

# Econ 753 HW3

Evan Wasner

10/16/2020

## Part 1 - Berndt Exercises

### 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
<b>Gender</b>					
Male	12.40	3.05	6.13	3.42	12256.64
Female	12.76	2.22	4.32	2.50	8632.72
<b>Race</b>					
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		%Change	1978		1985		%Change	
	Mean	SD	Mean	SD		Mean	Hourly	SD	Mean		Hourly
<b>Gender</b>											
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
<b>Race</b>											
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
<b>Total</b>											
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 p-value of 0.59. Thus we fail to reject the null hypothesis that the data fits a normal distribution. However,

Table 7: Counts and Proportions of lnwage and wage Distributions

	lnwage		wage		Normal Distribution Proportions
	Counts	Proportion	Counts	Proportion	
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

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.

### Question 6a

For exercise 6, we begin by regressing lnwage on dummy variables for whether a worker is female, in a union, Black, Hispanic, and on variables for years of education and experience (allowing for a quadratic effect of experience). The results are given in Table 9. The interpretation of the coefficients on nonwh and hisp are wage premiums (in percentage terms) for Blacks and Hispanics compared to Whites controlling for gender, unionization, education and experience. According to this regression, Blacks on average have wages 16% lower than Whites, which is statistically significant. The negative premium for Hispanics, however, is not statistically significant (-2.7%).

In order to test the hypothesis that the coefficients on nonwh and hisp are equal to each other but not necessarily zero, we run a Chow test to test the linear hypothesis: “nonwh - hisp = 0”. If we reject the null hypothesis, then the coefficients are not equal to each other and not 0. The results of the Chow test are a F-statistic of 2.46 and a p-value of 0.12, which means we fail to reject the null hypothesis at a 10% confidence level. Therefore, we cannot reject the hypothesis that the premiums for Blacks and Hispanics are equal and/or both 0.

### Question 6b

In 6b, we separate the data by race. First, we observe the fraction of the data source in each category, shown in Table 10. The majority of the sample is white (83%), with 10% of the sample Black and 7% Hispanic.

In Table 11, the means of several variables calculated by racial group is displayed, as well as the difference between each mean with those of Whites. The first thing to notice is that Whites on average have the highest wages. On average, Whites have the most education but the least experience. Blacks have 20% lower wages than those of Whites despite having several more years of experience and having a higher proportion of workers in unions. However, the Black population of the sample also has a higher percentage of females, and so the negative wage premium may also be including the negative premium for women. While Hispanics

Table 9: Simple Regression of log Wages to Determine Wage Discrimination

	<i>Dependent variable:</i>
	lnwage
Constant	0.488*** (0.098)
fe	-0.306*** (0.034)
union	0.207*** (0.037)
nonwh	-0.157*** (0.055)
hisp	-0.027 (0.069)
ed	0.075*** (0.007)
ex	0.026*** (0.005)
exsq	-0.0003*** (0.0001)
Observations	550
R <sup>2</sup>	0.392
Adjusted R <sup>2</sup>	0.385
Residual Std. Error	0.385 (df = 542)
F Statistic	50.012*** (df = 7; 542)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Table 10: Means and Number of Individuals in Demographic Groups

	Mean	Count
other	0.83	457
nonwh	0.10	57
hisp	0.07	36

Table 11: Comparison of Demographics Variables

	Education		Experience		Female		Unionized		Wages	
	ed	diff	ex	diff.1	fe	diff.2	union	diff.3	lnwage	diff.4
Others	12.81	0.00	18.11	0.00	0.37	0.00	0.30	0.00	1.71	0.00
NonWhites	11.72	-1.09	22.16	4.05	0.49	0.13	0.39	0.08	1.51	-0.20
Hispanic	10.31	-2.51	21.06	2.95	0.33	-0.03	0.19	-0.11	1.53	-0.18

have less experience, education, rates of unionization, and percentage of females compared to Blacks, they have less of a negative premium than Blacks (-18%).

