To What Extent Is Education Associated with Greater Household Net Worth for Males in Young Generation

Yiyang Jin [1, 2]

[1] Jinan University – University of Birmingham Joint Institute, Jinan University, Guangzhou, China

[2] School of Mathematics, University of Birmingham, Birmingham, United Kingdom

yxj364@student.bham.ac.uk

Abstract. Existing studies have mainly examined the impact of education level on income for the general population and lacked a detailed exploration of a single group. In order to fill the gap in the impact of educational attainment on income levels in the young male population and verify whether this impact was exhibited by self-esteem and living areas, this paper uses a dataset from the National Longitudinal Survey of Youth (NLSY) questionnaire and sets educational attainment(educ) as the independent variable, marital status, residential zone(urban) as the control variables as well as household wealth(net_worth1000) for dependent variable. At the same time, multiple linear regression, threshold regression, and multilevel regression are adopted in this study to reveal the plausibility and generalizability of the effect of educational attainment on household wealth in the young male population. Under 5% significant level, The regression results are as follows: In MLR, co = 11.3422,p-value = 0.000;In TR, co =6.7053, p-value= 0.000 for categories below threshold as well as co = 12.9303,p-value = 0.001 for groups above threshold, revealing impact on household wealth may be exhibited by self-esteem; In MR, coefficient is 11.481,ICC = 0.0068,implying that effect of education on wealth is not demonstrated through distinction of living areas. These outcomes demonstrate that the positive effect of educational attainment on household wealth shows a very high level of significance among young men, but the strength of the effect varies with fluctuations in the individual's level of self-esteem.

Keywords: Linear regression, social welfare, economic statistics, NLSY.

1. Introduction

Globally, education is widely seen as a critical means of improving the social economic status of individuals and household wealth-education relationship is gradually becoming a new method to measure this importance [1]. Especially in low- and middle-income countries, education is not only related to individual employment and income levels, but also largely influences the accumulation of household wealth. However, the relationship between education and wealth is not static and is influenced by multiple factors, including marital status, gender, level of regional development.

Many reasearches have been conducted to verify the effect of these factors on personal wealth. Marriage is a factor that cannot be ignored when exploring the impact of education level on individual wealth accumulation. Kapelle and Lersch based on German SOEP data, find that wealth growth is slower in the early years of a marriage, but that the level of wealth rises significantly with the duration of the marriage, and, more importantly, the study reveals that education level may indirectly, through educating homogenous marriages affects wealth accumulation [2]. This implies that the level of education not only directly affects wealth through income and occupation, but also may indirectly shape the wealth gap through the mechanism of marriage. Therefore, this paper takes marital status and its duration as important control variables when analyzing the impact of education on wealth. At the same time, different levels of development in the region also have an impact on wealth. A reasearch conducted in a U.S. Census Bureau study used multilevel regression with small region estimation (MRP) methodology to reveal significant regional differences in the impact of education on wealth [3]. They also point out that Median household wealth is significantly higher in states with generally higher levels of education, such as the Northeast and Pacific Coast, than in the South and Mountain states. This finding underscores the importance of using geography as a control variable

when modelling the education-wealth relationship to avoid using regional averages as a uniform rule. From a gender perspective, the level of education is associated with wealth income across gender groups. Under studies considering education as a prerequisite for improving financial literacy and thus wealth accumulation [4], Bannier and Schwarz, basing on a German sample, revealed that education does not uniformly amplify the wealth effect of financial literacy, i.e., highly educated females have an advantage over highly educated males in terms of their information-processing and asset-allocation abilities, which makes the marginal effect of 'literacy-wealth' significantly higher than that of equally educated males, suggesting that the path of education's effect on wealth may be significantly higher than that of the same level of education [5]. The marginal effect of 'literacy-wealth' is significantly higher than that of men with the same level of education, suggesting that the path of education on wealth may differ by gender.

However, the focus of most studies gathers on the qualitative analysis of education on wealth and the global tendency, with less research on aggregation of different methods as well as lacking in-depth research on the relationship between education and wealth in the youth group in terms of age stage. Therefore, most of the research variable settings that this paper draws from set the level of education as the core explanatory variable, the level of wealth as the response variable, and combined with control variables such as region, marital status, and intelligence factors. Based on the NLSY dataset, The paper employing a variety of regression methodologies to holistically reveals the ability of the different models to corroborate the relationship between men's wealth and income and the level of education.

2. Method

2.1. Data Preparation

The National Longitudinal Survey of Youth (NLSY) is a series of surveys sponsored by the U.S. Bureau of Labor Statistics (BLS), among others, designed to examine the entry of youth into the labour market and their transitions and exits within the market. The data used in this study comes from a dataset in NLSY called combined_1989.dta, which contains a total of 4,488 observed individuals including 2071 male observations. The dataset covers a total of 32 variables that reflect key information on the demographic characteristics, socio-economic status, and psychological characteristics of the respondents. These include continuous variables such as years of education, net wealth, and age; binary variables such as marital status and indicators of urban residence; and categorical variables such as individual self-esteem levels and police contact records. The dataset combines numerical and categorical data, providing valid values for multiple regression analyses of the impact of education level on individual wealth. The information of some variables is provided in Table 1.

