Chris Kalra

Econ 610

Homework 6

1. a)





If someone is black, we set *black* = 1, which causes a decrease in the estimated log(wage) by

.188. This means blacks have a lower monthly salary of approximately 18.8% compared to

non-blacks, and with |t| ≈ 5 and an associated p-value of <.00001, the difference is extremely

statistically significant.

b) By using the anova function in R, we receive an F value of 1.49 on 2 and 925

degrees of freedom, resulting in a p-value of .226, meaning *exper2* and *tenure2* are jointly

insignificant at the 20% level

c) By adding the interaction term *educ\*black* to the original equation, we see that it has |t| = 1.121 and an associated p-value of .262, meaning the interaction is not statistically significant, and that education does not depend on race

d) If we remove *single* and *black* as variables from our regression model, then single and black jointly serve as the baseline group for which we can compare the other 3 groups (and it helps us avoid perfect collinearity in our model). The estimated ceteris paribus wage differential between married blacks and non-married blacks = 0.009448 – 0.188915 = –.179467, meaning married blacks earn an estimated 17.95% less than married non-blacks

1. a) For men,



For women,



One key difference I notice is that men with a young child sleep an additional 60 minutes longer than men without a child, but women with a young child sleep nearly 2 hours (118 minutes) fewer than women without a young child.  
Furthermore, the intercepts are much different.  
Lastly, sleep increases for men with age, but decreases with squared age. For women, sleep decreases with age, but increases with squared age.

b) The chow statistic returned is 2.116351 on 6 and 694 degrees of freedom, resulting in a borderline p-value of .04949. Because the p-value is <.05, we reject the null hypothesis, and conclude that the 2 equations are not the same at the .05 level

c) The chow statistic returned is 1.2558 on 5 and 694 degrees of freedom, resulting in a p-value of .2814. Therefore, we fail to reject the null hypothesis, which states that the interaction terms involving *male* are jointly significant

d) From the above questions and answers, we can see that the interactions involving *male* are not jointly statistically significant, but the models for men and women do differ. Because of this, I believe the ideal model would have *male* as an explanatory variable, but it would not have any of the interaction terms

1. a)





Note that the coefficient on *tenure2* is positive even though the estimate for β7 is negative; this is because we are estimating – β7

When *educ* = 0 and we take the partial derivatives of both sides with respect to *female*, we have the gender differential when education is 0, resulting in



Therefore, the gender differential for someone with 0 years of education is a wage decrease of 22.679% for women

b) *female* is now the gender differential for someone with 12.5 years of education (and 12.5 is very close to sample mean of years of education), with an estimate of –.296. This means that the gender differential for someone with 12.5 years of education is a wage decrease of 29.6% for women

c) With an |t| = 8.27 and an associated p-value of <.00001, the coefficient of *female* from (b) is extremely statistically significant. Although the coefficients of *female* from both models are fairly close (a difference of less than .07), *female* is not statistically significant in the model from (a), as it has a |t| = 1.354 and an associated p-value of .1764

*Code and Output*

1a)

*wage2=data*

*full=lm(lwage~educ+exper+tenure+married+black+south+urban, data=wage2)*

*summary(full)*

Call:

lm(formula = lwage ~ educ + exper + tenure + married + black +

south + urban, data = wage2)

Residuals:

Min 1Q Median 3Q Max

-1.98069 -0.21996 0.00707 0.24288 1.22822

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 5.395497 0.113225 47.653 < 2e-16 \*\*\*

educ 0.065431 0.006250 10.468 < 2e-16 \*\*\*

exper 0.014043 0.003185 4.409 1.16e-05 \*\*\*

tenure 0.011747 0.002453 4.789 1.95e-06 \*\*\*

married 0.199417 0.039050 5.107 3.98e-07 \*\*\*

black -0.188350 0.037667 -5.000 6.84e-07 \*\*\*

south -0.090904 0.026249 -3.463 0.000558 \*\*\*

urban 0.183912 0.026958 6.822 1.62e-11 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3655 on 927 degrees of freedom

Multiple R-squared: 0.2526, Adjusted R-squared: 0.2469

F-statistic: 44.75 on 7 and 927 DF, p-value: < 2.2e-16

1b)

*full2=lm(lwage~educ+exper+tenure+married+black+south+urban+I(exper^2)+I(tenure^2), data=wage2)*

*summary(full2)*

*anova(full, full2)*

Analysis of Variance Table

Model 1: lwage ~ educ + exper + tenure + married + black + south + urban

Model 2: lwage ~ educ + exper + tenure + married + black + south + urban +

I(exper^2) + I(tenure^2)

