

The Influence of Education Level, Gender, Race, Marital Status, Age, and Occupation on the Wage of the General Population

Ronghai Xie^{1,*}

¹ Boston College Morrissey College of Arts and Sciences

* Corresponding author. Email: xierb@bc.edu

ABSTRACT

This article aims to analyze how education level, gender, race, marital status, age, and occupation influence the wage of the general population. Using the Family files of 2019 data from the Panel Study of Income Dynamics (PSID), this paper identifies the importance of maintaining equality in the labor market. After analyzing the data, this paper finds that years of education tend to have a more significant impact on people's income for people with more years of education and a higher degree of education. In addition, the discrimination of race and gender still exists in the labor market, and these discriminations deepened the wage inequality in our society today. Also, marriage impacts females' and males' income differently, and females' income is hurt more than males' income when getting married.

Keywords: wage, labor economics, wage equality, education, industry

1. INTRODUCTION

The labor market has been one of the most important markets in our economy, and wages, as the essential factor within the market, influence people's behaviors in the labor market. This research paper explores how different factors influence the annual income of the general population and how the wage differences imply the broader social issues that exist in our society today.

In the past decades, the income inequality in the United States significantly increased [1]. What factors can influence or determine the wage for the general population? The education level is undoubtedly one of the most important determinants of wages. People with higher levels of education are expected to earn higher salaries than those who do not since a higher level of education means more knowledge and better performance in the workplace. In modern society, people strive to attend better universities and earn a higher degree to receive better returns in the future. Gender and race are controversial topics throughout contemporary history that people have been calling and fighting for equal payment among different genders and races. The strong presence of African-American workers at the turn of the twentieth century became an influencing factor in the labor market. Also, according to the labor market trend in the past one hundred years, women's

participation rate in the labor market has increased drastically, and women now comprise 60% of the enrollment of universities and colleges. However, there still exist wage differences and discriminations around these two factors. Age also influences wages as people are expected to accumulate experience and knowledge over their lifetime. With more experience, the expectation for wages would increase accordingly. Besides, marital status could also influence wages. If people have family and children to take care of, they would be expected to allocate some time for the family instead of work. In this case, marital status could affect wages. Lastly, occupation can make a big difference in the workers' wages. Workers in different industries and different occupations earn different wages. For example, workers in the financial sector are generally expected to earn more than workers in the production sector.

2. LITERATURE REVIEW

Age, education, race, marital status, gender, and occupation affect the wages of the general population. As employment becomes a significant issue and catches more and more people's attention, wages also become part of the issue. Many researchers conducted experiments and researches to test how different factors make a difference in the workers' wages.

In Siwei Cheng's article "A Life Course Trajectory Framework for Understanding the Intracohort Pattern of Wage Inequality," one of the properties of the LCT framework points out that the influence of intracohort heterogeneity on the baseline wage and wage trajectory [1]. It shows that wages could be influenced by various reasons, including age, education, race, gender, marital status, and so on. For example, working mothers may give up jobs with higher wage growth rates in exchange for flexible schedules to find a balance between family time and work. Another example would be education. Since people with higher levels of education tend to have more networking and a "greater stock of human capital," it could influence their wages in the first place [1]. People with higher education levels might also have higher status and better job performance during their career, further lifting their confidence and motivation in the workplace. As Cheng points out in her article, these different personal factors over the baseline wage and trajectory are influential in determining wages.

What's more, in the article "Gender Difference in Employment and Income in China's Labor Market," Jongsung Kim pays closer attention to the gender inequality in the labor market and how this inequality affects the wages for males and females. He claims a significant difference in employment, experience, and wages. He states that although "women, on average, have lower educational attainment than men, the gap is less than one year (0.58 years). On the other hand, there is a significant gender gap in employment: 78.5 percent of men and 65.6 percent of women are 'working or employed'" [3]. The workplace also favors men instead of women because, on average, men "have on average 0.67 more years of work experience than women" [3]. The wages women earn are also less than men. "Women earn 9,320.3 yuan per year on average, about 87 percent of what men earn on average" [3].

