Labour Market Analysis
Seattle University
ECON 5300 - 02 19WQ Applied Econometrics

Ankita Pathak, Nancy Jain and Sanyukta Ghai

02/27/2019

Contents

In the state of Illinois, how do earnings vary by educational attainment? Does the premium for higher education vary by gender?

1. Introduction

The Labor Market Analysis is particularly important in understanding the factors that affect the earnings among the people. In general, it is expected that higher education increases earnings. It is also important to understand how the earnings differ for different education levels to analyze if the returns justify the investment in higher education. Therefore, in this paper, we want to address - In the state of Illinois, how do earnings vary by educational attainment. The investment in time and money for higher education is the same for males and females across the US. However, based on a report published by the U.S. Bureau of Labor Statistics which presents earnings data from the Current Population Survey(CPS), indicates that with higher educational attainment, an individual is likely to earn more and female earnings has increased significantly from 1979 to present, but the differences in gender earnings with higher educational attainment still exists (Chart 4, Table 14, Women in Higher Education, 26(1). (2017). Therefore, the other critical question we want to address is - does the premium for higher education vary by gender.

We are using individual-level data from the 2016 American Community Survey (ACS) for people living in the state of Illinois. We are analyzing the effect of higher education and gender on a person's earnings. One more factor which is prevalent in the society is dissimilarities among individual's earnings with respect to the race; therefore, this report also emphasizes on finding out how racial differences affect earnings. Only 17.4% of our sample belongs to races other than white. Therefore, we have combined all the other races as non-whites to understand the gap between white and other races. We are using the multiple regression models to quantify the differences in earnings across genders, age, educational levels, and race and understand the weight of these factors in determining a person's earnings.

These questions are very important as historically, we have seen earnings disparity across genders and races. Even after so many years of reform, the effects continue to exist in the labor market. Analyzing the current state will give economists an idea about the effects of these reforms and will help them devise strategies for labor market earnings equity.

Our findings are in accordance with our expectations. Higher education leads to higher earnings. Earnings for Doctorate degree are more than all other levels except Professional degree. This may be because the earnings in research areas are less and more varied than professional jobs. Our analysis also shows that the gap in earnings for genders continues to exist. Also, the gap maintains through all educational levels. The gap between the earnings of white and non-whites also continues to exist.

2. Econometric models and Estimation Methods

We have used multiple regression model (logarithmic regression) to address the research questions. It describes the impact of gender and educational attainment on earnings. Key explanatory variables are education levels and gender. Different levels of education considered are High School Dropouts, High School Degree, College dropout, Associate degree, Bachelor's degree, Master's Degree, Professional Degree and Doctorate Degree. The categories that we have considered for gender are male and female. Also, experience is an important factor that affects the outcome. Since, experience is not given in our data, we have considered age and (age)ˆ2 as the proxy variable for experience. Races that we have considered for our analysis are whites and non-whites, these are dummy variables that emphasize the impact of variation in earnings and labor market by race and ethnicity.

We have also included interaction terms between gender and education levels. Interaction terms signify whether the difference in earnings across gender and education is significant or not. However, separate gender equations are easier to analyze than the combined equation. If the interaction terms are statistically significant, it implies that the coefficient on educational levels will be statistically different for male and

female equation. Without interaction terms, the coefficient of gender would be interpreted as the unique effect of gender on earnings. The non-zero coefficients on interaction terms between gender and education level shows the effect of gender on earnings. So, the effect of gender on earnings is not limited to its coefficient but also depends on the coefficient of interaction term and education level. Interaction between the following terms are included in the model: Female and Whites, Female and High School Degree, Female and College dropout, Female and Associate Degree, Female and Bachelor's degree, Female and Master's degree, Female and Professional degree, Female and Doctorate Degree.

The response variable that we have considered is natural log of total earnings of a person. We have considered the logarithmic regression model for our analysis, because it is more appropriate to correct the skewness seen in earnings distribution data. Also, for this analysis, the percentage differences across groups are more relevant than unit differences. For example, change in education level or change in gender is associated with a percentage increase or decrease rather than dollar amount of change in earnings.

We have used Ordinary Least Squares method of estimation.

3. Data

The data used in this paper is from the the 2016 American Community Survey (ACS). The data includes most population and housing characteristics. A set of core questions is asked every year, including questions on household composition, labor market and occupational dynamics, and earnings. We have considered individual level data for this analysis.

3.1 Sample Data

Initial data contains 1,26334 observations and 284 variables for the state of Illinois. We have filtered this data on the basis of following criteria: people who are currently employed, as most of the earnings population are employed, people who have positive earnings, as most of the earnings population have earnings greater than zero and people who have educational attainment of Grade 10 and higher, as most of the earnings population have completed Grade 10. Key variables of our sample data include total earnings, gender, educational attainment, age and race. We have removed all incomplete data from the dataset, as those records are meaningless for this analysis. This leaves us with 58,539 observations and 5 key variables.

