Proceedings of ICEMGD 2025 Symposium: Resilient Business Strategies in Global Markets
DOI: 10.54254/2754-1169/2025.CAU26954

The Impact of Education on Income and Its Gender Heterogeneity

XINYUE CHEN

MACQUARIE University, Sydney, Australia
xinyue.chen@students.mq.edu.au

Abstract. The impact of education on personal income has always been an important topic in economics and sociology research, and gender differences add complexity to this relationship. This study aims to explore whether there is gender heterogeneity in the rate of return of education to income, and further analyze the social reasons and policy implications behind it. Methodologically, this study employed a large sample of observational data and a multivariate regression model to systematically analyze the impact of education level on wage income. Regression analysis and robustness tests were conducted separately for males and females. The model used income as the dependent variable and years of education as the core independent variable, and introduced control variables such as age, region, and marital status. By incorporating gender interaction terms, the model clearly identified differences in the returns to education between men and women. Empirical results show that education significantly increases individual wages, but this effect is more pronounced for women. In other words, for every additional year of education, women experience a larger increase in income than men. This finding aligns not only with the basic expectations of human capital theory but also with the "human capital compensation effect," which states that women's educational advantage stems in part from their ability to compensate for potential institutional and structural disadvantages in the labor market through higher education. Therefore, while education plays a key role in narrowing the gender income gap, sustained policy intervention and structural social reforms are also essential.

Keywords: Returns to Education, Human Capital Compensation Effect, Gender Wage Gap, Human Capital Compensation Effect

1. Introduction

With economic globalization, global attention to economic fairness and social mobility has continued to increase, and personal wage income has become a widely discussed topic among the public and in economics. Through a large amount of data statistics, it can be observed that there are many individual and group differences in wages, and the causes of these differences have always been what economists and the public want to know. According to many studies, there are numerous factors that affect personal income. A person's age, years of education, occupation, etc. will affect the money he can earn [1]. Among them, education, as the core element of human capital investment, is widely believed to have a deep connection with both income and income differences.

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/2025.CAU26954

Among all the differences, the most common one is gender gap. If the wage difference is caused by work ability or personal ability, people may understand its existence. However, if the individual income is different simply because of gender, people will think it is unfair [2]. Moreover, the difference of education level of labor will also affect gender gap in personal income. Based on this background, the impact of education level on income and the gender differences involved are very worthy of attention.

A large amount of literature has explored and studied the relationship between education and personal income, and research on gender and income differences has received increasing attention in labor economics. For instance, Säve-Söderbergh used large-scale field survey data from Sweden to study the salary negotiation behavior and starting salary differences of university graduates in their first employment [3]. The paper suggests that negotiation behavior alone explains a large portion of the gender gap in starting salaries, especially among high-income groups. This study provides new empirical evidence on the gender pay gap, but its analysis focuses on the "negotiation behavior" mechanism and does not delve into the gender heterogeneity between years of education and wage returns. Li et al. used the extended Mincer earnings equation model based on the latest survey data on China's labor market to study the relationship between the returns to education and the gender wage gap in different sectors (public and private) [4]. The contribution of this study is that it embeds the analytical framework of "education returns" and "gender wage gap" into the context of specific industries for discussion, revealing the sectoral heterogeneity of gender income inequality in China's economic structure. The contribution of this study is that it embeds the analytical framework of "education returns" and "gender wage gap" into the context of specific industries for discussion, revealing the sectoral heterogeneity of gender income inequality in China's economic structure. Widyanti used the extended Mincer wage equation and quantile regression model based on the Indonesian National Labor Force Survey data to reveal the heterogeneous relationship between the return on education and wage distribution, and emphasized the superior effect of educational investment on high-income groups. However, the study did not incorporate the gender dimension into the analytical framework and failed to explore whether there are significant differences in the returns to education between men and women [5].