Besides, in Paul Devereux's article "Effects of Industry Growth and Decline on Gender and Education Wage Gaps in the 1980s," he focuses on how education affects wages. He claims that there is a 30% to 50% increase in wages for educated groups than uneducated groups, indicating a huge difference between the wages of people with more education and people with less education. It shows that education is one factor that can determine wages [4].

Last but not least, in the article "Quality Education and Skills of Indonesian Labor, Towards Equality Wages in Foreign Countries," the author Febriani uses Test Crosstab and Chi Square test to prove that the quality of education and skills of Indonesia labor helps them earn wage equality in foreign countries [5]. Similarly, in the article "Decomposing Vietnamese gender equality in terms of wage distribution," the authors Vu and Yamada find that there is consistent gender wage gap in Vietnamese labor market. The price of skills being the

main contributor of the wage gap, while the labor market discrimination does not have a crucial influence [6]. They argue that the rise in women's education will contribute to the elimination of wage inequality in Vietnam.

3.RESULT & ANALYSIS

This research project analyzed the Family files of 2019 data from the Panel Study of Income Dynamics (PSID). The variables used for analysis include the Annual Income of Individuals, Education Level, Marital Status, Age, Gender, Race, and occupation. By applying multivariable regression analysis, this project discussed how the above variables influence the wage of the general population.

Table 1 shows the summary statistics for this research project. As shown in this table, the variable "Income" is the dependent variable, and the rest of the variables are independent variables. The variable "age" is a continuous variable representing the individual's actual age, and "age2" is age squared, allowing for the non-linear relationship between income and age. The variable "female" is a dummy variable. The variable female equals 1 when the individual is a female and equals 0 when the person is a male. The variable "education" is the education level of individuals, ranging from 0 to 17, and "educ2" and "educ3" each represent the education squared and the education to the third power. The variable "feduc" is the interactive variable between female and education, which allows for the different effects of education on income based on whether the individual is a female or a male. The variable "race" is a dummy variable that will be 1 when the individual is White and 0 when the individual is non-White. The variable "raceeduc" is an interactive variable between race and education. It represents the different effects of education on income based on whether the individual is white or not. The variable "marriage" is a dummy variable. When an individual is married, the variable marriage equals 1. Otherwise, it equals 0. The variable "femalemarr" is an interactive variable between females and marriage, accounting for the different effects of marriage on income based on whether the individual is a female or a male. The variables "management," "educjob," "foodindustry," "production," "sales," and "transportation" are all dummy variables and represent different occupations. If the individual is in the industry, the variable would be 1. Otherwise, the variable would be 0.

Table 1. Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Income	8628	41884.74	62359.95	0	2000000
age	8628	46.79659	16.60781	18	102
age2	8628	2465.709	1719.258	324	10404
female	8628	0.3275382	0.4693426	0	1
education	8561	13.51478	2.501177	0	17
educ2	8561	188.9043	64.22188	0	289
educ3	8561	2711.144	1317.562	0	4913
feduc	8561	4.380563	6.412132	0	17
race	8628	0.5807835	0.4934595	0	1
raceeduc	8561	8.044037	7.105582	0	17
marriage	8628	0.4364859	0.4959783	0	1
femalemarr	8628	0.0068382	0.082415	0	1
management	8628	0.2834956	0.450721	0	1
educjob	8628	0.0328002	0.1781236	0	1
foodindustry	8628	0.041261	0.01989048	0	1
production	8628	0.0579509	0.233664	0	1
sales	8628	0.0594576	0.2364928	0	1
transportation	8628	0.080204	0.2716245	0	1