Table 1. Explanation of Some Variables

Numble	Variables name	Explanation	Format
1	educ	Education years	Continuous variable
2	net_worth1000	Net wealth in thousands of USD	Continuous variable
3	age	age	Continuous variable
4	married	marital status	binary variable
5	urban	Indicator variable for urban residence	binary variable
6	sesteem_score1	Personal self-esteem level	categorical variable
7	policever	Any record of police contact	categorical variable

2.2. Data Processing

Prior to the empirical analyses, the raw data were preprocessed to ensure data consistency and quality. Firstly, all variables were checked for missing values, and the missing female observations for the net_worth1000 variable were dealt with by direct exclusion in this study. At the same time, binary variables such as urban and married were transformed by dummy variables (one-hot encoding). Categorical variables such as sesteem_score1 are standardised to simplify calculations. Next, multicollinearity diagnosis was also carried out prior to multivariate regression modelling to avoid the impact of redundant information on model stability. All data transformations were implemented using Python libraries such as pandas and statsmodels to ensure practicable and reviewable data handling.

2.3. Regression Models and Variants

2.3.1. Multiple Linear Regression

Initially, the study considered traditional method of analyzing the effect of coefficients on dependent variables. Here, this study considers net_worth1000 as implicit variable, educ as main independent variable and married urban as controlled variable. Further, constructing MLR model as follows:

$$net\_worth1000_i=\beta_0+\beta_1 educ_i+\beta_2 married_i+\beta_3 urban_i+\varepsilon_i \quad (1)$$

Where $\beta_0, \beta_1, \beta_2, \beta_3$represent the coeffecients of variables and ε is disturbance term.

2.3.2. Threshold Regression

Chatterjee in his study based on Rosenberg Self-esteem Scale and NLSY dataset states that self-esteem is positively correlated with increase in net worth, which indicates that strong self-esteem tends to promote wealth accumulation [6]. This suggests that the effect of education level on the extent of wealth accumulation may be perturbed by varying levels of self-esteem in younger age groups. Therefore, this design refers to the method of threshold regression of Vo's study on the effect of education level on wealth accumulation of Vietnamese households [7]. The study sets the threshold based on the self-esteem level of the sample, further performing regression separately. The specific regression equations are shown below:

$$net\_worth1000_i=\beta_0+\beta_1 educ_i+\beta_2 married_i+\beta_3 urban_i+\varepsilon_i \quad (2)$$

$$net\_worth1000_i=\gamma_0+\gamma_1 educ_i+\gamma_2 married_i+\gamma_3 urban_i+\delta_i \quad (3)$$

Where $\beta_0,\beta_1,\beta_2,\beta_3$ is the regression coefficient corresponding to when the level of self-esteem is less than the threshold, $\gamma_0,\gamma_1,\gamma_2,\gamma_3$ is the regression coefficient corresponding to when the level of self-esteem is less than the threshold.

2.3.3. Multilevel Regression

Chenevert pointed out in his study that there are significant regional differences in the impact of education on wealth, which suggests that the impact of education level on wealth may be constrained by the level of development in different regions [3]. In view of the randomness of the impact in different regions, this paper adopts a multilevel regression approach, using household wealth as the dependent variable, educational attainment as the independent variable, and taking into account the random effect of place of residence on top of the control variables, in order to measure the perturbations generated by the regional economic structure, house prices, and inequality of opportunity, among other factors. The regression model is expressed as follows:

$$net\_worth1000_{ij}=\beta_0+\beta_1 educ_{ij}+\beta_2 married_{ij}+\mu_j+\varepsilon_{ij} \quad (4)$$

Where j represents region number, i refers to the individual number,$\mu_j$is the Stochastic impact in region j.

3. Results and Discussions

3.1. Results of Multiple Linear Regression

Education level (educ) has a significant positive effect on wealth (net_worth1000) with a coefficient of 11.34 and the p-value is greatly less than 0.05 normal level, implying that controlling for urban and marital status (married), each additional year of education increases wealth by approximately 11.34 thousand dollars. The coefficient on the area variable (urban) is 23.90, indicating that individuals living in urban areas have significantly higher levels of wealth than those in non-urban areas. The coefficient on marriage is 36.94, indicating that married individuals have significantly higher levels of wealth than unmarried individuals. The overall model results support that education, urban residence and marital status have a significant positive effect on wealth. The value of coefficients are listed in Table 2.

