10.18

(a)

```
> cat( mean(data$MOTHERCOLL)*100, "%", "\n" )
12.14953 %
> cat( mean(data$FATHERCOLL)*100, "%", "\n" )
11.68224 %
```

(b)

```
> cor(data$educ, data$MOTHERCOLL, use = "complete.obs")
[1] 0.3594705
> cor(data$educ, data$FATHERCOLL, use = "complete.obs")
[1] 0.3984962
> cor(data$MOTHERCOLL, data$FATHERCOLL, use = "complete.obs")
[1] 0.3545709
```

EDUC與MOTHERCOLL·FATHERCOLL呈正相關

(c)

```
> confint(iv1, "educ", level = 0.95)
              2.5 %     97.5 %
educ -0.001440868 0.1534768
```

(d)

```
Anova Table (Type II tests)

Response: educ
             Sum Sq  Df F value    Pr(>F)
MOTHERCOLL   289.32   1 63.5631 1.455e-14 ***
exper          8.12   1  1.7838    0.1824
I(exper^2)    11.04   1  2.4254    0.1201
Residuals   1929.90 424
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

$H_0: \beta_{MOTHERCOLL} = 0$

$H_a: \beta_{MOTHERCOLL} \neq 0$

$|63.563| > 1.0$ don't reject H_0

⇒ MOTHERCOLL has a significant effect on EDUC. And MOTHERCOLL is a strong instrument.

```
Call:
lm(formula = educ ~ MOTHERCOLL + exper + I(exper^2), data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-7.4267 -0.4826 -0.3731  1.0000  4.9353

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 12.079094   0.303118  39.849  < 2e-16 ***
MOTHERCOLL   2.517068   0.315713   7.973 1.46e-14 ***
exper        0.056230   0.042101   1.336    0.182
I(exper^2)  -0.001956   0.001256  -1.557    0.120
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.133 on 424 degrees of freedom
Multiple R-squared:  0.1347,    Adjusted R-squared:  0.1285
F-statistic: 21.99 on 3 and 424 DF,  p-value: 2.965e-13
```

$EDUC = 12.079094 + 2.517068\,MOTHERCOLL + 0.056230\,EXPER - 0.001956\,EXPER^2 + v$

(e)

```
> confint(iv2, "educ", level = 0.95)
              2.5 %     97.5 %
educ 0.02734574 0.1483496
```

narrower than (c) [-0.00144, 0.15348]

(f)

```
Diagnostic tests:
                 df1 df2 statistic p-value
Weak instruments   2 423    56.963  <2e-16 ***
Wu-Hausman         1 423     0.519   0.472
Sargan             1  NA     0.238   0.626
---
```

$56.963 > 1.0$ ⇒ reject H_0

⇒ The IV is not strong.

(g)

```
Call:
ivreg(formula = log(wage) ~ educ + exper + I(exper^2) | MOTHERCOLL +
    FATHERCOLL + exper + I(exper^2), data = data)

Residuals:
     Min       1Q   Median       3Q      Max
-3.07797 -0.32128  0.03418  0.37648  2.36183

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.2790819  0.3922213  -0.712  0.47714
educ         0.0878477  0.0307808   2.854  0.00453 **
exper        0.0426761  0.0132950   3.210  0.00143 **
I(exper^2)  -0.0008486  0.0003976  -2.135  0.03337 *

Diagnostic tests:
                 df1 df2 statistic p-value
Weak instruments   2 423    56.963  <2e-16 ***
Wu-Hausman         1 423     0.519   0.472
Sargan             1  NA     0.238   0.626
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6679 on 424 degrees of freedom
Multiple R-Squared: 0.153,      Adjusted R-squared: 0.147
Wald test: 9.724 on 3 and 424 DF,  p-value: 3.224e-06
```

Do not reject H_0.

The IV on this model doesn't have any problems. The extra instruments are valid

/0.20

(a)

```
Call:
lm(formula = msft ~ mkt, data = capm5)

Residuals:
     Min       1Q   Median       3Q      Max
-0.27431 -0.04753 -0.00882  0.03821  0.35730

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.002737   0.006058   0.452    0.652
mkt         1.208834   0.122466   9.871   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.08079 on 178 degrees of freedom
Multiple R-squared:  0.3537,    Adjusted R-squared:  0.3501
F-statistic: 97.43 on 1 and 178 DF,  p-value: < 2.2e-16
```

estimate > | ⇒ mkt stock is risky

(b)

```
Call:
lm(formula = rm_rf ~ RANK, data = capm5)

Residuals:
      Min        1Q    Median        3Q       Max
-0.110497 -0.006308  0.001497  0.009433  0.029513

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -7.903e-02  2.195e-03   -36.0   <2e-16 ***
RANK         9.067e-04  2.104e-05    43.1   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.01467 on 178 degrees of freedom
Multiple R-squared:  0.9126,    Adjusted R-squared:  0.9121
F-statistic:  1858 on 1 and 178 DF,  p-value: < 2.2e-16
```

F-stat > 10 ⇒ RANK has a strong IV and very significant.

(c)

```
Call:
lm(formula = rj_rf ~ rm_rf + v, data = capm5)

Residuals:
     Min       1Q   Median       3Q      Max
-0.27140 -0.04213 -0.00911  0.03423  0.34887

