



A Brief Analysis on the Factors Affecting Higher Education Attainments and the Effects of Higher Education Attainment on Social Well-being

Yue Qu

Teachers College, Columbia University, New York, USA

Email: yq2280@tc.columbia.edu

Abstract. Education attainment refers to the highest levels of education attained by a specific population. In recent years, the importance of higher education attainment has been studied and published a lot. It has been shown that higher level of educational attainment leads to better performances in the job markets. Also, educational attainment is related to positive social outcomes, such as health, population growth, and income distribution. While education attainments have such great benefits, there is a huge disparity of education attainments among children. What factors cause this varying degree of educational attainments? Many studies have already shown that family background variables like family income and parents' educations are correlated to kids' educational attainments. In this research, using data from the 2018 Current Population Survey (CPS) and Annual Social and Economic Supplement, we examine the effects of demographic factors including age, gender, race, and immigration status affect people's levels of educational attainments. We include "age", "female", "race", and "immigration status" as our independent variables. The dependent variable, an individual's level of educational attainment is represented by "years of education". Based on the multivariate linear regression model, we find out that sex and immigration status have significant effects in predicting people's level of educational attainments while age is less important in determining individual's level of educational attainments.

Keywords: Higher education, Multivariate linear regression, Current Population Survey, Annual Social and Economic Supplement

1 Introduction and Motivation

Nowadays, educational attainment has been considered as a significant factor to influence the social well-being. Education attainment refers to the highest level of education attained by a specific population. Higher education has long been considered an integral part of the life of Americans. Various studies have shown that educational attainment nurtures people's social outcomes and promotes active participation in society. Additionally, it has been studied that educational attainment is a stronger predictor of labor market success than measures of cognitive skills and personality characteristics. In the

© The Author(s) 2024

Y. Chen et al. (eds.), *Proceedings of the 2023 3rd International Conference on Modern Educational Technology and Social Sciences (ICMETSS 2023)*, Advances in Social Science, Education and Humanities Research 784, https://doi.org/10.2991/978-2-38476-128-9_9

context of higher education, if education do matter, they will eliminate awkward anomalies in the economic behavior of colleges and have great implications for several policy issues. What factors affect the higher education attainment and why do people earn what they earn is at the heart of recent debates over educational reform.

Despite this longstanding reach, lack of consensus on higher education attainment reflects the principal obstacle for estimating factors and effects on educational attainment, which is endogenous formation of samples. It's difficult to separate the selection effect from actual effect as individuals generally self-select into different educational levels.

Studying factors affecting higher education attainments and the effects of higher education attainment on social well-being is important for a variety of reasons. One reason is that it's meaningful for government to understand the relative importance of higher education versus other inputs in the production function to make budget allocation decisions. Besides, answer such questions would help government to exploit those effects to further social objectives.

In this research, we will first explore and analyze the factors that affect people's levels of education attainment. Following up, we will analyze the benefits of education attainments on aspects of well-being. In the last section, we will summarize our findings and propose some potential policy implications.

2 Literature Review

In recent years, the importance of education attainment has been studied and published in the field of academics. Firstly, a higher level of educational attainment leads to better performances in the job markets. According to the research, workers who are better educated have higher abilities to complete the tasks and gain more from the complex task training compared to workers who had only attended schools for a several years. Additionally, more years of schooling mean that workers are more productive and skillful, thus leading to higher production and output levels. Ultimately, the overall higher levels of education attainment promote the economic development of the society ^[1].

Apart from positive labor market and economic outcomes, educational attainment is also related to positive social outcomes, such as health, population growth, and income distribution (for data set of educational attainment in the world between 1950 and 2010) ^[2].

While education attainments have such great positive benefits to society, there is a huge disparity of education attainment among children of different social and economic backgrounds. What factors cause this varying degree of education attainment among children coming with different backgrounds and social status? Various educational and economic studies have found that education attainments of children vary greatly depending on their family backgrounds. It has been found in many educational and economic studies that background variables including family income, family type, family size, and parents' education determine not only the amount of education children receive but also the quality of their education ^[3].

Specifically, it has been argued that household resources and investments are directly related to the educational attainment of children ^[4]. For parents with greater financial resources, they can not only identify communities with higher-quality schools but can also choose more-expensive neighborhoods and send their kids to high-quality schools. Meanwhile, wealthy parents are willing to invest more on education. For instance, these parents generally devote more resources to ensure that their children can have access to plenty of extracurricular activities not only at schools but also in their communities. On the contrary, for parents who are struggling economically, they simply do not mean or resources to help their kids. In short, wealthy and better educated parents are more capable of helping their kids to better develop competencies.