There is a lack of systematic exploration of whether there is significant heterogeneity in the returns to education by gender. This study pays special attention to the impact path of education on income and its differences among different groups and aims to empirically test whether the impact of education on income is significantly heterogeneous between genders. This paper firstly analyzes existing literature to obtain basic theoretical support. Secondly, this study will use multiple regression to study the relationship between education and income. Thirdly, gender regression is performed to determine whether there is potential heterogeneity. Then, the interaction term between education and gender is added to determine its significance. Finally, it is concluded whether education has an impact on income and whether there are gender differences.

2. Method

2.1. Dataset preparation

This study used a salary dataset containing 6,704 respondents, which was obtained from Kaggle [6]. This dataset covers six key variables: respondent age, gender, education level, years of work experience, occupation, and annual salary. Gender is a categorical variable, divided into male, female, and other. Education level covers four levels: high school, undergraduate, master's, and

DOI: 10.54254/2754-1169/2025.CAU26954

doctoral. Income is a continuous variable, presented as salary. Therefore, this paper takes salary, gender and education level as the main target variables.

The original dataset contains 6 columns and 6, 704 rows, but this study preprocessed the dataset to accommodate regression analysis. This study does not intend to use the occupation column. The core of this article is the impact of education on income and gender heterogeneity. Including occupation in the regression may obscure the role of education and easily lead to over-control in the model. Before conducting this study, observations with missing values were removed. For example, some data points lacked salary value, while others had all data missing. A total of five observations were removed, leaving 6,699 study values.

This study then converted the two character variables, gender and education level, into numeric codes to serve as dummy variables. For education level, 0 represents a bachelor's degree, 1 represents a high school degree, 2 represents a master's degree, and 3 represents a doctorate. For gender, 1 represents male, 0 represents female, and 2 represents other. All data were described and summarized using tools such as Stata. The age and experience variables have a wide distribution, and the salary variable has a wide distribution, with a large gap between the minimum and maximum values. Logarithmic processing can be used to mitigate heteroskedasticity and skewness in the income distribution. The following Figure 1 illustrates the distribution characteristics of education and logarithmic income. The box plot shows that the median logarithmic wage of individuals increases with increasing educational levels (e.g., from bachelor's to master's to doctoral), indicating that individuals with higher education levels earn higher incomes. This finding provides a basis for conducting this study on whether education leads to higher wages and lays a foundation for visualizing gender heterogeneity in the returns to education.

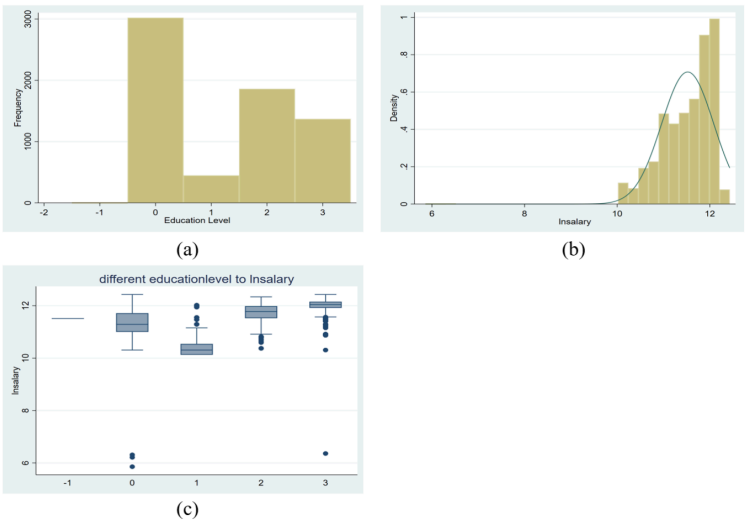

Figure 1. Relationship analysis diagrams: (a) educationlevel distribution (b) lnsalary distribution (c) different educationlevel to lnsalary (picture credit : original)

2.2. Multiple regression

This paper uses human capital theory as its research framework. Education is considered an investment that increases worker productivity and has economic returns [7]. Relied on data from multiple countries, they empirically found that individuals with more education generally have higher market income returns. Based on this theoretical research, this paper further explores whether gender affects the income returns brought by education from a gender perspective.

