The impact of education on income for the Northeast, Midwest, West, and South throughout high school and college

Income and Education: A Cross Regional Look into the Impact of Education on Income

Tom Barber[1]
✉ tbarber@umass.edu

[1] *University of Massachusetts Amherst*

Introduction

Income within the U.S. can vary by many different factors. In this research, we look specifically at age, sex, education level attained, and region of residence. The results show some clear conclusions: - Education has quite a large impact on total personal income, as most would expect. - There are regional differences within the total income variable, showing that region of residence does have an impact on income. - When looking at age and sex we see both have, to a lesser extent than education, an impact on income. Those that are older tend to make more, while women in general tend to make less.

Possible Omissions or Issues

This research takes into account quite a few variables that are very important in the income equation. There might be, however, some omissions of which might create an issue of endogeneity. The main concern here is the lack of information on socioeconomic status. This can impact income, education, and regional residence. Unfortunately, this could lead to omitted variable bias and endogeneity. Being unable to find proper variables that would allow for this to be taken into account we will have to be wary of the conclusions found from this research.

Methods

The Linear Regression Models equation:

Model 1

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i + \epsilon_i$$

Where β_0 = income of the Northeast when education is zero, β_1 = coefficient of education level, β_2 = coefficient of regional impact

Model 2

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i + \beta_3 x_i + \beta_4 x_i + \epsilon_i$$

Where β_0 = income of Northeast when other variables are zero and sex is male, β_1 = coefficient of education level, β_2 = coefficient of Regional impact, β_3 = coefficient of Age, β_4 = coefficient of Sex

Model 3

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i + \beta_3 x_i + \beta_4 x_i + \epsilon_i$$

Where β_0 = income of the Northeast when education is zero, β_1 = coefficient of education level, β_2 = coefficient of Midwest, β_3 = coefficient of South, β_4 = coefficient of West
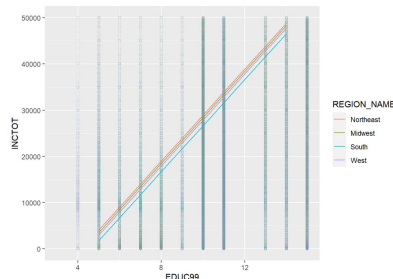
Results

Figure 1: Education vs Total Income controlled by Region
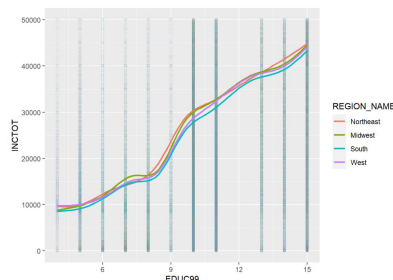

Figure 2: Education vs Total Income controlled by Region, Age, and Sex

Table 1: Linear Model of Total Income vs Education controlling for Region

term	estimate	std.error	statistic	pvalue
(Intercept)	-20732.18679	183.210140	-113.16070	0
EDUC99	4979.32726	11.923483	417.60676	0
REGION	-46.00095	3.735554	-12.34761	0

Table 2: Linear Model of Total Income vs Education controlling for Individual Regions

term	estimate	std.error	statistic	pvalue
EDUC99	4974.833	11.92313	417.2765	0
REGION_NAMENortheast	-20783.033	170.37566	-121.9836	0
REGION_NAMEMidwest	-21303.446	163.82381	-130.0388	0
REGION_NAMESouth	-23076.137	148.91720	-154.9387	0
REGION_NAMEWest	-21861.670	153.16321	-142.7531	0

Table 3: Linear Model of Total Income vs Education controlling for Region, Age, and Sex

term	estimate	std.error	statistic	pvalue
(Intercept)	-5162.0775	208.034152	-24.813606	0.0000000
EDUC99	4915.1210	11.548320	425.613523	0.0000000
REGION_NAMEMidwest	-470.8356	126.223117	-3.730186	0.0001914
REGION_NAMESouth	-2141.1104	110.988783	-19.291233	0.0000000
REGION_NAMEWest	-987.5186	115.814067	-8.525285	0.0000000
AGE	228.2001	2.002151	113.977476	0.0000000
SEX	-16542.9175	74.138791	-223.136437	0.0000000

Discussion

When looking at the data we can clearly see some disparities. There is a strong positive relationship between income and education, as most would expect. Even when controlling for different variables impacts we see the coefficient for education maintaining close to 5100 units. This implies that for each unit increase in education we expect an increase of about the same amount in total income. To a lesser extent, we also find that after controlling for age and sex the positive relation between income and education remains significant. These coefficients for age and sex show that on average the older you are the more you make, and that men typically make more than men. From the regional standpoint we observe negative coefficients. This implies that individuals within these regions make less than those in the reference region, which in this case is the Northeast.

When interpreting the coefficients of the other regions in terms of the reference region we see how each compares to the Northeast. The intercept for the Northeast is -20783, and represents the estimated total income for the region when education is theoretically zero. When subtracting this from the intercept for the Midwest we find a value around -520. Statistically, this means that on average individuals in the Midwest have an estimated total income that is 520 dollars lower than those in the Northeast. By similar processes, we observe that income in the West is 1078 dollars lower than in the Northeast. Finally, in the South we find, on average, the estimated total income is 2293 dollars lower than in the Northeast. These conclusions provide insight into income disparities regionally and are statistically significant, as very small p-values and numerically significant confidence intervals are obtained.

It is not very surprising that education has such a large impact on income. It is, however, interesting that income varies quite substantially by region. This is something that can be addressed by public policies. We see that the region that is particularly lacking is the South, which could be due to many reasons. An article published by Forbes discussed some of the best colleges in the U.S. and their location. These colleges were predominantly in the Northeast or West. One could then say that investment in education could be useful in lessening the income gap in the South. It could also be an issue of affordability and access to education, whether that be college or even high school. This can also be changed through public policy and economic or developmental programs.

References

- Data set was pulled from IPUMS, https://www.ipums.org/
- Further reading on regional differences in higher education: https://www.forbes.com/sites/ccap/2016/07/06/regional-differences-in-higher-education-where-is-the-best-place-to-go-to-college/?sh=3db14b848d87