In addition, parents' education is a powerful predictor of their children's educational attainment ^[5]. On the one hand, through stronger and wider social networks, highly educated parents can use their social capital to help students with their academics and better prepare for future career developments. On the other hand, educated parents can also apply their knowledge and experiences to teach their children certain behaviors and knowledge that are valued by the educational and professional elites. In this way, educated parents can help their kids accumulate human capital directly. In short, parents' education can help children accumulate both human capital and social capital.

Additionally, some studies have found a link between educational attainment and race. According to William Julius Wilson, compared to black people who do not suffer from adverse background characteristics, blacks who experience terrible background conditions are more likely to live in poor neighborhoods which are characterized by relatively high level of poverty and very low-quality schools ^[6]. These people also tend to have much lower levels of educational attainment. Conversely, for poor white people, they are less likely to live in poor neighborhoods. Instead, this group of people generally live in the same neighborhoods as white middle class. Therefore, they can share similar school systems and other resources with the white middle-class people. In short, poor black people living in poor communities are often more likely to be exposed to severe crimes and be deprived of many educational and job opportunities. More importantly, they must deal with living environments that are not suited for learning and education. On the contrary, poor whites often tend to have better living environments and are less likely to be exposed to these negative influences and problem. These characteristics indeed contribute to effective learning. Based on this logic and reasoning, we expect that minorities who have adverse background characteristics will experience greater negative effects on educational attainment.

Gender differences in education attainment has also received a lot of attention in the academic field. In many countries, gender often determines whether a child attend a school. It has been found that across the world, girls are more likely than boys to drop out schools. Also, the poorest girls/women from the most disadvantaged and least developed regions tend to have the lowest educational attainment levels. Women's lower level of education attainment can be attributed to social structures and norms which define certain codes and roles for women. In short, women are less likely to enroll in school at any point in life. Additionally, some research also records the significant gender bias in education attainment. On the other hand, while much research has recorded

the gender bias in education attainment, one study shows that gender gaps is increasingly favor rather than discriminate against females ^[7]. In this study, using data from 38 countries with multiple demographic and Health surveys, Grant and Behrman reexamine the gender gaps in schooling in the developing countries. They found out rather than discriminating against females, gender gaps in education has increasingly favor females. Specifically, they found that among children who have ever attended school, girls younger than 16 years of age have equal or greater schooling progress than boys of the same age in all regions of the developing world. Also, it has been found that in recent years, young females have consistently had higher secondary school completion rates and higher rates of college completion rates than young males ^[8]. Do the gender gaps in education discriminate or favor females? We will explore and answer this question using our dataset and regression model in the following section.

Furthermore, nowadays it has become increasingly important to understand ethnic and immigrant variation in education attainment and achievement for the U.S. population has become increasingly diverse. According to The Census Bureau, while in 2000, 34% of all youth aged 15-19 were from minority groups; it estimates that by 2025, this will increase to 46% (U.S. Census Bureau 2000). While the ethnic gaps in educational attainments have decreased in recent years, substantial gaps still exist, particularly between less advantaged groups such as African Americans, Hispanics, and Native Americans and more advantaged groups such as whites and Asian. Given the increasing share of minority people in the overall population and the large differences in education attainment among different ethnicity groups, it is important to understand the impacts of ethnicity on people's education attainments.

In short, various studies research have already examined and show that family background variables like family income and parents' education are correlated to kids' educational attainments ^[9]. Additionally, many literatures have also indicated that demographics also affect people's education achievements and attainments. While there is literature in these areas, the research is relatively limited. Considering this observation and finding, we decide to further explore this issue. How do people's demographic factors influence their levels of educational attainments? We will investigate this notion by using the Multivariate linear regression model.

3 Description of Data

The data used in this study is an extract of variables from the 2018 Current Population Survey (CPS) and Annual Social and Economic Supplement (ASEC). The Current Population Survey (CPS) is a monthly survey of approximately 60,000 U.S. households conducted by the United States Census Bureau for the Bureau of Labor Statistics (BLS). The sample households are selected by a statistical sampling scheme. There are 66,571 observations in our dataset. This dataset provides information about demographics, educational attainment, and various measures of health and social well-being of the citizens. Additionally, this survey is one of the largest, and most well-recognized surveys in the United States and provides plenty of information on work, our earnings, and our education. Given the authority and comprehensiveness of the CPS survey source, we

think this data is a good fit for our research focus. The data from ASEC is collected in March of 2018, and generally records information for the previous calendar year. In particular, this paper uses the dataset to explore and understand how demographic characteristics like people's age, gender, ethnicity, and immigration status affect people's level of educational attainment. We have cautiously checked for missing variables, thought about appropriate sample restrictions for each analysis, and recoded some of the raw variables to make the analysis more straightforward like recode "educ" to "educat" (a new categorical variable indicating an individuals' highest level of educational attainment).