DOI: 10.54254/2754-1169/2025.CAU26954

The merit of this study lies in its focus on gender heterogeneity. Compared with traditional studies on the impact of education on income, the inclusion of the gender dimension enhances the depth and significance of the analysis. It provides a basis for solving structural problems of wage disparity, allocation of educational resources, and gender equality policies.

The core objective of this study is to investigate whether education affects individual income and whether there are significant gender differences. Firstly, examine the relationship between education and income through a global regression. Then conduct separate regressions for each sex, including interaction terms, to assess heterogeneity in the returns to education between men and women. It also used significance tests combined with t-tests to determine the presence of heterogeneity, and employ robustness tests such as heteroskedasticity.

The multiple regression functions are:

$$\ln(\text{Salary}_i) = \beta_0 + \beta_1 * \text{EducationLevel}_i + \beta_2 * \text{Age}_i + \beta_3 * \text{Experience}_i + \beta_4 * \text{Gender}_i + \varepsilon_i \quad (1)$$

Gender-specific OLS regression:

$$\ln\left(Salary_i^{female}\right) = \beta_0 + \beta_1 * EducationLevel_i + \beta_2 * Age_i + \beta_3 * Experience_i + \varepsilon_i \quad (2)$$

female, if gender = 0.

$$\ln\left(Salary_i^{male}\right) = \beta_0 + \beta_1 * EducationLevel_i + \beta_2 * Age_i + \beta_3 * Experience_i + \beta_4 * Male_i + \varepsilon_i \quad (3)$$

male, if gender=1.

In these regressions, $ln(Salary_i)$ is depended variable, and $Educationlevel_i$ is main explanatory variable. Age and Experience is the control variable. β_0 is the constant term. ε_i is the random error term.

Using interaction regression to test gender heterogeneity in education:

$$\ln(Salary_i) = \beta_0 + \beta_1 * EducationLevel_i + \beta_2 * Gender_i + \beta_3 * (Educationlevel_i * Male_i) + \beta_4 * Age_i + \beta_5 * Experience_i + \varepsilon_i \quad (4)$$

β_1 is the rate of return to education when gender is 0, that is, the benchmark group. β_3 is the additional return on education for men compared to women

3. Results and discussion

3.1. The multiple regression results

This study first established a multiple regression model with the logarithm of salary as the dependent variable, education level as the main independent variable, and age, work experience, and gender as control variables.

The regression results are shown in Table 1 and Table 2 below. The resulting coefficients indicate a significantly positive impact of education on income. This means that, ceteris paribus, a one-unit increase in education leads to an average wage increase of 4.81%. This result confirms the core conclusion of human capital theory: education, as an investment in human capital, can improve workers' productivity and thus enhance their competitiveness in the labor market [8]. Firms are

DOI: 10.54254/2754-1169/2025.CAU26954

willing to pay higher wages to more productive workers, so wages rise with higher education levels. The results of the first regression model of this study show that education is an important driving force for personal income growth, and it can be inferred that education promotes personal income and overall economic growth [9].

Table 1. Analysis of variance and model fit statistics

Source	SS	df	MS		
Model	1243.39456	4	310.848639	Number of obs	6,699
Residual	885.69	6694	0.132311923	F(4, 6694)	2349.36
Total	2129.09057	6698	0.317869599	Prob > F	0.000
				R-squared	0.5840
				Adj R-squared	0.5838
				Root MSE	0.36375

Table 2. OLS regression of log salary on education, gender, and controls

Variable	Coefficient	Standard error	T value	P value	95% CI Lower	95% CI Upper
EducationLevel	0.0480904	0.0043911	10.95	0.000	0.0394824	0.0566984
Age	-0.0234642	0.0016944	-13.85	0.000	-0.0267858	-0.0201426
YearsOfExperience	0.0914965	0.0021127	43.02	0.000	0.087327	0.095666
Gender	0.0831612	0.0089968	9.24	0.000	0.0655247	0.1007977
_cons	11.46511	0.0412501	277.94	0.000	11.38424	11.54597

Work experience also has a significant positive impact on income. This indicates that for every additional year of work experience, income generally increases by 9.15%. This suggests that the accumulation of work experience not only enhances proficiency and competence, but also potentially increases workers' remuneration as positions and responsibilities increase. However, the age variable showed a significant negative correlation, which may reflect the career life cycle effect: in the early stages of a career, income increases with experience, but in the middle and late stages of a career, income may decline due to factors such as pressure to update skills, changes in health status, or approaching retirement.