3.2 Sample Descriptives

Table 1 provides descriptive statistics for our sample data used in the analysis. 51% of the sample is male and 49% is female. 83% of the sample is white and 17% is non-white. 49% of people in our sample have Bachelor's or higher level of education. Only 3% are highschool dropouts. From our sample, highest percentage of people have education level as High School Degree which is 24.8%, where as lowest pecentage have have Doctorate degree which is only 1.5%. The average age of our sample data is 43.14 years. Maximum total earnings of a person is $807,000 and the mean earnings is $54,908 with age ranging from 16 to 94.

4. Results

Results of the labor market analysis are presented in Table 3. Table 2 presents the OLS estimates of logarithmic model. Value of r-squared from log model, given in Table 2, is 35.15%, which means 35.15% of variation in log(total earning) can be explained by our model. Table 2 shows that controlling for age, race, education levels - earnings for men are more than earnings for females. Controlling for other variables such as age, gender, and race - earnings increase with increase in education level compared to the reference category of high school dropouts.

Earnings also vary by age while controlling for gender, race and education levels - earnings increase at decreasing rate upto 47.06 years and decrease at increasing age after that. Age also captures the effect of experience, since people with more experience are generally older than those with less experience. Table 2 also shows that white males and female earn more than their counterparts, while controlling for other variables - age and education levels. Table 3 presents the robust standard errors of all the parameter estimates.

4.1 Male vs Female

From Table 2 we can infer that, after controlling for age, race and education levels - on average, females earn 20.58% less than males, that is, there is penalty for females when compared with males. Controlling for age, gender and education levels - on average, whites earn 24.63% more than non-whites. Controlling for age and education level - on average, white males earn 25% more than non-whites males and white females earn 2.42% more than non-whites females. Also, controlling for age and education level - on average white females earn 34.74% less than white males, and non-white females earn 20% less than non-white males.

4.2 Education levels

By comparing parameter estimates in the models of males and females, we can infer that for higher level of education, earnings of females vary more than the earnings of males. Reference category for education level is high school dropouts. On average, after controlling for age and race, a person having high school degree earns 73.76 % more than a high school dropout and professional degree earns 580% more than a high school dropout.Controlling for age and race, on average, females earn 30.12% less than males with professional degree. Controlling for age and race, on average, females earn 29.69% less than males with professional degree and females earn 17.09% less than males with Doctorate Degree.

Parameter estimates of all the interaction terms between gender and education levels are not statistically significant at 5% level of significance, whereas parameter estimates of education levels, genders, age, (age)ˆ2, race and interaction term between gender and race are statistically significant at 1% level of significance.

5. Conclusions

The analysis in this paper shows that inequality in earnings exists across gender and race. The effect remains statistically significant for gender, even after controlling for age, race and education level. The effect, also remains statistically significant for race, even after controlling for age, gender and education level. Also, as level of education increases average earnings also increases, this pattern is seen across all levels of education except for Doctorate degree. We think the reason for that maybe, the years of research required for doctorate study replaces the years that could account as experience. Average earnings for Doctorate degree is less than that of Professional Degree, even though Doctorate degree is the highest education level for our data. This effect is statistically significant after controlling for age, gender and race.

Also the interaction between gender and education level is not statistically significant after controlling for age and race. It means that the effect of gender on earnings or effect of education level on earnings is not affected much by the interaction between gender and education level. It is interesting to observe that the interaction between gender and race is statistically significant after controlling for age and level of education, which means that the effect of gender on earnings or effect of race on earnings is affected by the interaction between gender and race.

The fact that there is a large earnings gap between males and females, and, whites and non-whites, implies that there is gender and racial discrimination in the state of Illinois. One possible explanation for the earnings gap between males and females is because females are expected to take sabbaticals or leave their jobs to take care of their family or other commitments. Possible explanation for the earnings gap between whites and

non-whites could be historical bias. We also, do not have the information about the skill levels across this set of people. Hence, we cannot generalize these observations for other states as well.