4 Theoretical Framework

We adopt multivariate linear regression model which can be represented by the equation:

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \varepsilon_i = \mathbf{x}_i^T \boldsymbol{\beta} + \varepsilon_i, i = 1, \dots, n$$

We choose this model for three reasons. Firstly, multiple regression allows us to generate more accurate predictions. Secondly, multiple regression enables us to see how mediators account for an association. Finally, multiple regression can estimate the causal effect of independent variable on the dependent variable. In other words, it can control for bias from the correlated variables. In short, multivariate linear regression allows us to determine the extent to which factors influence our dependent variable while keeping the effects of other variables constant.

In Econometrics, ordinary least squares (OLS) is a type of linear least squares method for estimating the unknown parameters in a linear regression model. OLS selects a set of parameters of linear function to interpret variables according to the least square principle: minimize the sum of squares of the difference between the observed dependent variable (the value of the observed variable) in each data set and the dependent variable predicted by the linear function. When the regression function is exogenous, the OLS estimation is consistent; when the error is homogeneous and the sequence is uncorrelated, the OLS estimation is optimal. Under these conditions, when the error has finite variance, the OLS method provides unbiased estimation of the mean value of the minimum variance. Under the additional assumption that the error obeys the normal distribution, OLS is the maximum likelihood estimator.

Statistical hypothesis test is a method of statistical inference. Generally, two statistical data sets are compared, or the data set obtained by sampling is compared with the composite data set from the idealized model. This paper proposes a hypothesis for the statistical relationship between two data sets and compares it as an alternative to the idealized null hypothesis.

The logic of hypothesis testing is: If the null is true, we can estimate the probability of observing a mean in our sample as extreme than \bar{x} (the "p-value"). If the p-value is sufficiently small (under .05 is a common rule of thumb), we conclude that the null must be false (we reject the null). We will falsely reject the null (reject when the null is true) with probability .05 (that is the standard we use).

Besides, our study on the relationship between higher education attainment and wages is linked the human capital and signaling models. Human capital model argues that education is only a tool of socialization and does not affect the individual's later income, while signaling model believes that education is a screening method, which helps employers to identify which individuals have better talents, attitudes or characteristics, and are competent for vocational training. On this basis, the higher wage of educated labor only means that education plays a role of screening or filtering, but not that education improves the productivity of labor. In human capital model, education is not a simple consumption good but is an investment, which means it's not separable from the person, cannot be sold or transferred or obtained without some participation from the human. In signaling model, acquiring education is a way of signaling pre-existing skills to the employer. Individuals will evaluate costs of obtaining the signal against expected increase in wages and choose the option with the greatest net benefit. If education only plays a role in identifying the merits and demerits of individual talents, then social investment in education is undoubtedly a huge waste of resources, because people can use other more convenient ways to achieve the same purpose.

5 Empirical Framework

Our analytic strategy was like that used in previous studies since the variables in this study allow a reasonable replication of the educational attainment models most estimated. We extracted the data of education attainment and imported it into Stata where we performed the statistical analyses. From the data files, we first extracted: age, gender ("female"), ethnicity ("racecat") and immigration status ("immigrant") as independent variables. This was done for each age group. Then we created a new categorical variable "educat" indicating an individuals' highest level of educational attainment that takes on the following 6 values: less than a HS degree = 1, a HS degree only = 2, some college but no degree = 3, an associate degree = 4, a bachelor's degree = 5, or any graduate degree = 6. The dependent variable is the years of education ("yrsed" "educ"). More deeply, we explored the relationship with education attainment and wages and health. The CPS and ASEC were used to create the variables described below.

1. Age

First, we limit our sample to individuals aged 25 to 54 and address on how educational attainment levels vary across the population. The reason we set the age range is that most younger individuals may not have finished their education yet. Consequently, they are not valid comparison groups (for the variable "yrsed"). So, we do this to capture individuals mostly finished with their schooling, and to avoid the years when workers begin to retire in greater numbers.

Using the variable educ, we describe how educational attainment has changed across 5-year age cohorts (e.g., 26-30 year olds, 31-35 year olds, 36-40 year olds, etc.) Then we transform the highest grade attained information into a year of schooling measure, assigning years to each 'grade' (e.g., high school = 12, college = 16, masters = 18, etc. as in the class notes).

2. Gender

For the past fifty years, there has been a huge gap in the educational achievement of males and females in the United States, but which gender has been disadvantaged has fluctuated over the years. Thus, in this study we choose the variable “sex” (“female” equals to 1 if the sample members are female) to explore how gender effects the educational attainment.

3. Ethnicity

While during the 1990s the educational attainment of all races went up, with the gap between African Americans and non-Hispanic whites decreasing, differences between the races remain, especially among higher education. To study the relationship between ethnicity and educational attainment, we pick the variable “race” (“Hispanic” equals to 1 if the sample members are Hispanic). These racial achievement gaps have many social economic implications and there have been many efforts in education reform to narrow the gaps.

