Máté Schieszler
Péter Szilvási

Data Analysis 2 - Assignment 1

Occupation: Physicians and surgeons (2010 CENSUS CODE: 3060)

Data on 343 female and 439 male physicians and surgeons. Taking the natural logarithm of earnings is crucial for us as we wish to examine the percentage difference between male and female income.

Task 1 - Unconditional Gender Gap

Model 1:
Female physicians and surgeons earn on average 8% less than their male colleagues in our data. The R-squared is extremely low (0.002), meaning the variance in the ln of wages is poorly explained by the model. The CI suggests that the true value of this wage difference lies between 20% less or 4% more wage for women in this occupation. With a p-value of 0.192 we lack statistically significant evidence to claim that the wage for female and male physicians differs at 0.05 significance level.

Model 2:
Despite adjusting for potential heteroscedasticity, Model 2 shows the same slope but with a wider CI and higher p-value indicating more uncertainty in the estimation of the true value of the slope coefficient.

Task 2 - Variance on Level of Education

Model 3:
The Baseline model is identical to Model 1 from Task 1. Controlling for professional education and Ph.D. degrees narrows the wage difference to 5.8% less (7.6% for Master's) on average for women. However, this coefficient remains statistically insignificant for us. Notably, this model tells us about individuals with professional education earn on average 29.4% more than those with a Master's degree. Similarly, individuals with Ph.D. degrees see a 26.4% increase in earnings, both coefficients are statistically significant. Introducing Master's education shows that individuals with a Master's degree in our data earn on average 12.3% less than those with a Ph.D. Controlling for education also improves the model's explanatory power as we see an increased R-squared of 0.014.

Model 4:
This model looks at age and the ln of wages. It shows for each year of age, females can expect to get 1,8% higher salary at the 99% confidence level. However, for males, at the same confidence level this value is 2.4%. The interaction coefficient also shows that there is a 0.6% differential effect of age on the natural logarithm of wages for females compared to males. However, since the interaction coefficient is not statistically significant, we cannot do inference about gender differences.

Model 5:
This model combines Model 3 and Model 4. This now better describes the variance in the dependent variable as this model demonstrates increased values for R-squared. In our data females with a Professional or Ph.D. degree earn 2.8% less than males compared to those who only have Master's degree. This coefficient is 4.4% when looking at female Master's degrees compared to those who have a Ph.D.

Summary:

In conclusion, our analysis reveals that among Physicians and Surgeons in our dataset, the wage gap diminishes progressively with higher levels of education. However, this conclusion is specific to our data, as our coefficients are not statistically significant at the 0.05 significance level. Nonetheless, we can infer that for each year of age males can expect to receive a 0.6% higher wage than females.