The conclusions above are subject to a number of limitations. First, we have not considered the interaction between race and educational levels. The race is expected to have some effect on the educational attainment because of the social structure and wealth. If we compare the earnings of both white and non-white people, while controlling for all the other external factors, for example, earnings of a white person having professional degree when compared with that of a non-white person having the same level of education might not vary that much. Second, we have considered age as a proxy for experience since information on work experience is not available directly. However, in order to estimate the pure effects of work experience, age must be controlled for, and vice versa. Therefore, it does not completely justify the effects of age and experience on earnings. Third, we have categorized race into whites and non whites only. Non whites could be separated into more categories to get a better insight about variation in earnings. for example: let's say, on average Asians earn more than African American, hence if we separate non whites further, we can provide better interpretation about the impact of race on earnings. Fourth, since variation in error terms for our model is not constant, it continues to show heteroskedastic. Hence, we reject the null hypothesis. Therefore, the results show heteroskedasticity. Finally, there may be other variables that affect earnings, e.g. skills, differences across industries, marital status or dependents on the person who is earning. Including these in the regression might increase the precision of our estimates as well as eliminate potential omitted variable bias.

I. References

Economic Importance and Statistical Significance: Guidelines for Communicating Empirical Research(n.d.). http://bloustein.rutgers.edu/wp-content/uploads/2014/10/Miller-and-Rodgers-Fem-Ec.pdf

Patten, E., & Patten, E. (2016, July 01). Racial, gender wage gaps persist in U.S. despite some progress. http://www.pewresearch.org/fact-tank/2016/07/01/racial-gender-wage-gaps-persist-in-u-s-despite-some-progress/

Women in Higher Education, 26(1). (2017). doi:10.1002/whe.2017.26.issue-1

Wooldridge, Jeffrey M., 1960-. (2012). Introductory econometrics : a modern approach. Mason, Ohio :South-Western Cengage Learning

II. Appendix

Table 1: Descriptive Statistics

Statistic	Mean	St. Dev.	Min	Max
Total Person Earnings USD	54,472.280	66,696.800	1	807,000
Age (in years)	43.124	14.528	16	94
Female	0.488	0.500	0	1
Male	0.512	0.500	0	1
White	0.826	0.379	0	1
Non-White	0.174	0.379	0	1
High School Dropout	0.030	0.172	0	1
High School Degree	0.247	0.431	0	1
College Dropout	0.235	0.424	0	1
Bachelor Degree	0.243	0.429	0	1
Associate Degree	0.096	0.294	0	1
Master Degree	0.107	0.309	0	1
Professional Degree	0.026	0.159	0	1
Doctorate Degree	0.015	0.122	0	1

Total number of records 58,539

Contributions

We have collectively worked through all the parts of this assignment in the following manner:

Sanyukta : Responsible for handling the coding and generating required tables for the report. Ankita : Responsible for ensuring that the requirements and output are documented correctly. Nancy : Responsible for interpretaion and analysis of data.

Table 2: Results for Logarithmic Models - Earnings, Gender, Education, Age, Race

	Dependent variable:		
	Log(Total Person Earnings USD)		
	Males	Females	Combined
	(1)	(2)	(3)
Age (in years)	0.180***	0.153***	0.167***
	(0.002)	(0.002)	(0.002)
Age x Age	−0.002***	−0.002***	−0.002***
	(0.00002)	(0.00003)	(0.00002)
Female			−0.230***
			(0.047)
White	0.214***	0.027*	0.220***
	(0.014)	(0.015)	(0.015)
Female x White			−0.196***
			(0.021)
High School Degree	0.511***	0.560***	0.553***
	(0.032)	(0.037)	(0.032)
College Dropout	0.674***	0.711***	0.710***
	(0.032)	(0.036)	(0.032)
Associate Degree	0.819***	0.878***	0.865***
	(0.035)	(0.039)	(0.036)
Bachelor Degree	1.205***	1.252***	1.248***
	(0.032)	(0.037)	(0.032)
Master Degree	1.344***	1.420***	1.398***
	(0.035)	(0.039)	(0.035)
Professional Degree	1.860***	1.862***	1.917***
	(0.043)	(0.052)	(0.044)
Doctorate Degree	1.501***	1.672***	1.561***
	(0.052)	(0.059)	(0.052)
Female x High School Degree			−0.055
			(0.047)
Female x College Dropout			−0.054
			(0.047)
Female x Associate Degree			−0.053
			(0.051)
Female x Bachelor Degree			−0.057
			(0.047)
Female x Master Degree			−0.049
			(0.051)
Female x Professional Degree			−0.128*
			(0.066)
Female x Doctorate Degree			0.038
			(0.078)
Intercept	5.554***	5.852***	5.822***
	(0.050)	(0.054)	(0.043)
Observations	29,913	28,544	58,457
R^2	0.361	0.299	0.352
Adjusted R^2	0.361	0.298	0.351
Residual Std. Error	0.919 (df = 29902)	0.956 (df = 28533)	0.938 (df = 58437)

Note: *p<0.1; **p<0.05; ***p<0.01
Reference Category for Males is a non-white male high school dropout
Reference Category for Females is a non-white female high school dropout
Reference Category for Combined is a non-white male high school dropout

