Truett Bloxsom

Problem Set 2

Professor: Tom Vogl

TA: Bokyung Kim

ECO 348K (34340)

1. The Mincerian Wage Equation conceptually predicts that wage as a function of number of years of education and years of experience or working years with an experience squared term to account for range of significance of experience. Both education and experience lead to a larger wage but a person can only be in school or working not both. Beta 1 can be defined as the effect of each year of education on wage holding experienced fixed. There is a squared experience term since the first few years of experience are much more important that the last few years of experience and the squared term takes this into account.

2. Table 1 shows a summary of the CPS data after cleaning it



The mean age of the data set is 43 years old. 56% of the sample were male and 44% were female. 78% of the sample were white and 12% being black with the rest being other minorities. Mean hours worked is 43.49. The average education is 14 years so the average person had an associates degree or 2 years of college. The average experience was 24 years.

3. Table 2 shows the univariate regression of log hourly wage on education



Table 3 shows the summary statistics of education if not missing wage data



A one point increase in education is predicted to increase wage by 11.91%.

(2.762406)^2 \* .1024975 = .78214683 this is the covariance between educ\_new and lwage.

.78214683 / (2.762406\*.7316781) = .38697305 this is the correlation between educ\_new and lwage which is within 5 decimal places of the correlation found by the regression in table 2. The correlation between education and log wage is .38697305, and if you square the correlation, you get .14974814 which is exactly the R-squared of the regression in table 2 if you round to 4 decimal places. This is because R-squared is the square of correlation between the true and predicted outcomes.

4. Table 4 shows the Mincerian Wage Equation regression using robust standard errors



A one point increase in education is predicted to increase wage by 11.06% holding experienced fixed. This is intriguing since the return to education went up adding experience to the regression.

5. Table 5 shows the extended Mincerian Wage Equation regression



The education coefficient does change after controlling for race and sex. A one point increase in education is predicted to increase wage by 11.47% holding other variables fixed.

6. Using an F test, I found that the difference between the female-male log wage gap and the black-white wage log gap was highly statistically different with an F value of 84.41.

7. Table 6 shows the extended Mincerian Wage Equation regression with the gender-education interaction term



The estimated return to education is statistically different for men and women looking at the t-score for the educ\_male interaction term with a value of -6.22. This means that a 1 year increase in education is predicted to increase a females wage by 1.45% more than if a male increased his education by 1 year.

Table 7 shows the estimated ratio of the return of education for men and women



The ratio of the return to education for men and women is statistically different from 1 with a z-score of 49.91.

8. Table 8 shows summarized for the NSLY data



The mean age is 46 but the variance is much lower than the CPS data with a minimum age of 42 and a maximum of 50. 14% of the children spoke a foreign language in their childhood household. 76% of the children grew up in an urban environment. 55.29% of the dataset were male. The average experience is 26.6 years. 81.65% of the dataset is white, 11.61% is black, and 6.74% is Hispanic.

9. Table 9 shows the regression of the extended Mincerian Wage Equation with controls for race and sex



The return to education coefficient for the CPS dataset was 12.34% and the return to education for the NSLY dataset is 11.97% which is similar to the CPS dataset. There is a difference in the t-stat for both: the CPS t-stat for education was 95.01 and for the NLSY dataset it is 14.77. Both t-stats are statistically significant, but the CPS t-stat is more than 6 times greater than that of the NSLY dataset. For experience the CPS data set coefficient was 2.22% and for the NSLY dataset it is -2.00%. The t-stat for the NLSY experience variable is -0.47 which is not statistically significant at even the 10% level. The CPS t-stat for experience is 18.80 which is highly statistically significant at the 1% level.

The reason why the CPS experience variable is significant is because of the variance of age of the dataset. The standard deviation of the age variable in the CPS dataset is 10.68 with a mean of 43.38. This means that there are more people in the data set with few years of experience compared to that of the NLSY dataset with a mean of 45.50 and a standard deviation of 2.23. Experience increases earnings more when said person has few years of experience compared to someone with many years of experience. This can be proven by comparing the experience squared terms of both datasets. In the CPS dataset, the experienced squared coefficient is significant at the 1% level and is negative. This means that with every extra year of experience, the return to wage of experience decreases. The experience squared statistic in the NLSY dataset is not significant. Since the CPS data set has more young people with less average experience, the gain of 1 more year of experience is more significant than the average person in the NLSY dataset who on average has more experience already.

10. I would argue that Beta 1 does not represent the causal effect of education on wages. The education coefficient and t-stat are highly significant while including other significant variables in the regression. But for education to be represent a causal effect in regression, the regression must include all other predictors of wage to be accurate. As seen by the R-squared statistic, the model only explains 21.08% of the variation in wage. This could lead to omitted variable bias in the model which is considered in the next regression in table 10.

11. Table 10 shoes the extended Mincerian Wage Equation regression with the AFTQT test, mothers education, and fathers education variables added



Both the coefficient and t-stat for education dropped significantly when the added variables were included especially the AFQ test. A one point increase in education is predicted to increase wage by 6.77% holding other variables fixed. So the significance of education dropped almost 50% with including the new independent variables. The most significant variable included was the AFQ test variable with a t-score of 7.78. The test variable can be interpreted as an ability or IQ variable. So once intelligence was included in the model, the gain to education dropped.

12. As seen in question 11, adding an ability variable changed the coefficients on the other significant variables. The correlation between education and the AFQ test is .5774 and the correlation between AFQ test and log wage is .3835. This leads to the conclusion that there was omitted variable biased before the AFQ test was added to the regression variables. Wage is a very hard dependent variable to predict and even with the many independent variables included in the regression in table 11, the variation in the independent variables could only explain 23.08% of the variation in the wage variable. Because of these conclusions the OLS does not deliver accurate casual estimates of the Mincerian Wage Equation.