Table 2. Regression Analysis

Source	SS	df	MS	Number of obs = 8561		
				F (17,8543) = 134.15		
Model	7.0475e+12	17	4.1456e+11	Prob > F = 0.0000		
Residual	2.6401e+13	8543	3.0904e+09	R-squared = 0.2107		
Total	3.3448e+13	8560	3.9075e+09	Adj R-squared = 0.2091		
				Root MSE = 55591		
Income	Coefficient	Std. Error	t	P> t	[95% Confidence Interval]	
age	2742.025	215.3843	12.73	0.000	2319.82	3164.23
age2	-31.14078	2.09401	-14.87	0.000	-35.24554	-27.03601
female	36336.21	7350.688	4.94	0.000	21927.08	50745.34
education	5625.339	3932.303	1.43	0.153	-2082.926	13333.6
educ2	-898.978	388.4941	-2.31	0.021	-1660.52	-137.4357
educ3	46.26986	12.06217	3.84	0.000	22.62509	69.91464
feduc	-3228.15	540.5477	-5.97	0.000	-4287.754	-2168.546
race	-7353.703	7140.43	-1.03	0.303	-21350.67	6643.267
raceeduc	1270.519	526.402	2.41	0.016	238.6442	2302.395
marriage	15957.66	1618.142	9.86	0.000	12785.72	19129.61
femalemarr	-8407.958	7571.794	-1.11	0.267	-23250.5	6434.588
management	-16297.13	1637.701	-9.95	0.000	-19507.42	-13086.84
educjob	-20706.58	3510.938	-5.90	0.000	-27588.87	-13824.3

foodindustry	-13329.23	3146.156	-4.24	0.000	-19496.46	-7162.005
production	564.0181	2695.3	0.21	0.834	-4719.422	5847.459
sales	6407.94	2628.76	2.44	0.015	1254.936	11560.94
transportation	-1566.216	2366.963	-0.66	0.508	-6206.037	3073.604
_cons	-46292.49	15079.61	-3.07	0.002	-75852.16	-16732.81

Table 2 presents the regression result of all variables on income. The variables overall are statistically significant except for variables such as “production,” “transportation,” and “race.”

The variable “age” has a positive impact on income (+2,742), while age2 has a negative coefficient (-31). This variable has a quadratic relationship with income, and the function is:

$$2,742 \times \text{age} - 31 \times \text{age}^2 \tag{1}$$

As discussed above, the variable “age” is a continuous variable representing the individual’s actual age, and “age2” is age squared. The interpretation of this result is consistent with our intuition. In our early life stage, when age increases, our wages will also increase. However, as we get older and enter the later life and career stage, our wages will begin to decline. Therefore, this result shows the trend of wage change as a result of age increasing. Figure 1 shows this trend graphically; we could clearly see that the line begins by increasing, peaks around the age of 47, and then begins to decline over time.

The variable “female” has a positive coefficient. This means that a female, on average, earns 36,336 more dollars of annual income than men, holding other variables equal zero and constant.

For the variable “education,” a cubic function was used to represent its relationship with the annual income. The result of the regression analysis shows that the variable “education” has a positive coefficient, which indicates that education is correlated with the variable wage and it has a positive impact on wage. Besides, the variable “educ2” has a negative coefficient, and the variable “educ3” has a positive coefficient. The function represented by these regression results is:

$$5625 \times \text{education} - 898 \times \text{education}^2 + 46 \times \text{education}^3 \tag{2}$$

The graph of this equation is illustrated in figure 2, where education initially has a smaller impact on wages and only increases at a decreasing rate. On this figure, annual income represents the money earned by an individual in a given year in U.S. dollars. However, after nine years of education, the effect of education on wages begins to rise dramatically and grow at an increasing rate.

An interactive variable “feduc” was created by multiplying the variable education with the variable females to see the correlation between education and income under the context of gender. The new variable “feduc” has generated a negative coefficient in the regression analysis. Since it is statistically significant, it can be inferred that there exists a different effect of education on income for females and males. The interpretation of these results could be that for the same year of education, a female, on average, earns 3,228 dollars less than men.

The “race” variable also has a negative relationship with income. This means that for individuals who are white, their earnings, on average, will be 7,353 dollars less than non-white individuals, holding everything else constant. However, the interactive variable “raceeduc” shows the different effects of education on income based on different races, and the coefficient is positive 1,270. The interpretation of this variable is that for the same year of education, white individuals, on average, will earn 1,270 dollars more than non-white individuals. Therefore, after approximately six years of education, individuals who are white, on average, will make more income than non-white individuals, and the pay gap will increase for every additional year of education.