Table 3: Results for Linear Models - Earnings, Gender, Education, Age, Race

	Dependent variable:		
	Total Person Earnings USD		
	Males	Females	Combined
	(1)	(2)	(3)
Age (in years)	6,025.159***	3,616.600***	4,842.154***
	(158.932)	(106.416)	(97.229)
Age x Age	−58.912***	−36.247***	−47.840***
	(1.766)	(1.200)	(1.088)
Female			5,821.824**
			(2,916.179)
White	13,302.590***	631.387	13,798.850***
	(1,074.563)	(664.848)	(904.388)
Female x White			−13,389.080***
			(1,262.186)
High School Degree	−2,402.897	−2,238.207	1,242.930
	(2,352.693)	(1,664.209)	(1,970.451)
College Dropout	7,811.286***	4,362.365***	11,063.610***
	(2,369.625)	(1,647.179)	(1,987.340)
Associate Degree	10,878.670***	6,501.151***	14,936.100***
	(2,619.618)	(1,764.566)	(2,193.214)
Bachelor Degree	42,957.070***	26,044.960***	46,828.400***
	(2,381.089)	(1,656.302)	(1,991.699)
Master Degree	62,759.910***	35,958.320***	67,555.590***
	(2,600.860)	(1,750.715)	(2,171.974)
Professional Degree	145,490.500***	84,774.740***	150,523.700***
	(3,209.547)	(2,367.442)	(2,688.389)
Doctorate Degree	79,737.300***	51,839.440***	84,960.190***
	(3,847.108)	(2,688.762)	(3,227.966)
Female x High School Degree			−9,017.736***
			(2,916.282)
Female x College Dropout			−11,510.160***
			(2,920.010)
Female x Associate Degree			−14,316.920***
			(3,157.317)
Female x Bachelor Degree			−26,149.330***
			(2,915.660)
Female x Master Degree			−37,915.150***
			(3,116.964)
Female x Professional Degree			−72,206.570***
			(4,089.443)
Female x Doctorate Degree			−39,654.970***
			(4,772.434)
Intercept	−106,795.800***	−53,665.710***	−82,784.710***
	(3,748.648)	(2,442.879)	(2,659.643)
Observations	29,913	28,544	58,457
R^2	0.243	0.199	0.252
Adjusted R^2	0.242	0.199	0.252
Residual Std. Error	68,424.050 (df = 29902)	43,374.240 (df = 28533)	57,698.950 (df = 58437)

Note: *p<0.1; **p<0.05; ***p<0.01
Reference Category for Males is a non-white male high school dropout
Reference Category for Females is a non-white female high school dropout
Reference Category for Combined is a non-white male high school dropout

Table 4: Logarithmic Models with Robust Standard Errors

	Dependent variable:		
	Log(Total Person Earnings USD)		
	Males	Females	Combined
	(1)	(2)	(3)
Age (in years)	0.180***	0.153***	0.167***
	(0.003)	(0.003)	(0.002)
Age x Age	−0.002***	−0.002***	−0.002***
	(0.00003)	(0.00003)	(0.00002)
Female			−0.230***
			(0.057)
White	0.214***	0.027*	0.220***
	(0.015)	(0.015)	(0.015)
Female x White			−0.196***
			(0.021)
High School Degree	0.511***	0.560***	0.553***
	(0.041)	(0.041)	(0.040)
College Dropout	0.674***	0.711***	0.710***
	(0.041)	(0.041)	(0.041)
Associate Degree	0.819***	0.878***	0.865***
	(0.042)	(0.042)	(0.043)
Bachelor Degree	1.205***	1.252***	1.248***
	(0.041)	(0.041)	(0.041)
Master Degree	1.344***	1.420***	1.398***
	(0.042)	(0.042)	(0.043)
Professional Degree	1.860***	1.862***	1.917***
	(0.058)	(0.058)	(0.052)
Doctorate Degree	1.501***	1.672***	1.561***
	(0.060)	(0.060)	(0.058)
Female x High School Degree			−0.055
			(0.057)
Female x College Dropout			−0.054
			(0.057)
Female x Associate Degree			−0.053
			(0.059)
Female x Bachelor Degree			−0.057
			(0.057)
Female x Master Degree			−0.049
			(0.059)
Female x Professional Degree			−0.128*
			(0.077)
Female x Doctorate Degree			0.038
			(0.082)
Intercept	5.554***	5.852***	5.822***
	(0.064)	(0.064)	(0.056)

Note: *p<0.1; **p<0.05; ***p<0.01
Reference Category for Males is a non-white male high school dropout
Reference Category for Females is a non-white female high school dropout
Reference Category for Combined is a non-white male high school dropout
Values in () are the Robust Standard Errors