Assignment 8

```r
library(here)
```

- **lwage**: Log of hourly wages.
- **educ**: Years of schooling.
- **exper**: Years of work experience (age - years of schooling - 6).
- **expersq**: Years of work experience squared.
- **union**: Is the individual part of a union? 1 if yes, 0 if not.
- **female**: Is the individual female? 1 if yes 0 if not.
- **y85**: Is the observation for the year 1985? 1 if yes, 0 if not.
- **y85educ**: Years of education for observations in the year 1985, 0 if year of observation isn't 1985.
- **y85fem**: Is the individual female and is the observation in 1985? 1 if yes, 0 if otherwise.

```r
data <- haven::read_dta(here("data", "CPS78_85.DTA"))
```

```r
ols <- lm(lwage ~ educ + exper + expersq + union, data = data)
summary <- summary(ols)
summary
```

```
##
## Call:
## lm(formula = lwage ~ educ + exper + expersq + union, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.61051 -0.30649 -0.01713  0.30776  2.19462
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.483e-01  8.501e-02   4.097 4.50e-05 ***
## educ         8.724e-02  5.769e-03  15.122  < 2e-16 ***
## exper        3.170e-02  4.071e-03   7.786 1.61e-14 ***
## expersq     -4.400e-04  8.855e-05  -4.968 7.84e-07 ***
## union        1.855e-01  3.392e-02   5.469 5.62e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4719 on 1079 degrees of freedom
## Multiple R-squared:  0.2469, Adjusted R-squared:  0.2441
## F-statistic: 88.43 on 4 and 1079 DF,  p-value: < 2.2e-16
```

Question 1

Part A

Where we are estimating the following equation:

$log(wage_{it}) = \beta_0 + \beta_1 educ_{it} + \beta_2 exper_{it} + \beta_3 exper_{it}^2 + \beta_4 union_{it} + \epsilon_{it}$

H_0: $\beta_3 \geq 0$

H_a: $\beta_3 < 0$
$Tstatistic$: -4.9684174

This leads to a very small p-value which would lead to a rejection of the null hypothesis in almost all cases. This would mean we reject the null hypothesis that the returns of work experience to wages is the same for all years of work experience.

Part B

The value of an additional year of experience for an individual with 2 years of experience on their log(wages):

$\beta_2 + 2\beta_3(exper_{it})$ where experience is 2 years.
0.0299398

Part C

The value of an additional year of experience for an individual with 20 years of experience on their log(wages):

$\beta_2 + 2\beta_3(exper_{it})$ where experience is 20 years.
0.0141008

Question 2

Where we are estimating the following equation:

$log(wage_{it}) = \beta_0 + \delta_0 y85_{it} + \beta_1 educ_{it} + \delta_1 y85_{it} * educ_{it} + \beta_2 exper_{it} + \beta_3 exper_{it}^2 + \beta_4 union_{it} + \beta_5 female_{it} + \delta_2 y85_{it} * female_{it} + \epsilon_{it}$

```
ols2 <- lm(lwage ~ y85 + educ + y85educ + exper + expersq + union + female + y85fem, data = data)
summ2 <- summary(ols2)
summ2
```

```
##
## Call:
## lm(formula = lwage ~ y85 + educ + y85educ + exper + expersq +
##     union + female + y85fem, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.56098 -0.25828  0.00864  0.26571  2.11669
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.589e-01  9.345e-02   4.911 1.05e-06 ***
## y85          1.178e-01  1.238e-01   0.952   0.3415
## educ         7.472e-02  6.676e-03  11.192  < 2e-16 ***
## y85educ      1.846e-02  9.354e-03   1.974   0.0487 *
## exper        2.958e-02  3.567e-03   8.293 3.27e-16 ***
## expersq     -3.994e-04  7.754e-05  -5.151 3.08e-07 ***
## union        2.021e-01  3.029e-02   6.672 4.03e-11 ***
## female      -3.167e-01  3.662e-02  -8.648  < 2e-16 ***
## y85fem       8.505e-02  5.131e-02   1.658   0.0977 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4127 on 1075 degrees of freedom
## Multiple R-squared:  0.4262, Adjusted R-squared:  0.4219
## F-statistic:  99.8 on 8 and 1075 DF,  p-value: < 2.2e-16
```

Part A

$E(log(wage_{it})|y85 = 0) = \beta_0$

Part B

$E(log(wage_{it})|y85 = 1) = \beta_0 + \delta_0$

Part C

δ_0 is the difference in the logged mean earnings between 1978 and 1985.

Part D

$E(log(wage_{it})|educ_{it}, y85 = 0) = \beta_0 + \beta_1 educ_{it}$

Part E

$E(log(wage_{it})|educ_{it}, y85 = 1) = \beta_0 + \delta_0 + (\beta_1 + \delta 1)educ_{it}$

Part F

The difference in coefficients concerning the returns of education in 1978 vs. 1985 is δ_1.

H_0: $\delta_1 = 0$, there is no difference in the returns of education in 1978 vs. 1985
H_a: $\delta_1 \neq 0$, there is a difference in the returns of education in 1978 vs. 1985
T-statistic: 1.9735085
P-value: 0.0486934

At the 0.05 significance level, we could reject the null hypothesis and say the returns to education have changed between 1978 and 1985.

Part G

$E(log(wages_{it}|female = 1, y85 = 0) = \beta_0 + \beta_5$
$E(log(wages_{it}|female = 0, y85 = 0) = \beta_0$

```
meanFem1978 <-
    summ2[["coefficients"]]["(Intercept)", "Estimate"] +
    summ2[["coefficients"]]["female", "Estimate"]

meanMale1978 <- summ2[["coefficients"]]["(Intercept)", "Estimate"]
```

Mean Female Log Earnings in 1978: 0.1422242
Mean Male Log Earnings in 1978: 0.4589329
Sex Gap in Log Earnings in 1978 (Men made this much more money than women): 0.3167086

Part H

$E(log(wages_{it}|female = 1, y85 = 1) = \beta_0 + \delta_0 + \beta_5 + \delta_2$
$E(log(wages_{it}|female = 0, y85 = 1) = \beta_0 + \delta_0$

```
meanFem1985 <-
    summ2[["coefficients"]]["(Intercept)", "Estimate"] +
    summ2[["coefficients"]]["y85", "Estimate"] +
    summ2[["coefficients"]]["female", "Estimate"] +
    summ2[["coefficients"]]["y85fem", "Estimate"]

meanMale1985 <-
```

```
    summ2[["coefficients"]]["(Intercept)", "Estimate"] +
    summ2[["coefficients"]]["y85", "Estimate"]
```

Mean Female Log Earnings in 1985: 0.3450824
Mean Male Log Earnings in 1985: 0.5767391
Sex Gap in Log Earnings in 1985 (Men made this much more money than women): 0.2316567

Part G

δ_2 is the difference in the sex gap between 1978 vs. 1985. We can show this algebraically.

Sex gap in 1978: $\beta_0 - \beta_0 + \beta_5 = \beta_5$
Sex gap in 1985: $\beta_0 + \delta_0 - \beta_0 + \delta_0 + \beta_5 + \delta_2 = \beta_5 + \delta_2$
Meaning of δ_2: $\beta_5 - \beta_5 + \delta_2 = \delta_2$ (Subtracting the sex gaps)