Table 2. Results of Multiple Linear Regression

	coef	Std err	t	P >|*t*|	R-squared
Intercept	-117.4675	21.489	-5.446	0.000	0.040
urban	23.8971	9.614	2.486	0.013	
married	36.9395	8.338	4.430	0.000	
educ	11.3422	1.610	1.610	0.000	

3.2. Results of Threshold Regression

In the below-threshold self-esteem score group, the coefficient corresponding to the independent variable educ is 6.7053 and has a p-value of less than 0.01, which is highly significant; meanwhile, the coefficient of the control variable married is 32.9144 and has a p-value of approximately 0, which suggests that the level of education and marital status have a greater impact on net worth. However, the control variable urban in this subgroup exhibits a high p-value of about 0.202, presenting a non-significant effect of urbanisation. In the above-threshold self-esteem score group, the coefficient corresponding to the independent variable educ is 12.3903 and has a p-value of less than 0.01, demonstrating a stronger level of significance, while the coefficients of the control variables married and urban are 0.9815 and 0.2668 and have a smaller p-value, suggesting that the effect of education level, marital status and urbanisation on net worth is more significant. In addition to this, the methodology shows that the effect of education and marital status on net worth is significant in both low and high self-esteem score groups despite the different levels of self-esteem of the respondents, while self-esteem score itself fails to show a significant regression effect. However, when controlled variables are all fixed, the coefficient of educ is obviously different. This reveals that the effect of education level on net worth may be exhibited through self-esteem level, same education level young adults with higher self-esteem tend to have more accumulation in wealth. The values of coefficients are provided in Table 3.

Table 3. Results of Threshold Regression

	coef	Std err	t	P >|t|	R-squared
Below threshold					
Intercept	-78.6747	30.156	-2.583	0.01	0.030
urban	13.2436	10.370	1.277	0.202	
married	32.9144	9.455	3.481	0.001	
educ	6.7053	1.995	3.361	0.001	
Above threshold					
Intercept	-138.6079	48.131	-2.880	0.004	0.034
urban	32.1167	15.810	0.031	0.042	
married	35.8819	13.093	2.741	0.006	
educ	12.3903	2.563	4.835	0.000	

3.3. Results of Multilevel Regression

Multilevel Regression analyses revealed that years of education and marital status had a significant positive effect on net wealth. With the introduction of urban and rural (urban) as the secondary random intercept, the model estimates an urban-rural stratum variance of 482.15 and a residual variance of 34896.42, corresponding to an intra-group correlation coefficient (ICC) of 0.0068. This suggests that about 0.68% of the wealth difference can be attributed to urban-rural differences. The random effects results show that the effect of education level on the level of wealth remains significant and rural-urban differences ‘s effect on the level of wealth can be ignored.This implies effect of education on wealth is not demonstrated through distinction of urban and rural areas.The results are shown in Table 4.

Table 4. Results of Multilevel Regression

	coef	Std err	z	P >|z|	ICC
Intercept	-106.218	23.996	-4.427	0.000	0.0068
married	36.571	8.346	4.382	0.000	
educ	11.481	1.616	7.103	0.000	
Group Var	239.796	2.160			

In contrast, the results of the three models have a high degree of similarity. Comparing the threshold regression and multivariate linear regression, the gap between the explanatory ability of the dependent variable fluctuation of the two models is controlled within 0.01 and the coefficients of the educ variables are closer, which indicates the similarity and referability of the models' explanations of the real situation; comparing the multivariate linear regression and multilevel regression, the difference between the coefficients of the educ variables of the two models is small, which indicates that under the premise of the control variables being fixed, the two models of the independent variables can all have an approximate effect on the dependent variable.

4. Conclusion

Overall, all three models point out that the level of education affects wealth income at a high level of significance (the figure is 11.3422 for multiple linear regression;6.7053 for groups below threshold and 12.9303 for categories above threshold;11.481 for multilevel regression), with the p-value of coefficients are all approximately tantamount to zero. This outcome fits with the results of many papers and further confirms that higher education tends to lead to more wealth accumulation in the young male cohort, but the strength of the effect varies with fluctuations in the individual's level of self-esteem. However, this study still has areas that can be improved in the process of model construction. This paper adopts the median method in determining the critical value of the threshold regression, and different determination methods will have a greater impact on the data grouping, and future research can further consider the use of parameter estimation or weighted quartiles to improve the accuracy of the selection of the critical value. In addition, the dataset still has a large potential for exploration, and future studies can start from the remaining variables to further reveal the influencing factors of wealth accumulation among the youth group and continuously fill in the missing fields.

References

[1] Pfeffer FT. Growing wealth gaps in education. Demography. 2018 Jun 1;55 (3):1033-68.

[2] Kapelle N, Lersch PM. The Accumulation of Wealth in Marriage: Over-Time Change and Within-Couple Inequalities. Eur Sociol Rev. 2020;36 (4):580-93.

[3] Chenevert R, Gottschalck A, Klee M, Zhang X. Where the Wealth Is: The Geographic Distribution of Wealth in the United States. U.S. Census Bureau; 2016.

[4] Behrman JR, Mitchell OS, Soo CK, Bravo D. How Financial Literacy Affects Household Wealth Accumulation. Am Econ Rev. 2012;102 (3):300-4.

[5] Bannier CE, Schwarz M. Gender-and education-related effects of financial literacy and confidence on financial wealth. J Econ Psychol. 2018 Aug 1; 67:66-86.

[6] Chatterjee S, Finke M, Harness N. Individual wealth management: Does self-esteem matter? J Appl Bus Econ. 2009 Nov 1;10 (2).

[7] Vo DH, Ho CM. Does Educational Attainment and Gender Inequalities Affect Wealth Accumulation? Evidence from Vietnam. Heliyon. 2022;8: e12502.