Econ 753 HW3

Evan Wasner

10/16/2020

Part 1 - Berndt Exercises

Should summarize all of quesiton 1 here

Question 1a

In part a, we calculate the mean hourly and annual wage using two methods. The first takes the arithmetic mean of the logarithm of hourly wages and exponentiates it to get the geometric mean. The second method *begins* by exponentiating the logarithm of hourly wages and *then* takes the arithmetic mean. The annual wage is computed by multiplying the two derived mean hourly wages by 2,000 (assuming a working year contains 2,000 hours). The results are given in Table 1. The arithmetic mean is 13% larger than the geometric mean. The fact that there is a difference between the two is due to the fact that performing a logarithm (and its inverse, exponentiation) is not a linear transformation, and so the exponent of a mean logarithm does not equal the mean of an exponentiated logarithm. The fact that the geometric mean is smaller is due to the fact that the logarithm operator decreases the values of large numbers more than it decreases the values of small numbers, thus taking the logarithm of a set of numbers greater than one will decrease the mean.

In Table 2, the means and standard deviations of years of schooling and years of potential experience are given. We can see that the mean years of education is 12.54, which indicates that the average worker does not have a college education. The mean years of experience is 18.7, and the average age is likely the around the sum of the average years of education and the average years of experience. The standard deviation for years of experience is quite high (13.4), which indicates a wide range of ages in the sample. The standard deviation for years of education is lower (2.7), as the years of education will likely hover around the number of years of a high school education.

Table 1: Geometric and Arithmetic Mean for 1978 Wages

	Geometric Mean	Arithmetic Mean
Hourly	5.37	6.06
Annual	10741.87	12125.53

Table 2: Mean and Standard Deviation of Education and Experience

	Mean	Standard Deviation
Education	12.54	2.77
Experience	18.72	13.35

Table 4: Years of Education and Mean Wages by Race

	Education		lnWage		
	Mean	SD	Mean Hourly	SD	Mean Annual
Gender					
Male	12.40	3.05	6.13	3.42	12256.64
Female	12.76	2.22	4.32	2.50	8632.72
Race					
Whites	12.81	2.49	5.55	3.38	11100.34
NonWhites	11.72	3.44	4.54	2.55	9084.33
Hispanic	10.31	3.66	4.62	1.96	9233.10

Question 1b

In part b, we calculate the mean values for the dummy variables nonwh (indicates whether the individual is nonwhite and non-Hispanic), hisp (indicates whether the individual is nonwhite and Hispanic), and fe (indicates whether the individual is female). Since they are dummy variables, these means give the proportion of individuals in the sample that are Black, nonwhite Hispanic, and female, respectively. We can then multiply these proportions by the sample size (550) in order to find the number of individuals in the sample in each demographic group. The results are given in Table 3. There are 57 Black individuals, 36 nonwhite Hispanic individuals, and 207 female individuals (nearly half of the sample) in the sample.

Table 3: Means and Number of Individuals in Demographic Groups

	Mean	Count
nonwh	0.10	57
hisp	0.07	36
fe	0.38	207

Question 1c

In part c, we subdivide the data first by gender and then by race. Means and standard deviations for years of education as well as geometric mean wages and standard deviations of wages are calculated for all demographic groups. This data provides means and standard deviations of the raw data, without controlling for factors such as human capital. The results are given in table4.

We can see that, on average, men make more than $1.50 more than women despite having slightly less years of education. Whites make around $1 more than both Blacks and Hispanics, although Whites also on average have more years of education than the other two groups. Males have a higher standard deviation of wages than women, and Whites have a high standard deviation of wages than Blacks and Hispanics. This may be due to the fact that women and racial minorities are often excluded from certain occupations and thus may be more crowded into certain occupations with similar wage rates, as opposed to Whites who have access to a wide spectrum of pay grades.