Res.Df RSS Df Sum of Sq F Pr(>F)

1 927 123.82

2 925 123.42 2 0.39756 1.4898 0.226

1c)

*full3=lm(lwage~educ+exper+tenure+married+black+south+urban+(educ\*black), data=wage2)*

*summary(full3)*

Call:

lm(formula = lwage ~ educ + exper + tenure + married + black +

south + urban + (educ \* black), data = wage2)

Residuals:

Min 1Q Median 3Q Max

-1.97782 -0.21832 0.00475 0.24136 1.23226

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 5.374817 0.114703 46.859 < 2e-16 \*\*\*

educ 0.067115 0.006428 10.442 < 2e-16 \*\*\*

exper 0.013826 0.003191 4.333 1.63e-05 \*\*\*

tenure 0.011787 0.002453 4.805 1.80e-06 \*\*\*

married 0.198908 0.039047 5.094 4.25e-07 \*\*\*

black 0.094809 0.255399 0.371 0.710561

south -0.089450 0.026277 -3.404 0.000692 \*\*\*

urban 0.183852 0.026955 6.821 1.63e-11 \*\*\*

educ:black -0.022624 0.020183 -1.121 0.262603

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3654 on 926 degrees of freedom

Multiple R-squared: 0.2536, Adjusted R-squared: 0.2471

F-statistic: 39.32 on 8 and 926 DF, p-value: < 2.2e-16

1d)

*wage2$black*

*wage2$marrnonblack = (wage2$married)\*(1-(wage2$black))*

*wage2$marrblack=(wage2$married)\*(wage2$black)*

*wage2$singleblack=(1-(wage2$married))\*(wage2$black)*

*full4=lm(lwage~educ+exper+tenure+south+urban+marrblack+marrnonblack+singleblack, data=wage2)*

*summary(full4)*

Call:

lm(formula = lwage ~ educ + exper + tenure + south + urban +

marrblack + marrnonblack + singleblack, data = wage2)

Residuals:

Min 1Q Median 3Q Max

-1.98013 -0.21780 0.01057 0.24219 1.22889

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 5.403793 0.114122 47.351 < 2e-16 \*\*\*

educ 0.065475 0.006253 10.471 < 2e-16 \*\*\*

exper 0.014146 0.003191 4.433 1.04e-05 \*\*\*

tenure 0.011663 0.002458 4.745 2.41e-06 \*\*\*

south -0.091989 0.026321 -3.495 0.000497 \*\*\*

urban 0.184350 0.026978 6.833 1.50e-11 \*\*\*

marrblack 0.009448 0.056013 0.169 0.866083

marrnonblack 0.188915 0.042878 4.406 1.18e-05 \*\*\*

singleblack -0.240820 0.096023 -2.508 0.012314 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3656 on 926 degrees of freedom

Multiple R-squared: 0.2528, Adjusted R-squared: 0.2464

F-statistic: 39.17 on 8 and 926 DF, p-value: < 2.2e-16

2a)

*sleep75=data*

*ex2male=lm(sleep~totwrk + educ + age + I(age^2) + yngkid, data=subset(sleep75, sleep75$male==1))*

*summary(ex2male)*

Call:

lm(formula = sleep ~ totwrk + educ + age + I(age^2) + yngkid,

data = subset(sleep75, sleep75$male == 1))

Residuals:

Min 1Q Median 3Q Max

-1793.96 -216.05 7.93 244.57 1141.21

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 3648.20826 310.03933 11.767 < 2e-16 \*\*\*

totwrk -0.18212 0.02449 -7.438 6.45e-13 \*\*\*

educ -13.05238 7.41422 -1.760 0.0791 .

age 7.15659 14.32037 0.500 0.6175

I(age^2) -0.04477 0.16841 -0.266 0.7905

yngkid 60.38021 59.02278 1.023 0.3069

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 402.3 on 394 degrees of freedom

Multiple R-squared: 0.1562, Adjusted R-squared: 0.1455

F-statistic: 14.59 on 5 and 394 DF, p-value: 3.952e-13

*ex2female=lm(sleep~totwrk + educ + age + I(age^2) + yngkid, data=subset(sleep75, sleep75$male==0))*

*summary(ex2female)*

Call:

lm(formula = sleep ~ totwrk + educ + age + I(age^2) + yngkid,

data = subset(sleep75, sleep75$male == 0))

Residuals:

Min 1Q Median 3Q Max

-2485.02 -244.18 7.24 270.64 1376.91

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 4238.72933 384.89226 11.013 < 2e-16 \*\*\*

totwrk -0.13995 0.02766 -5.060 7.33e-07 \*\*\*

educ -10.20514 9.58885 -1.064 0.288

age -30.35657 18.53091 -1.638 0.102

I(age^2) 0.36794 0.22334 1.647 0.101

yngkid -118.28256 93.18757 -1.269 0.205

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 437 on 300 degrees of freedom

Multiple R-squared: 0.09768, Adjusted R-squared: 0.08264

F-statistic: 6.495 on 5 and 300 DF, p-value: 9.42e-06

2b)

*pooled=lm(sleep~ male + totwrk + educ + age + I(age^2) + yngkid + male\*totwrk + male\*educ + male\*age + male\*I(age^2) + male\*yngkid, data=sleep75)*

*pooledss=sum(resid(pooled)^2)*

*pooled2=lm(sleep~ totwrk + educ + age + I(age^2) + yngkid, data=sleep75)*

*pooledss2=sum(resid(pooled2)^2)*

*chow2=((pooledss2 - pooledss)/6)/(pooledss/(706-12))*

*chow2*

[1] 2.116351

2c)

*pooled3=lm(sleep~ male + totwrk + educ + age + I(age^2) + yngkid + male\*totwrk + male\*educ + male\*age + male\*I(age^2) + male\*yngkid, data=sleep75)*

*pooledss3=sum(resid(pooled3)^2)*

*pooled4=lm(sleep~ male + totwrk + educ + age + I(age^2) + yngkid, data=sleep75)*

*pooledss4=sum(resid(pooled4)^2)*

*chow3=((pooledss4-pooledss3)/5)/(pooledss3/(694))*

*chow3*

[1] 1.2558

3a)

*wage1=data*

*ex3=lm(lwage ~ female + educ + female\*educ + exper + I(exper^2) + tenure + I(-1\*tenure^2), data=wage1)*

*summary(ex3)*

Call:

lm(formula = lwage ~ female + educ + female \* educ + exper +

I(exper^2) + tenure + I(-1 \* tenure^2), data = wage1)

Residuals:

Min 1Q Median 3Q Max

-1.83265 -0.25261 -0.02374 0.25396 1.13584

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.3888060 0.1186871 3.276 0.00112 \*\*

female -0.2267886 0.1675394 -1.354 0.17644

educ 0.0823692 0.0084699 9.725 < 2e-16 \*\*\*

exper 0.0293366 0.0049842 5.886 7.11e-09 \*\*\*

I(exper^2) -0.0005804 0.0001075 -5.398 1.03e-07 \*\*\*

tenure 0.0318967 0.0068640 4.647 4.28e-06 \*\*\*

I(-1 \* tenure^2) 0.0005900 0.0002352 2.509 0.01242 \*

female:educ -0.0055645 0.0130618 -0.426 0.67028

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4001 on 518 degrees of freedom

Multiple R-squared: 0.441, Adjusted R-squared: 0.4334

F-statistic: 58.37 on 7 and 518 DF, p-value: < 2.2e-16

3b)

*wage1$educMinus=wage1$educ - 12.5*

*ex3b=lm(lwage ~ female + educ + female\*educMinus + exper + I(exper^2) + tenure + I(-1\*tenure^2), data=wage1)*

*summary(ex3b)*

Call:

lm(formula = lwage ~ female + educ + female \* educMinus + exper +

I(exper^2) + tenure + I(-1 \* tenure^2), data = wage1)

Residuals:

Min 1Q Median 3Q Max

-1.83265 -0.25261 -0.02374 0.25396 1.13584

Coefficients: (1 not defined because of singularities)

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.3888060 0.1186871 3.276 0.00112 \*\*

female -0.2963450 0.0358358 -8.270 1.14e-15 \*\*\*

educ 0.0823692 0.0084699 9.725 < 2e-16 \*\*\*

educMinus NA NA NA NA

exper 0.0293366 0.0049842 5.886 7.11e-09 \*\*\*

I(exper^2) -0.0005804 0.0001075 -5.398 1.03e-07 \*\*\*

tenure 0.0318967 0.0068640 4.647 4.28e-06 \*\*\*

I(-1 \* tenure^2) 0.0005900 0.0002352 2.509 0.01242 \*

female:educMinus -0.0055645 0.0130618 -0.426 0.67028

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4001 on 518 degrees of freedom

Multiple R-squared: 0.441, Adjusted R-squared: 0.4334

F-statistic: 58.37 on 7 and 518 DF, p-value: < 2.2e-16