Meanwhile, this study found that the coefficient of the gender variable is positively significant. Because Gender = 0 represents female and Gender = 1 represents male, this result means that men's wages are, on average, about 8.32% higher than women's. Even when human capital factors such as education and experience are constant, the gender wage gap persists and is significant. This suggests that the gender difference is not entirely explained by observable productivity variables but may involve implicit discrimination in the labor market, differential promotion opportunities, and differences in job distribution (e.g., men are more concentrated in high-paying industries and management positions) [10]. This result provides a foundation for subsequent research.

DOI: 10.54254/2754-1169/2025.CAU26954

Table 3. Variance Inflation Factor (VIF)

Variable	VIF	1/VIF
Age	8.43	0.118632
YearsOfExperience	8.41	0.118999
EducationLevel	1.46	0.686838
Gender	1.03	0.970463
Mean VIF	4.83	

Table 4. Robust regression results

Variable	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]
EducationLevel	0.0480904	0.0047108	10.21	0.000	0.0388558, 0.057325
Age	-0.0234642	0.0017641	-13.30	0.000	-0.0269223, -0.020006
YearsOfExperience	0.0914965	0.0027584	33.17	0.000	0.0860893, 0.0969038
Gender	0.0831612	0.0094166	8.83	0.000	0.0647016, 0.1016207
_cons	11.46511	0.0412614	277.87	0.000	11.38422, 11.54599

Table 3 shows a multicollinearity test conducted on this regression model revealing an average VIF of 4.83. The VIF values for Age and Years of Experience were close to 8.4, indicating a strong correlation between the two, but not severe collinearity. The Breusch-Pagan heteroscedasticity test which shows in Table 4 indicated heteroscedasticity ($p < 0.01$), so robust standard errors were used to correct for this. The corrected results showed little change in the coefficients and significance levels of the variables, demonstrating the robustness of the research conclusions.

3.2. Gender-specific OLS regression

Focusing on the female sample, it can be observed that it from Table 5 that the coefficient for the education level variable is significant at the 1% level. This indicates that, after controlling for other factors, each level increase in women's education leads to an approximately 6.96% increase in wages. This result suggests that women have a higher rate of return on education, with higher education significantly boosting their income. This may reflect that women rely more heavily on the accumulation of human capital to boost their income in the labor market. The regression coefficient for work experience is significantly positive, indicating a higher return on experience for women than for men. The age variable, however, is significantly negative, consistent with men. This suggests that both men and women experience some level of wage suppression as they age, but the degree of this constraint is less pronounced for men than for women.

DOI: 10.54254/2754-1169/2025.CAU26954

Table 5. Gender-specific regression results

Variable	Female (Gender=0)Coef. (Robust SE)	Male (Gender=1)Coef. (Robust SE)
Education Level	0.0696*** (0.00644)	0.0259*** (0.00684)
Age	-0.0378*** (0.00344)	-0.0141*** (0.00191)
Years of Experience	0.1118*** (0.00520)	0.0793*** (0.00301)
_cons	11.7532*** (0.0812)	11.3607*** (0.0462)
N	3,013	3,672
R^2	0.580	0.593

Notes: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. All regressions are estimated with robust standard errors.

For the male sample, the variable for education level is significantly positive, indicating that for every one-level increase in men's education level, the logarithm of their wages increases by approximately 2.59%. Clearly, education level has a significantly greater positive impact on women's wages than men's, consistent with the fact that women rely more heavily on education to accumulate human capital and boost their incomes in the labor market. This difference suggests that women may face higher barriers to entry and advancement in the labor market, necessitating a higher level of education to demonstrate their abilities and, therefore, to enjoy higher relative returns from education [8]. This result also reflects a harsh reality: educated women are more likely to enter high-skilled, high-paying industries (such as finance, education, and healthcare), while men with lower education levels can enter some high-paying technical positions (such as engineering and IT). This explains why education's marginal contribution to men's wages is relatively low.