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.003018   0.005984   0.504   0.6146
rm_rf        1.278318   0.126749  10.085   <2e-16 ***
v           -0.874599   0.428626  -2.040   0.0428 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.08012 on 177 degrees of freedom
Multiple R-squared:  0.3672,    Adjusted R-squared:   0.36
F-statistic: 51.34 on 2 and 177 DF,  p-value: < 2.2e-16
```

(d)

part(a): OLS

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.003250   0.006036   0.538    0.591
rm_rf       1.201840   0.122152   9.839   <2e-16 ***
```

part(d): 2SLS

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.003018   0.006044   0.499    0.618
rm_rf       1.278318   0.128011   9.986   <2e-16 ***
```

(e)

```
Call:
lm(formula = rm_rf ~ RANK + POS, data = capm5)

Residuals:
      Min        1Q    Median        3Q       Max
-0.109182 -0.006732  0.002858  0.008936  0.026652

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.0804216  0.0022622  -35.55   <2e-16 ***
RANK         0.0009819  0.0000400   24.55   <2e-16 ***
POS         -0.0092762  0.0042156   -2.20   0.0291 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.01451 on 177 degrees of freedom
Multiple R-squared:  0.9149,    Adjusted R-squared:  0.9139
F-statistic: 951.3 on 2 and 177 DF,  p-value: < 2.2e-16
```

⇒ RANK has strong IV.

(f)

```
Call:
lm(formula = rj_rf ~ rm_rf + v2, data = capm5)

Residuals:
     Min       1Q   Median       3Q      Max
-0.27132 -0.04261 -0.00812  0.03343  0.34867

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.003004   0.005972   0.503   0.6157
rm_rf        1.283118   0.126344  10.156   <2e-16 ***
v2          -0.954918   0.433062  -2.205   0.0287 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.07996 on 177 degrees of freedom
Multiple R-squared:  0.3696,    Adjusted R-squared:  0.3625
F-statistic: 51.88 on 2 and 177 DF,  p-value: < 2.2e-16
```

```
Analysis of Variance Table

Response: rm_rf
           Df  Sum Sq Mean Sq   F value  Pr(>F)
RANK        1 0.39955 0.39955 1897.6825 < 2e-16 ***
POS         1 0.00102 0.00102    4.8421 0.02907 *
Residuals 177 0.03727 0.00021
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

P-value > 0.01, do not reject Ho. The market return is exogenous.

(g)

```
Call:
ivreg(formula = rj_rf ~ rm_rf | RANK + POS, data = capm5)

Residuals:
     Min       1Q   Median       3Q      Max
-0.27168 -0.04960 -0.00983  0.03762  0.35543

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.003004   0.006044   0.497     0.62
rm_rf       1.283118   0.127866  10.035   <2e-16 ***

Diagnostic tests:
                  df1 df2 statistic p-value
Weak instruments    2 177   951.262  <2e-16 ***
Wu-Hausman          1 177     4.862  0.0287 *
Sargan              1  NA     0.558  0.4549
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.08093 on 178 degrees of freedom
Multiple R-Squared: 0.3507,     Adjusted R-squared: 0.347
Wald test: 100.7 on 1 and 178 DF,  p-value: < 2.2e-16
```

Because β is higher (1.2018 < 1.2831) ⇒ catch some variables that OLS doesn't catch.

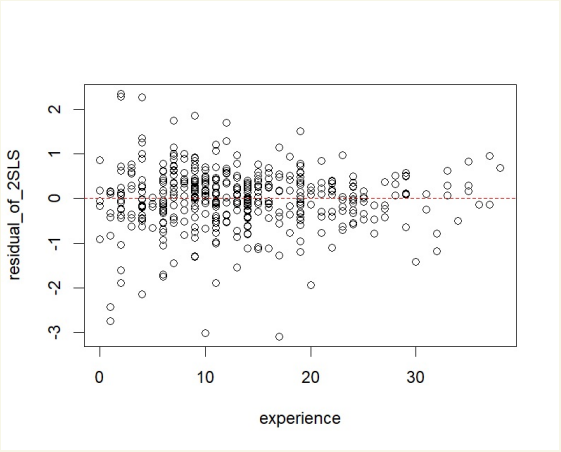
(h)

n	180L
p_sargan	0.454880022254587 > 0.05
R2	0.00310257448755873
sargan_stat	0.558463407760572

do not reject Ho. RANK and POS are both valid IV.

10.24

(a)



It seems no significant patterns ⇒ homoskedasticity holds

(b)

```
> NR2
[1] 7.438552
> p_value
[1] 0.006384122
```

< 0.05 ⇒ reject H_0.
means heteroskedasticity exists.

(c)

```
Robust SE for EDU: 0.03333859
> cat("95% CI for EDU using robust SE: [", lower_robust, ",", upper_robust, "]\n")
95% CI for EDU using robust SE: [ -0.003947005 , 0.1267403 ]
```

(d)

```
Bootstrap SE for EDU: 0.03234547
> cat("95% CI for EDU using bootstrap: [", boot_ci[1], ",", boot_ci[2], "]\n")
95% CI for EDU using bootstrap: [ 0.003650778 , 0.1285871 ]
```