational attainment levels by degrees and year. The difference between less than 9th grade and a bachelor’s degree is $62,727 for the year 2000, $60,082 for the year 2005, $57,160 for the year 2010, and $63,227 for
the year 2016. The difference between a high school diploma and a bachelor’s degree is $42,951 for the year 2000, $43,075 for the year 2005, $41,246 for the year 2010, and $47,946 for the year 2016. These differences range between $40,000 and $60,000 every year, which is not even equivalent to the salary that is obtained for a high school diploma graduate. This validates educational attainment as a major determinant in an individual’s future income or salary.

3.5 Results

Findings consistently show family income playing a significant role in a child's, or an individual, future educational attainment levels. Finding also show society doing better as a whole until the last year 2016. Children growing up in poor families are often observed as having less educational attainment for many reasons. However, there has also been important in some areas for females and all races. Findings also show gender and race playing a significant role in a child's future educational attainment levels if you are a female, and if your race is white. Males were assumed to have more education than females, however, statistics show there have been big strides where females have reached in educational attainment levels. Whites were assumed to have more educational attainment than other races, however, statistics show there have also been strides where other races such as African American and Hispanics are showing improvements. Even though females show improvements in educational attainment over the years, they are still facing discrimination and inequalities in the workforce and pay. Even though other races are showing improvement in educational attainment over the years, they are still facing inequalities and discrimination in the workforce, and opportunities in general.

An inquiry was raised on whether an individual must just graduate from college to gain higher income, or do they also need to do well? There is no real way for me to look at this, however, this study was able to identify that there is a certain level of educational attainment
needed to earn a certain amount of income or salary. There is a strong relationship between educational attainment and family income. The more income a family has the higher a child's educational attainment will be, which implies a higher lifetime income. However, for a child starting out in a poor family they are less likely to receive higher education, and because they do not receive much education, they do not receive much income. Because they do not receive much income, they are now poor, which is the repeating cycle within the United States. This is the cycle that overtime keeps individuals’ poor from generation after generation.

The issue is pressed on the fact that richer people tend to have a history of sending their children to college. If parents have a college education, they tend to want to send their child to college too. If parents only have a high school degree, they may not think college is as important, or if they do not have the funds to send their child to college, they may just simply not send them to college. My theory is the richer you are, the more likely you are to send your child to college. Family income, or background, determines how much education an individual receives in the future. The amount of education a child receives is dependent on the circumstances and incomes of their parent(s). I controlled for gender because I questioned whether wealthy families are more likely to send men to college as opposed to women. The variable race was controlled to get a better idea of how we are doing as a society.

Chapter 4: Implications and Predictions

Family income was expected to be statistically significant, and the main factor in determining a child's future educational attainment or the success of a child. Findings show that it is significant, and being that it is statistically significant, this implies the higher the family income the more education a child receives in the future, vice versa. Gender was expected to be
statistically significant when determining a child's future education level if you are a male. Results show as being not statistically significant, which implies females are more likely to have higher educational attainment levels than males. Race is expected to be statistically significant when determining a child's future education level if you are white. Results show that it is statistically significant, and this implies that if you are white are more likely to have a higher future educational attainment level.

It has always been hypothesized and tested that the more family income an individual has, the more education their child will receive in the future. The analysis clearly shows the positive impact on educational attainment levels and family income. The more education received, you get more money earned. Various studies also indicated education and family income as playing a major role in the widening of the income and achievement gaps. Some authors even claim the increase in the achievement gaps is a direct influence from gender and race. It is particularly true; gender and race do play a role in the widening of the achievement gap. The amount of money an individual receives only increases drastically as their educational attainment levels do. On the contrary, not all individuals are given the same opportunities to earn a high salary paying job, have resources or gain higher education. This implies family income is important and matters significantly when it comes to obtaining a higher educational level.

Gender and race have many factors attributing to the challenges they face with educational attainment and family income. For gender, males are always looked at as the head of the household, however the increase in single-family homes it changed the game. For the race variable, it is significant, but not as significant as expected because of the rise in educational attainment against all races. Especially for the races that were never in the playing field from the start. The problem is with the system and the fact that the inequality gaps are continuing to widen
all across the board, and never really seems to get smaller. The individuals who do not have money are unable to get a higher education, and because they are unable to get a higher education they end up in a vicious cycle. They then remain stuck in this vicious cycle that is extremely hard to get out of, leaving them in poverty and unsuccessful.

Chapter 5: Conclusions

The United States statistics is constantly showing repetition in high levels of significance when it comes to the inequalities in family incomes and opportunities. How can America be looked at as a country full of endless opportunities when the gap between the rich family income and the poor family income is continuously growing apart and remains that way throughout generations. This study focused on how are we honestly doing as a society? When it comes to family income and education attainment results shows we are doing better in some aspect, however, there are more results showing how we are can improve as a society. Education is supposed to be a way to break the constant cycle and level out the playing field for everyone. Since education is supposed to be a way to break the cycle, the equation tested above questioned whether it is true, or whether it is even working? The answer is education does help to break the cycle, it also shows society doing very good for the first three years and then declining the final year.

Family income and education will always be a major concern in research, and as a topic in society. These two components are proven to be major concerns because of their significance when determining future success for a child or an individual. Various studies have investigated the correlations and impacts of these two factors. This study focused on testing the hypothesis that family income determines how much education a child will receive in the future. Its findings show to be consistent with other studies because the amount of family income obtained does
determine how much education an individual will be received in the future. It also focused on exploring the possible relationship between education and family income, and other factors such as gender and race. This linear regression analysis on family income, race, and gender versus the educational attainment received showed findings of family income continuing to play a significant role in a child's future educational attainment. Findings also showed gender and race playing a significant role in a child's future educational attainment especially if you are a female, and if your race is white.
References


https://www.census.gov/content/dam/Census/library/publications/2017/demo/P60-259.pdf

The United States Census Bureau, “Table 1. Educational Attainment of the Population 18 Years and Over, by Age, Sex, Race, and Hispanic Origin: 2016,” (March 2017), accessed May 1, 2018
https://census.gov/data/tables/2016/demo/education-attainment/cps-detailed-tables.html

Alicia Brown, “Educational Equity: The New Institution Revolution” (2016), accessed May 1, 2018
https://www.advanc-ed.org/source/educational-equity-new-institution-revolution