In summary, the wage returns of education for women are significantly higher than those for men. In most countries, the gender wage gap stems not only from differences in career choices and work experience, but also from differences in returns to education. This does not mean that women earn more than men overall, but rather that education plays a key role in narrowing the gender wage gap.

3.3. Gender-education interaction regression

This study included from Table 6 an interaction term between education level and gender in this regression to more precisely examine whether there are gender differences in the return to education on wages. The regression results show that the coefficient for this interaction term is -0.0546, indicating that after controlling for other influencing factors, the return to education for men is 5.46% lower than that for women. With each level of education increase, women's wages increase more than men's. The coefficient for males is positive, indicating that men's wages continue to be higher than women's. Although men generally have an advantage in wages, women's educational attainment increases faster and more significantly than men's, and education has a stronger impact on women's wages. This is reflected in the fact that highly educated women are able to overcome occupational segregation and enter higher-paying industries.

DOI: 10.54254/2754-1169/2025.CAU26954

Table 6. Regression results with gender-education interaction

Variable	Coef.	Std. Err.	t	P>|t|	95% Conf. Interval
EducationLevel	0.0793362	0.0060916	13.02	0.000	[0.0673948, 0.0912777]
male	0.1592372	0.0128376	12.40	0.000	[0.1340714, 0.1844030]
educ_male	-0.054565	0.0074784	-7.30	0.000	[-0.0692225, -0.0399049]
Age	-0.022061	0.0017038	-12.95	0.000	[-0.025401, -0.018721]
YearsofExperience	0.0903018	0.0021271	42.45	0.000	[0.086132, 0.0944716]
_cons	11.38336	0.0426366	266.99	0.000	[11.29978, 11.46694]

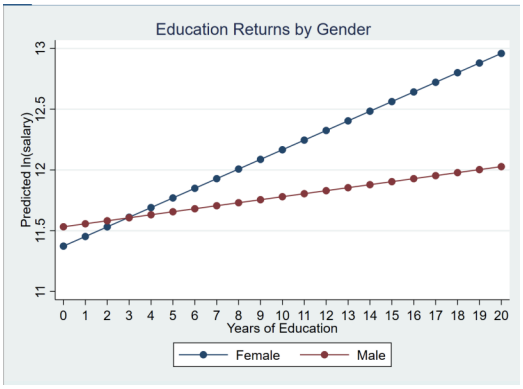

Figure 2. Education return by gender (picture credit: original)

Figure 2 also clearly shows that the female curve is steeper than the male curve, and the gap between the two curves widens with increasing years of education.

This paper therefore concludes that even after controlling for variables such as education, experience, and region, the gender wage gap persists, with men generally earning higher wages than women. Furthermore, the income returns from education differ somewhat across genders. This may reflect a "compensatory effect" for women's education in the labor market—women significantly narrow the wage gap with higher education levels, while men, even with higher education, experience relatively small marginal gains due to higher starting salaries. This also suggests that education plays an important role in narrowing the gender wage gap.

Table 7. Robustness test results

Variable	Coef.	Robust Std. Err.	t	P>|t|	95% Conf. Interval
EducationLevel	0.0793362	0.0060788	13.05	0.000	[0.0674199, 0.0912526]
male	0.1592372	0.0138220	11.52	0.000	[0.1321418, 0.1863326]
educ_male	-0.054565	0.0072080	-7.57	0.000	[-0.068695, -0.0404349]
Age	-0.022061	0.0017569	-12.56	0.000	[-0.025505, -0.018617]
YearsExperience	0.0903018	0.0023732	33.05	0.000	[0.0849462, 0.0956574]
_cons	11.38336	0.0429853	264.82	0.000	[11.2999, 11.46762]

Notes: The number of observations is 6,699, R-squared: 0.5885 and all coefficients significant at the 1% level.