4. Immigration Status

Nowadays, immigration status has been considered as a significant indicator to the educational attainment. A trend becomes visible when comparing the foreign-born to the native-born populace of some races. According to the U.S census, nearly 43.8% of African immigrants achieved the most college degrees, compared to 42.5% of Asian-Americans, 28.9% for immigrants from Europe, Russia and Canada and 23.1% of the U.S. population. So, we create a figure that plots the distributions of educational attainment separately by immigrant status, defined as whether a person is foreign born (i.e., $\text{gen immigrant} = \text{nativity} == 5$) using for the histogram command.

5. Income

Educational attainment is strongly correlated with income in the United States. According to 2003 Educational Attainment in the United States, the most significant average income difference was between those who had some college education or an associate degree and those who had a Bachelor's degree. While those with college averaged \$31,046, those with a bachelor's degree averaged \$51,194, over \$20,000 (64.9%) a year more. Thus, we generate a measure of hourly wages, ‘topcode’ the measure at \$200 an hour, and also recode all hourly wage measures below \$3.63 (half the federal minimum wage) to be missing.

6. Health

Another important outcome that may be affected by education is health. We create a dummy variable “greathealth” equal to 1 if an individual reports being in excellent or very good health, and 0 otherwise.

6 Discussion and Interpretation

6.1 The Effects of Demographic Characteristics on Education Attainment

Based on our intuition and insights gained from various academic articles, we think that demographic characteristics like people’s age, gender, ethnicity, and immigration status affect people’s level of educational attainment. To test for this basic intuition, we plot

some graphs to see if these variables might correlate with education attainments. Figure 1 to 5 are the graphs we obtained using Stata. For instance, figure 1 allows us to visually examine how education attainment changes as people get older. Specifically, we notice that between age 16 and 26, there is a strong positive relationship between years of education. Meanwhile, we also observe that between age 31 and 91, there is a slightly decrease in the years of schooling. The overall pattern suggests that educational attainment has changed across 5-year cohorts. For the purpose of clear and detailed analysis, we limit our focus on the population with age between 25 and 54 years old. We also notice another feature in Figure 3. Figure 3 shows people’s average years of education by age cohorts and race. the overall pattern suggests that compared to White non-Hispanics and Black non-Hispanics, Hispanics in general tend to have fewer years of education. In short, these graphs do suggest that age, gender, race and people’s immigration status are indeed relate to people’s educational attainments. These graphs give us more confidence to include these variables into our raw model.