### Question 6c

In part c, we run three separate regressions of  $\ln wage$  for the different racial data subsets. In this way, we are able to observe any differences in the effects of different variables by race. The results are given in Table 12. There are differences in the coefficients across races. For education, each year of education has 7% impact on wages for others, but that reduces to a 5.5% impact on wages for nonwhites and Hispanics. For others, each year of experience provides a statistically significant 3% impact on wages for others, but the coefficients on experience cease to be statistically significant for nonwhites and others. There is a negative impact for being female across all racial groups, but this impact is only statistically significant for others and Hispanics. Being in a union provides a positive impact on wages across all racial groups, but is not statistically significant for nonwhites. The effect of being in a union is strongest for Hispanics, indicating a 42% impact on wages. An interpretation of these results may be that, for nonwhites and Hispanics, employers may assume that experience and education do not provide as much as they do for others.

### Questions 6d-h

I am not completing the remainder of question 6 because I have already spent 25 hours on this problem set and there are simply not enough hours in each day for me to possibly complete it. Please, for the love of God, have mercy.

## Part 2 - Replication of

In part 2

Table 12: Regressions of lnwage by Racial Category

	<i>Dependent variable:</i>		
	Others	lnwage NonWhites	Hispanics
	(1)	(2)	(3)
Constant	0.456*** (0.111)	0.755* (0.389)	0.682*** (0.234)
fe	-0.338*** (0.037)	-0.059 (0.144)	-0.235** (0.104)
union	0.197*** (0.040)	0.205 (0.139)	0.417*** (0.123)
ed	0.077*** (0.008)	0.054** (0.026)	0.055*** (0.015)
ex	0.028*** (0.005)	0.003 (0.020)	0.021 (0.014)
exsq	-0.0003*** (0.0001)	-0.00001 (0.0004)	-0.0003 (0.0003)
Observations	457	57	36
R <sup>2</sup>	0.420	0.149	0.489
Adjusted R <sup>2</sup>	0.414	0.066	0.404
Residual Std. Error	0.376 (df = 451)	0.484 (df = 51)	0.287 (df = 30)
F Statistic	65.423*** (df = 5; 451)	1.789 (df = 5; 51)	5.752*** (df = 5; 30)

*Note:*

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Table 13: Test

Variable	Stores By State			Stores in NJ			Differences in NJ	
	PA	NJ	NJ-PA	Wage=4.25	Wage=4.26-4.99	Wage>4.99	Low-high	Midrange-high
FTE before	23.33 (1.35)	20.44 (0.51)	-2.89 (1.44)	19.56 (0.77)	20.08 (0.84)	22.25 (1.14)	-2.69 (1.37)	-2.17 (1.41)
FTE after	21.17 (0.94)	21.03 (0.52)	-0.14 (1.08)	20.88 (1.01)	20.96 (0.76)	20.21 (1.03)	0.66 (1.44)	0.74 (1.27)
Change in mean FTE	-2.17 (0.73)	0.59 (1.65)	2.75 (1.34)	1.32 (1.13)	0.87 (1.53)	-2.04 (1.27)	3.36 (1.3)	2.91 (1.22)

Table 14: Test

Variable	Wage=4.25	Wage=4.26-4.99	Wage>4.99
FTE before	19.56 (0.77)	20.08 (0.84)	22.25 (1.14)
FTE after	20.88 (1.01)	20.96 (0.76)	20.21 (1.03)
Change in mean FTE	1.32 (1.13)	0.87 (1.53)	-2.04 (1.27)

Table 15: Test

Variable	Wage=4.25	Wage=4.26-4.99	Wage>4.99
FTE before	19.56 (0.77)	20.08 (0.84)	22.25 (1.14)
FTE after	20.88 (1.01)	20.96 (0.76)	20.21 (1.03)
Change in mean FTE	1.32 (1.13)	0.87 (1.53)	-2.04 (1.27)