Table 7 demonstrates the stability of the variables' impacts. The significance and direction of the coefficients for each variable are consistent with those in the original regression, indicating that the

DOI: 10.54254/2754-1169/2025.CAU26954

model results are not affected by heteroskedasticity and that the conclusions are robust. Even after adjusting for robust standard errors, education level still has a significant positive impact on log wages. The interaction term remains significantly negative, reinforcing the evidence for gender differential returns to education.

3.4. Suggestions

After all the analyses, this study concludes that the impact of education on income is gender-heterogeneous. Education has a significant positive impact on income, with a greater impact on women's wages than on men's. So this paper thinks measures should be taken to encourage education, narrow the gender income gap, increase average wages, and promote economic development. Governments and relevant authorities can strengthen wage transparency through legislation and regulation, ensure equal pay for equal work, and reduce the wage disadvantage of women with the same education level. This will reduce gender discrimination in the labor market. Society and the state should work together to increase social welfare and thereby increase investment in education. Develop programs for continuing education and professional skills development for women to ensure that women continue to receive the salary returns of education at different stages of their careers. Governments and businesses can provide more flexible work arrangements, childcare support, and maternity leave policies to reduce the negative impact of marriage and childbearing on women's careers, thereby fully realizing the economic returns of education investment.

4. Conclusion

Based on large-scale crowd-watching data, this study analyzes the impact of education on income and explores whether there are gender differences in this impact. This paper primarily utilizes three OLS regressions: standard regression, gender-specific regression, and interaction regression. The results show that education level significantly and positively affects wages, and that the returns to education for women are higher than those for men. This is consistent with the human capital compensation effect and some inferences from the glass ceiling theory. The research's contribution lies in incorporating gender interaction terms into the analytical framework, empirically revealing the differential impact of education on different gender groups and providing new evidence for the causes of the gender wage gap. The results suggest that while simply improving education levels can narrow some of the gaps, institutional and cultural barriers still constrain long-term increases in women's incomes. However, this paper also has certain limitations. First, the limited selection of variables fails to control for a wider range of potential variables, such as urban area and family background. Second, the limited cross-sectional data makes it impossible to analyze and observe the dynamic evolution of the returns to education over time. Future work could incorporate longitudinal tracking data to explore the dynamics of gender income over the career life cycle and introduce institutional variables at the organizational level to more fully explain the relationship between education and gender income differences.

References

[1] Bowles, S., Gintis, H. and Osborne, M. (2025) DSpace. [online] Handle.net. Available at: https://hdl.handle.net/20.500.14394/22402 [Accessed 15 Aug. 2025].

[2] Iwasaki, I. and Ma, X. (2020) Gender wage gap in China: a large meta-analysis. Journal for Labour Market Research, 54(1).

DOI: 10.54254/2754-1169/2025.CAU26954

[3] Säve-Söderbergh, J. (2019) Gender gaps in salary negotiations: Salary requests and starting salaries in the field. Journal of Economic Behavior & Organization, 161, 35–51.
[4] Li, M., Hu, X. and Jin, K. (2025) The Return on Education and the Gender Wage Gap in China: A Sector Perspective. SAGE Open, 15(2).
[5] Widyanti, R.D. (2018) Wage Inequality and Return to Education in Indonesia: Quantile Regression Analysis. Signifikan: Jurnal Ilmu Ekonomi, 7(1), 27–44.
[6] Mohithsairamreddy, M. (2023) Salary Data. [online] Kaggle. Available at: https: //www.kaggle.com/datasets/mohithsairamreddy/salary-data.
[7] Fleischhauer, K.-J. (2007) A Review of Human Capital Theory: Microeconomics. [online] SSRN Electronic Journal.
[8] Becker, G.S. (1962) Investment in Human Capital: A Theoretical Analysis. Journal of Political Economy, 70(5, Part 2), 9–49.
[9] Boser, U. (2011) Return on Educational Investment: A district-by-district evaluation of U.S. educational productivity. [online] Available at: https: //files.eric.ed.gov/fulltext/ED535853.pdf.
[10] Blau, F.D. and Kahn, L.M. (2017) The Gender Wage Gap: Extent, Trends, and Explanations. Journal of Economic Literature, 55(3), 789–865.