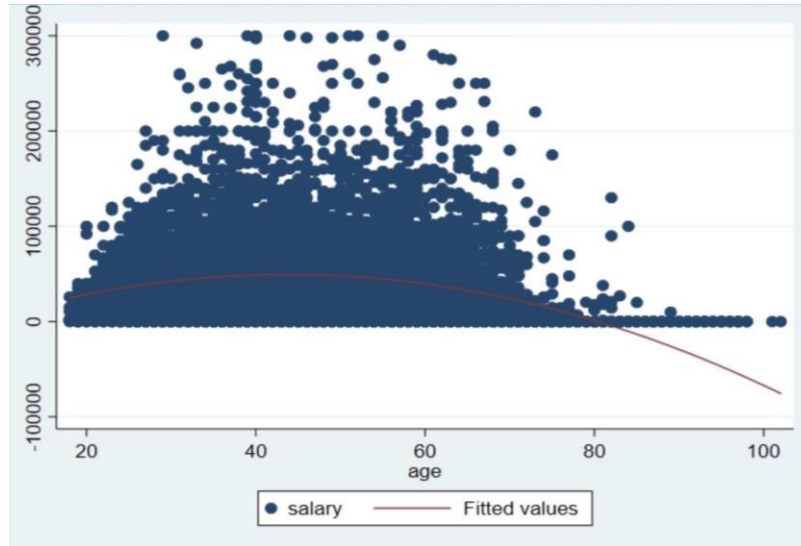


Figure 1. Trend of Wage Change

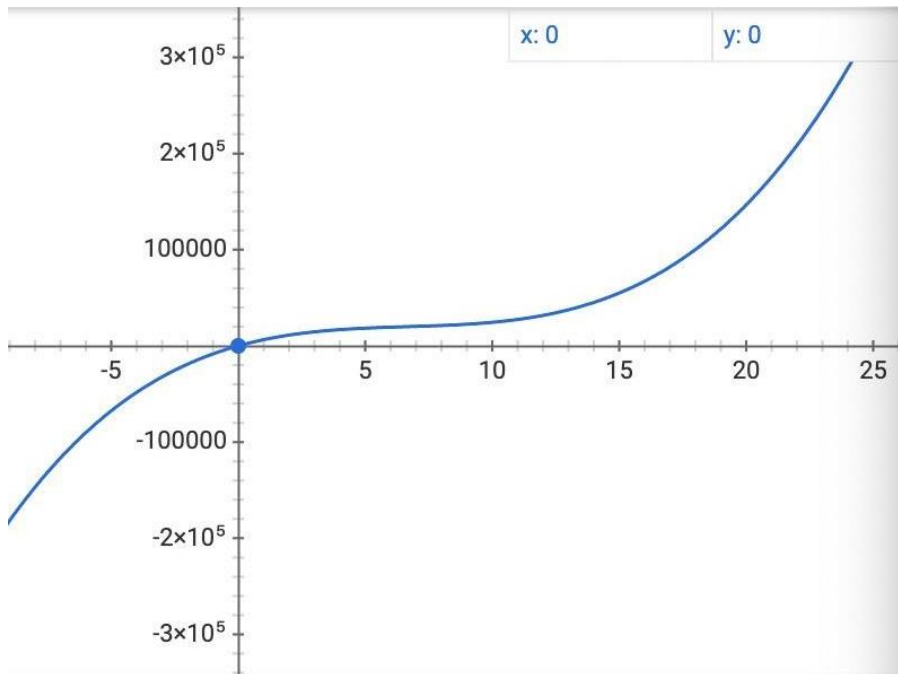


Figure 2. Relationship between Education and Annual Income

The coefficient of the variable “marriage” is positive, meaning that the individuals who are married, on average, tend to earn a higher income than those who are not married. However, the coefficient of the interactive variable “femalemarr” indicates that there is a different impact of marriage on the income for females and males. The coefficient of “femalemarr” is negative, which suggests that a female’s income, on average, is hurt more and typically earns less than men when getting married.