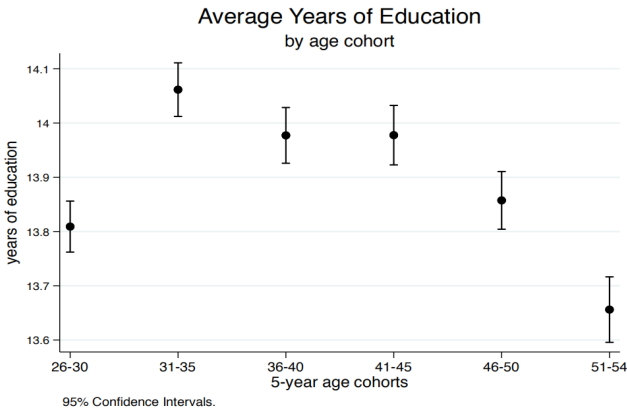


Fig. 1. Average Years of Education by Age Cohort

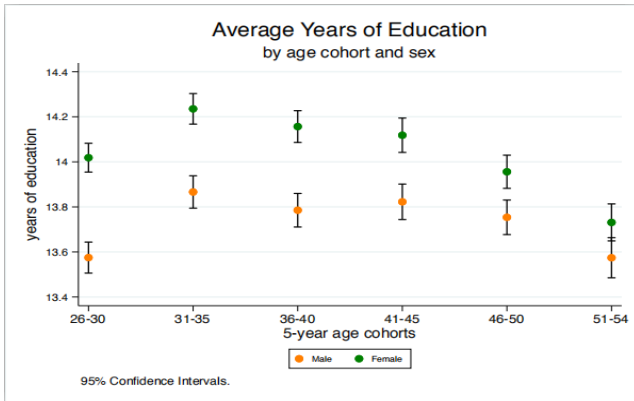


Fig. 2. Average Years of Education by Age Cohort and Sex

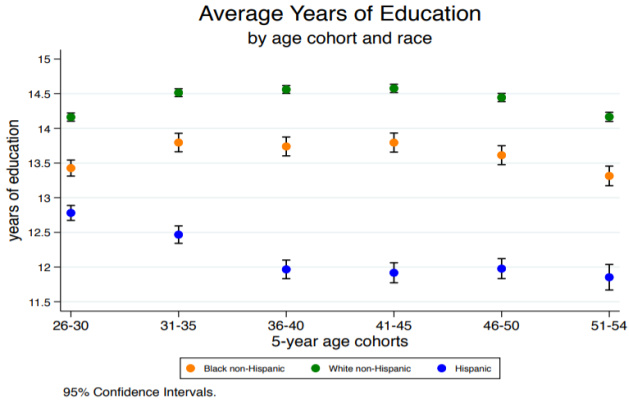


Fig. 3. Average Years of Education by Age Cohort and Race

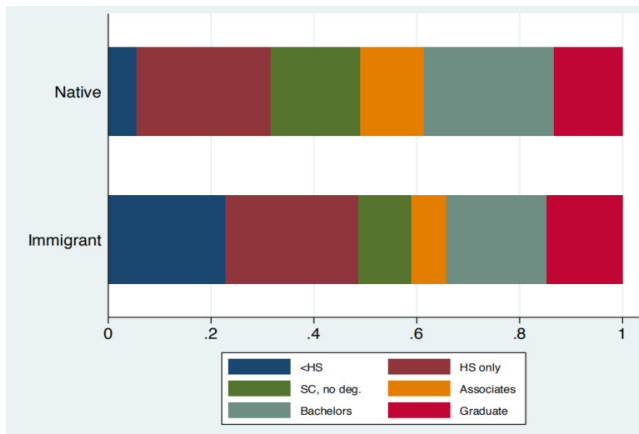


Fig. 4. Average Years of Education by Age Cohort and Race

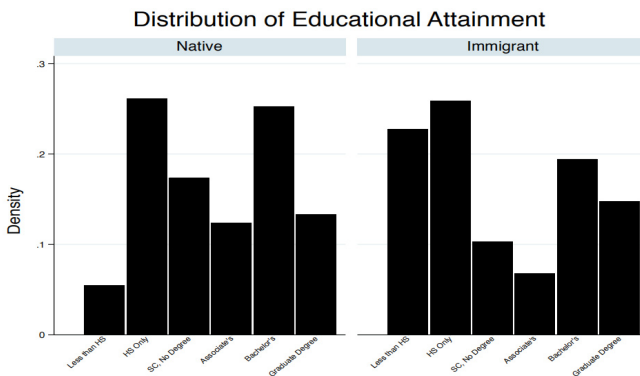


Fig. 5. Distribution of Educational Attainment between Natives and Immigrants

For better model interpretation, we recode the original variable “sex”. Particularly, we generate a new categorical variable called “female” which takes on value of 1 when an individual is a female and equals zero when the observation is a male.

For variable race, we generate a new variable named “racecat” to represent race. “racecat” takes on the following values: “White non-Hispanic” =1, “Black non-Hispanic” =2, “Hispanic” =3, and “Asian non-Hispanic”=4. White non-Hispanic is the reference group.

Additionally, since there is no direct measure for nationality in the data, we generate a categorical variable called “immigrant” (immigrant=1 means immigrant, immigrant=0 corresponds to non-immigrant). Figure 2 shows the distribution of educational attainment between immigrants and natives.

We consider two approaches to represent our dependent variable. In the first place, we recode “educ” education attainment recode, and use this variable to generate a new (numerical) categorical variable named “edcat”. Edcat indicates an individual’s higher educational attainment level that takes on the following values: less than a HS degree = 1, a HS degree only = 2, some college but no degree = 3, an associate degree = 4, a bachelor’s degree = 5. However, since we are treating the categorical variable as if it was a continuous one and this makes the regression model hard to interpret, we turn to look for an alternative. We instead generate a new variable called “yrsed” and use it as our dependent variable. Figure below shows the distribution of higher attainment for our population.

Based on the data summary and data exploring process, we decide to keep age, female, racecat, and immigrant as our independent variables in our regression model. We use “yrsed” as the dependent variable to represent an individual’s level of educational attainment.

Table 1 displays the statistical results estimating the extent to which age, gender, race, and nationality contribute to students’ education attainment levels. Specifically, for variable “age”, holding other variables constant, for one additional year increase in age, on average, we expect the years of schooling to decrease by 0.0077 unit. Since the t-statistic for “age” is larger than the critical value of 1.96 ($p < 0.05$), this variable is statistically significant. However, compared to other variables, “age” variable is less important in predicting people’s level of educational attainments.

For variable “female”, holding other variables the same, female on average has 0.337 units more years of education than men. Since the t-statistic for “female” is larger than the critical value of 1.96 ($p < 0.05$), this variable is statistically significant. Given this observation, our result confirms with the previous finding that rather than discriminating against females, the gender gaps in education not only decrease but also seem to favor females. Additionally, based on the magnitude of the estimated coefficient for “female”, we can see that sex has a relatively larger effects on people’s level of educational attainments.