Question 1d

In question 1d, we repeat exercises 1a-c with the 1985 data set and compare results. These results are summarized in Tables 5 and 6. In Table 5, we compare the demographic makeup of the survey samples. The racial composition of the sample did not change substantially, with the percentage of Blacks increasing from

Table 5: Comparison of demographics between 1978 and 1985

	1978		1985	
	Mean	Count	Mean	Count
nonwh	0.10	57	0.13	67
hisp	0.07	36	0.05	27
fe	0.38	207	0.46	245

Table 6: Comparison of education and real wages between 1978 and 1985

	Education					lnWage				
	1978		1985			1978		1985		
	Mean	SD	Mean	SD	%Change	Mean Hourly	SD	Mean Hourly	SD	%Change
Gender										
Male	12.40	3.05	13.01	2.77	0.05	6.13	3.42	5.29	3.21	-0.14
Female	12.76	2.22	13.02	2.43	0.02	4.32	2.50	4.19	2.86	-0.03
Race										
Whites	12.81	2.49	13.17	2.48	0.03	5.55	3.38	4.89	3.17	-0.12
NonWhites	11.72	3.44	12.64	2.60	0.08	4.54	2.55	4.33	2.47	-0.05
Hispanic	10.31	3.66	11.52	4.05	0.12	4.62	1.96	3.74	3.35	-0.19
Total										
Whole Sample	12.54	2.77	13.02	2.62	0.04	5.37	3.26	4.75	3.12	-0.11

10% in 1978 to 13% in 1985 and the percentage of nonwhite Hispanics decreasing from 7% to 5%. However, there was a large increase in the gender composition of the samples, with the percentage of females increasing from 38% in 1978 to 46% in 1985.

In Table 6, years of education and wages are compared for demographics of each year. The mean years of education increased across all demographic groups. With regards to gender, whereas females had 0.36 more years of education than men in 1978, in 1985 mean years of education were about equal for both. However, peculiarly, mean real wages decreased more substantially for males than for females even though male education increased by more than female education, with male real wages falling by 14% and female real wages falling by 3%. Similar patterns can be seen across race, as Hispanics had the largest percent increase in years of education while also seeing the largest percent decrease in real wages. Overall, all demographics saw a decrease in real wages, but the disproportionate changes in education are not broadly consistent with the human capital model for the various subgroups.

Question 1e

In question 1e, we compare how the logarithm of wages and wages themselves fit into a normal distribution. We begin by splitting the data into six brackets by standard deviations centered around the means. We then calculate the number of observations sitting within each bracket and compare the proportion of the sample size within each bracket to the proportion that would be expected from a normal distribution. From the results in Table 7, it is very clear that the proportion of observations within each standard deviation bracket of the logarithm of wages matches much better with that of a normal distribution than wages themselves.

We can then run chi-squared goodness-of-fit tests for both lnwages and wages against a normal distribution. The results are given in Table 8. The null hypothesis is that the given set of values fits with a normal distribution. As expected based on the data in Table 7, for lnwage, we receive a statistic of 3.74 and a

Table 7: Counts and Proportions of lnwage and wage Distributions

	lnwage		wage		
	Counts	Proportion	Counts	Proportion	Normal Distribution Proportions
wi<=w-2s	9	0.02	0	0.00	0.02
w-2s<wi<=w-s	80	0.15	45	0.08	0.14
w-s<wi<=w	194	0.35	278	0.51	0.34
w<wi<=w+s	193	0.35	164	0.30	0.34
w+s<wi<=w+2s	63	0.11	45	0.08	0.14
w+2s<wi	11	0.02	18	0.03	0.02

Table 8: Chi-Squared Goodness-of-Fit Tests for Normal Distribution

	Statistic	pValue
lnwage	3.74	0.59
wage	84.99	0.00

p-value of 0.59. Thus we fail to reject the null hypothesis that the data fits a normal distribution. However, again as expected, running the test on wages gives us a statistic of 85 and a p-value very close to 0, meaning we reject the null hypothesis that the data fits a normal distribution.