The rest of the variables are occupation variables. For the variable “management,” the coefficient is negative, meaning that people working at management positions, on average, earn 16,297 fewer dollars than the rest of the occupations. The same interpretation can also apply to

the variable “educjob.” The negative coefficient of educjob means that individuals who have an occupation related to education, on average, earns 20,706 fewer dollars than those at other industries. The variable “foodindustry” is typically the same; its negative coefficient implies that individuals who work in the food industry, on average, earn 13,329 fewer dollars than those who work at other industries or occupations. The variable “production” is positively related to income, meaning that individuals who work in the production industry, on average, earn 564 more dollars than those who work at other industries or occupations. However, since the “production” variable has a very high p-value, it might not have much explanatory power over income level. For the “sales” variable, its coefficient is positive,

showing that individuals who work as sales, on average, earn 6,407 more dollars than those who work in other industries. The “transportation” variable is negatively associated with income, meaning that individuals who work at transportation, on average, earn 1,566 fewer dollars than those who work in other industries. This variable also doesn’t have much explanatory power over income, given its high p-value.

4.DISCUSSION

Based on the analysis for the regression models, it has been shown that there exist wage differences between these personal factors. Age is an interesting factor that its impact increases at a relatively young age, peaks around the age of 47, and then begins to decline. This situation could be explained by the fact that when young people first enter into the labor market, they tend to gain more expertise and experience as they grow older. However, when people are at their later career stage or near their retirement age, they might spend more time with their family, doing exercise, and relaxing, rather than devoting themselves to work [7]. This regression result has been consistent with reality, which doesn’t seem to explain or contribute to the wage inequality of our society.

Education also proves to be one of the most important factors that can influence income. As people first go to school, the influence on wages is relatively small. However, as people have a higher degree of education, education tends to have a more significant impact on wages. Since the critical point for the function of education on income is around 9 to 10 years of education, this regression results could be explained by the fact that people, received less than ten years of education tend to learn more fundamental knowledge during the education process. They don’t tend to learn anything specific like a major in college. Therefore, when people with fewer years of education enter the labor market, their knowledge learned at school doesn’t seem to make a big difference on their wage and work performance, and years of education initially only increase at a decreasing rate. However, when people have a higher degree of education, such as undergraduate and graduate, their work performance would be more influenced by the knowledge they learned at universities or colleges [8]. Therefore, years of education tend to have a more significant impact on people’s income for people with more years of education and a higher degree of education.

Although with zero years of education, the regression model indicates that the females can earn a higher level of income than males, the wage differences between females and males at the same years of education still show that there is a potential discrimination against female workers in the labor market. The result of the regression model shows that for the same year of education, female workers also earn less than male workers by a significant amount. The wage gap between

female workers and male workers could lead to the wage inequality in our society today. The lower payment for female workers has proven that the discrimination against female workers still existed; such discrimination could lead to a larger wage gap between females and males as years of education increase [9]. The larger the wage gap existed, the more likely there existed wage inequality and discrimination in our society. So, the different effects of education on income among females and males could lead to potential wage inequality in society. Not only is wage different for females and males at the same level of education, but there also exist wage differences for races at the same level of education [10]. The result of the regression model have shown that for the same year of education, white people will earn a higher level of annual income than non-white people. This result can also prove the still existence of discrimination of race and gender in the labor market, and these discriminations deepened the wage inequality in our society today.

One of the variables in this research project that is not consistent with the previous research done by Siwei Cheng is marriage. The results of the regression models show that being married is correlated with a higher payment than those not married. One explanation might be that people getting married nowadays are more financially prepared, and people also tend to marry at a later age. This could result in the situation that people getting married tend to be older than those unmarried, thus accumulating more wealth and having a higher income level; Once people get married, they need to take more responsibility and take care of their family, so this might make them have financial capability before marriage. However, the interactive variable between females and marriage shows that marriage has affected females and males income differently. A female typically earns less income when married than when not married. This might be because of the social norm and stereotype, in which when females get married, the expected social role for them is to take care of family and children [11]. This social norm may influence many females to stay home after marriage or choose to allocate less time to work and more time for family issues. These factors may explain why marriage impacts females and males’ income differently and how females’ income is hurt more than males’ income when getting married.