For variable “immigrant”, holding other variables constant, compared to natives, immigrants on average has 0.887 unit less years of education. Since the t-statistic for “immigrant” is larger than the critical value of 1.96 ($p < 0.05$), this variable is statistically

significant. Furthermore, based on the magnitude of the coefficient estimates for “immigrant”, we observe that immigration status has a significant impact in determining people’s level of educational attainments.

In short, Since the t-statistics for all of our variables are greater than the critical value of 1.96 ($p < 0.05$), we learn that these variables are all statistically significant. These results again support our decision to include these variables in our model. In other words, age, gender, race, immigrant status are indeed determinants of an individual’s level of educational attainments.

Table 1. Linear Regression Model Showing the Effects of Demographic Characteristics on Education Attainment

yrsted	Coef.	Robust Std. Err.	t	P > t 	[95% Conf. Interval]	
age	-.0076331	.0014414	-5.30	0.000	-.0104583	-.004808
female	.3367489	.0240565	14.00	0.000	.2895982	.3838995
immigrant	-.8877352	.0408623	-21.73	0.000	-.9678253	-.807645
racecat	-.2764388	.015119	-18.28	0.000	-.3060721	-.2468055
_cons	14.62713	.0649196	225.31	0.000	14.49989	14.75437

6.2 Benefits of Education Attainments on Aspects of Well-Being

Having examined some of the demographic differences in educational attainment among prime-age individuals, we now turn to analysing how education relates to some important measures of well-being: wages and health.

First, we generate a measure of hourly wages, ‘topcode’ the measure at \$200 an hour, and recode all hourly wage measures below \$3.63 (half the federal minimum wage) to be missing. Then we created 4 race/ethnicity groups as follows: White non-Hispanic (race=100 & hispan=0), Black non-Hispanic (race=100 & hispan==0), Hispanics (hispan = 0), and Asian, non-Hispanic (race=651 & hispan=0). To be specific, we limited the sample to these 4 groups, and use 3 different ways (Figure 6-Figure 8) to plot the conditional expectation functions of hourly wages as a function of educational attainment (use the categorical variable created above) separately by race. Figure 6 shows the difference in levels of hourly wages, which is helpful to see the raw differences between races, within each educational category. But is not very helpful to compare the growth by race, as each race has a different starting level. Figure 7 uses a common point of comparison (HS Only) within each racial group, so is useful to compare the growth of hourly wages by educational group, separately by race. The problem is that we can’t see the raw differences. Figure 8 shows a monotonic transformation (natural logarithm) of the first figure, so it preserves the orders, but could give inaccurate impressions about the level of wages across different race groups.

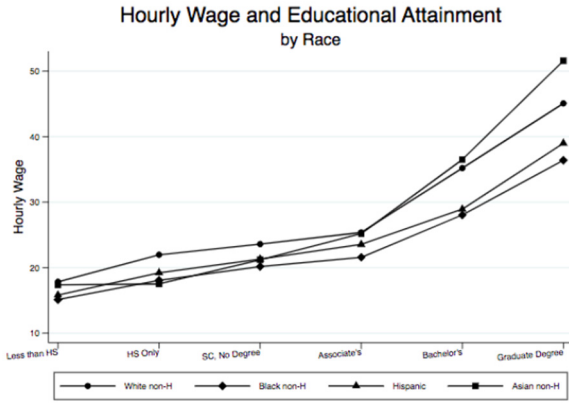


Fig. 6. Hourly Wage and Educational Attainment by Race

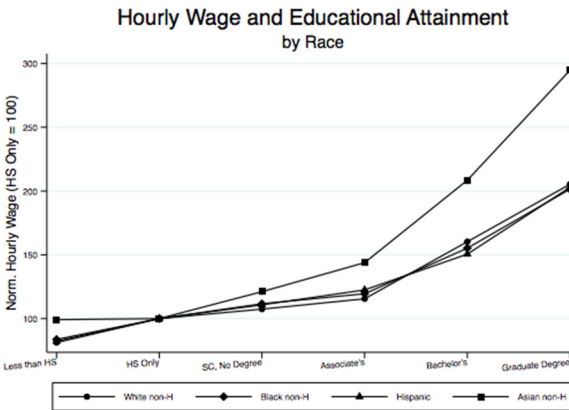


Fig. 7. Hourly Wage and Educational Attainment by Race

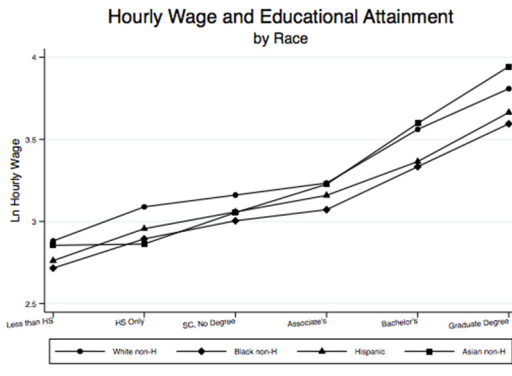


Fig. 8. Hourly Wage and Educational Attainment by Race

Another important outcome that may be affected by education is health. To explore the relation between higher education attainment and health, we create a dummy variable “greathealth” equal to 1 if an individual reports being in excellent or very good health, and 0 otherwise. Figure 9 illustrates how education is related to both wages and better health. The sum of the constant and the “baplus” estimated coefficient, represents the proportion of individuals with Bachelor’s degree or more (baplus = 1), that declared having great health.

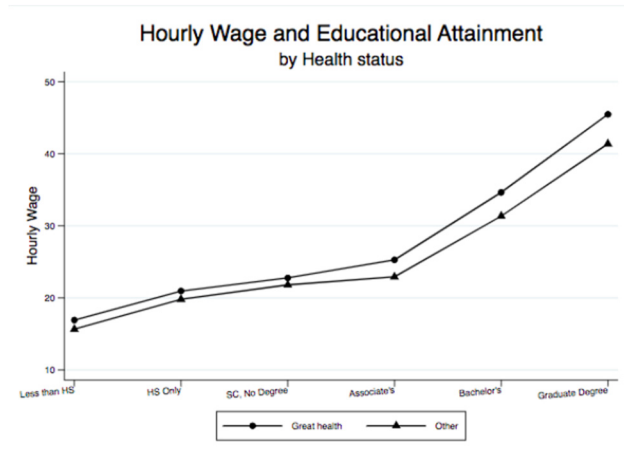


Fig. 9. Hourly Wage and Educational Attainment by Health Status

Table 2 displays the statistical results estimating the influence of yrsed (years of education) on hwage (hourly wage). It represents the predicted average hourly wage for individuals between 25 and 54 years of age that have zero years of education and indicates that an increase of 1 year of education is associated, in average, with 3.09 dollars more of hourly wages for individuals between 25 and 54 years of age.

First, we create a dummy (equal to 1 or 0) variable “baplus” indicating whether a person has at least a bachelor’s degree. Table 3 displays the statistical results estimating the influence of baplus on hwage. The constant 21.20 is the average hourly wage for the group of individuals with less than a BA degree. $_b[baplus]$ represents the average difference between the group of individuals with less than a BA degree and the group with a BA degree or more is 16.38.

Then we create a categorical variable indicating an individuals’ highest level of educational attainment that takes on the following 6 values: less than a HS degree = 1, a HS degree only = 2, some college but no degree = 3, an associate degree = 4, a bachelor’s degree = 5, or any graduate degree = 6. Next, we generate 6 dummy variables edcat1 to edcat 6. Table 4 displays the statistical results estimating the influence of all 6 of these categorical variables on hourly wage. The constant term is the average hwage for the reference group, in this case, category 1 (less than HS). The rest of the estimated coefficients represent the average difference of hwage between each group of edcat and the reference group.

Table 2. Linear Regression Model Showing the Effects of Education Attainment on Hourly Wage

hwage	Coef.	Robust Std. Err.	t	P > t	[95% Conf. Interval]	
yrstd	3.086356	.0410461	75.19	0.000	3.005906	3.166807
_cons	-15.79856	.5490463	-28.77	0.000	-16.87469	-14.72242

Table 3. Linear Regression Model Showing the Effects of Bachelor Degree on Hourly Wage

hwage	Coef.	Robust Std. Err.	t	P > t	[95% Conf. Interval]	
baplus	16.38489	.2178466	75.21	0.000	15.95791	16.81188
_cons	21.19822	.098482	215.25	0.000	21.00519	21.39124

Table 4. Linear Regression Model Showing the Effects of Educational Level on Hourly Wage

hwage	Coef.	Robust Std. Err.	t	P > t	[95% Conf. Interval]	
_edcat1	0	(omitted)				
_edcat2	3.964219	.2846965	13.92	0.000	3.406212	4.522227
_edcat3	5.834817	.3078008	18.96	0.000	5.231525	6.438109
_edcat4	7.93041	.3379985	23.46	0.000	7.267931	8.59289
_edcat5	17.11897	.3247748	52.71	0.000	16.48241	17.75553
_edcat6	28.06211	.4343123	64.61	0.000	27.21085	28.91336
_cons	16.4159	.2416027	67.95	0.000	15.94236	16.88944

7 Limitations and Directions for Future Research

Studies on education attainment are confronted and sometimes complicated by various empirical challenges. First, random assignment of population is wanted for consistent estimation. Endogenous formation of population and selective distribution of education itself often plague studies on education attainment. The internal validity problem is that people with higher education attainment could have systematic differences in observed and unobserved characteristics that affect their wages (for example, family background, ability, effort/motivation, etc.). A simple comparison of the well-being between these two groups of people will capture both the effects of education and any selection bias. To overcome limitations arising from endogenous effects, instrumental variable (IV) would be a useful method for future research. If the data set is infinite, we could exploit extremely comprehensive data sets that contain rich information on one's educational background (e.g., the characteristics of school, teacher, peer groups and family, a history of one's growth in test scores, etc.).