Occupations also play an essential role in the determination of income. People in management occupations, education-related occupations, food industry, and transportation industry, on average, all earn less income than other occupations[12]. Since these variables, except the transportation industry, are all statistically significant, they all have very strong explanatory power on income, and thus people, in each of the industries earn less income than other occupations. These negative coefficient variables could show the poor prospects of the occupation or industry in the short term.

Since the production variable has a very high p-value and not much explanatory power of the model, the variable would bring minimal usefulness for the discussion. However, one job position has a positive relationship with income that is statistically significant: sales. People in sales positions tend to earn more than any other occupation, holding everything else in the model constant; this could provide meaning for the great prospect of this occupation.

Finally, the regression results provide meaningful information about the different factors influencing wages. The regression model shows how age and education influence wages in a non-linear way and then present more realistic insights about the society nowadays. It also suggests how education and marriage impact income differently for females and males, and white and non-white individuals; lastly, it provides valuable information about the prospects of each industry in the short-term and each industry's income level compared to the other industries.

5. CONCLUSION

In the last thirty years, the income inequality has increased dramatically in the United States among workers with high-skills and low-skills. Fewer and fewer people in control of the majority of the wealth has led to the increasingly unequal wage distribution among the population. Factors influencing people's annual income may provide an explanation for income inequality. Age, gender, race, education, and marital status all influence the income levels and the wage growth trajectory over a lifetime. Among these variables, the results have shown the existence of wage inequality between workers and wage discrimination against workers. There are maybe more variables relevant but not included in this research regression results, but the current results have shown the social issues of wage inequality and discrimination. This should remind us to continue encouraging and advocating wage equality for all individuals.

REFERENCES

- [1] Barany, Z. "The Minimum Wage and Inequality" *Journal of Labor Economics*, 34 (1) (2016), pp. 237-274.
<https://www.journals.uchicago.edu/doi/10.1086/682346>.
- [2] Cheng, Siwei. "A Life Course Trajectory Framework for Understanding the Intracohort Pattern of Wage Inequality." *The American journal of sociology* 120.3 (2014): 633–700. Web.
- [3] Kim, Jongsung. "Gender Difference in Employment and Income in China's Labor Market." *Institute for National Security Strategy. The Journal of East Asian Affairs*, Vol. 27, No. 2 (Fall/Winter 2013), pp. 31-53.
- [4] Devereux, Paul. "Effects of Industry Growth and Decline on Gender and Education Wage Gaps in the 1980s" *Sage Publications*. Vol. 58, No. 4 (Jul., 2005), pp. 552-570 Sage Publications, Inc.
- [5] Febriani, Febriani. "Quality Education and Skills of Indonesian Labor, Towards Equality Wages in Foreign Countries." *Jurnal ekonomi pembangunan (Surakarta, Indonesia)* 14.2 (2013): 203–213. Web.
- [6] Vu, Tien M, and Hiroyuki Yamada. "Decomposing Vietnamese Gender Equality in Terms of Wage Distribution." *Pacific economic review (Oxford, England)* 23.5 (2018): 705–731. Web.
- [7] Lazear, Edward P., and National Bureau of Economic Research. *Age, Experience and Wage Growth*. Cambridge, Mass: National Bureau of Economic Research, 1974. Print.
- [8] Martins, Pedro S, and Pedro T Pereira. "Does Education Reduce Wage Inequality?" *Labour economics* 11.3 (2004): 355–371. Web.
- [9] Huang, Tung-Chun. "The Impact of Education and Seniority on the Male-Female Wage Gap: Is More Education the Answer?" *International journal of manpower* 20.6 (1999): 361–374. Web.
- [10] Budig, Michelle J, Misun Lim, and Melissa J Hodges. "Racial and Gender Pay Disparities: The Role of Education." *Social science research* 98 (2021): 102580–102580. Web.
- [11] Hersch, Joni, and Leslie S Stratton. "Household Specialization and the Male Marriage Wage Premium." *Industrial & labor relations review* 54.1 (2000): 78–94. Web.
- [12] Helland, H, T Bol, and I nge. "Wage Inequality Within and Between Occupations." *Nordic journal of working life studies* 7.4 (2017): 3–27. Web.