Second, the lack of the pre-determined ability of the population in our data may expose this study to a limitation of omitted variable bias. The data is likely to be irregular and hard to collect as to be of little practical value. To detect the long-run effects of education, the sample must be restricted to a certain age range. For instance, the influence of education attainment on people in their twenties and early thirties are plagued by the late arrival on the labor market of those with more formal schooling; To be specific, it's difficult for medical students to earn their professional salaries until they reach their thirties ^[10]. Next researcher in this area should try to address the problem by collecting a sample of population today, and then doing field study to explore into their past to find out how they performed in their youth or try to find a uniform data base collected years ago and try to collect uniform data on the current performance of sample members. A fruitful area of future research would be to examine similar data in other educational settings and other countries to conduct a comparison.

8 Conclusions

To conclude, Higher levels of education attainment contribute to not only better job market outcomes but also generate various social benefits for the society. While these benefits are significant and large, a huge disparity of education attainment exists among different children across the world. In using multivariate linear regression model, we have found out that demographic characteristics including individuals' age, gender, race/ethnicity, and immigration status are determinants for people's levels of education attainment. To be more specific, female on average is likely to have 0.33 more years of schooling than men; immigrants on average tend to have 0.88 years of schooling than the non-immigrants/natives. In short, among these four variables, our model indicates that gender and immigration status have relatively stronger effects on individuals' levels of educational attainment.

Since people have studied the impact of education on income, analytical techniques and data collection methods have made great progress. Multiple regression analysis and income function have also been used to analyze income. We have every reason to believe that educational attainment and age are the most powerful determinants of income, even when multiple variables are considered in a broad area. In other words, gender, race, occupation, talent, luck and so on may be important factors affecting income, but all these factors together are not enough to explain the reasons for income differences.

Although higher education attainment and social well-being are not necessarily directly related, with the continuous improvement of economic development, higher education should also be able to become a signal of ability ^[11]. We know that when enterprises employ workers, they are not clear about the ability of each worker. However, people with different abilities have different contributions to the enterprise. If they pay the same salary without analysis, it will not be conducive for those with high abilities to play their advantages and create more value. According to the educational signaling theory of information economics, education can be used as a filter to classify individuals with different productivity, to reveal the hidden ability information to employers ^[12].

Reference

1. Barro, R. J., & Lee, J. W. (2001). International data on educational attainment: updates and implications. *Oxford Economic Papers*, 53(3), 541-563.
2. Barro, R. J., & Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*, 104, 184-198.
3. Hall, R. E., & Jones, C. I. (1999). Why do some countries produce so much more output per worker than others?. *The Quarterly Journal of Economics*, 114(1), 83-116.
4. Becker, G. S. (1993). Nobel lecture: The economic way of looking at behavior. *Journal of Political Economy*, 101(3), 385-409.
5. Haveman, R., & Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature*, 33(4), 1829-1878.
6. Wilson, W. J. (1991). Studying inner-city social dislocations: The challenge of public agenda research: 1990 presidential address. *American Sociological Review*, 1-14.
7. Grant, M. J., & Behrman, J. R. (2010). Gender gaps in educational attainment in less developed countries. *Population and Development Review*, 36(1), 71-89.
8. Buchmann, C., & DiPrete, T. A. (2006). The growing female advantage in college completion: The role of family background and academic achievement. *American Sociological Review*, 71(4), 515-541.
9. Jang, S. T. (2020). The schooling experiences and aspirations of students belonging to intersecting marginalisations based on race or ethnicity, sexuality, and socioeconomic status. *Race Ethnicity and Education*, 1-22.
10. Beutel, A. M., & Axinn, W. G. (2002). Gender, social change, and educational attainment. *Economic Development and Cultural Change*, 51(1), 109-134.
11. Taylor, L. K., Bentley, J., Hunt, J., Madden, R., McKeown, S., Brandt, P., & Baker, D. (2012). Enhanced reporting of deaths among Aboriginal and Torres Strait Islander peoples using linked administrative health datasets. *BMC Medical Research Methodology*, 12, 1-8.
12. Xie, Y., & He, J. (2023). Re-examining cultural reproduction theory: cultural capital and adolescent academic achievement in China. *Asia Pacific Education Review*, 1-15.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

