計量作業 0428

Q10.18

(a)

```
> mean(mroz_sub$MOTHERCOLL, na.rm = TRUE)
[1] 0.1214953
> mean(mroz_sub$FATHERCOLL, na.rm = TRUE)
```

(b)

```
> cor(mroz_sub[, c("educ", "MOTHERCOLL", "FATHERCOLL")], use = "complete.obs")
              educ MOTHERCOLL FATHERCOLL
educ       1.0000000  0.3594705  0.3984962
MOTHERCOLL 0.3594705  1.0000000  0.3545709
FATHERCOLL 0.3984962  0.3545709  1.0000000
```

(c)

```
> confint(iv_model, level = 0.95)["educ", ]
       2.5 %        97.5 %
-0.001219763   0.153255678
```

(d)

```
> summary(first_stage)

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

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 ***
exper        0.056230   0.042101   1.336    0.182
I(exper^2)  -0.001956   0.001256  -1.557    0.120
MOTHERCOLL   2.517068   0.315713   7.973 1.46e-14 ***
---
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
```

F value is 63.5631>10 ，fail to reject H0

MOTHERCOOL is a strong instrumental variable

(e)

```
> confint(iv_model_2, level = 0.95)["educ", ]
      2.5 %       97.5 %
0.02751845 0.14817686
```

(f)

```
Call:
lm(formula = educ ~ exper + exper2 + MOTHERCOLL + FATHERCOLL,
    data = mroz_sub)

Residuals:
    Min      1Q  Median      3Q     Max
-7.2152 -0.3056 -0.2152  0.7627  5.0620

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 11.890259   0.290251  40.965  < 2e-16 ***
exper        0.049149   0.040133   1.225    0.221
exper2      -0.001449   0.001199  -1.209    0.227
MOTHERCOLL   1.749947   0.322347   5.429 9.58e-08 ***
FATHERCOLL   2.186612   0.329917   6.628 1.04e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.033 on 423 degrees of freedom
Multiple R-squared:  0.2161,    Adjusted R-squared:  0.2086
F-statistic: 29.15 on 4 and 423 DF,  p-value: < 2.2e-16
```

```
> linearHypothesis(first_stage_2, c("MOTHERCOLL = 0", "FATHERCOLL = 0"))

Linear hypothesis test:
MOTHERCOLL = 0
FATHERCOLL = 0

Model 1: restricted model
Model 2: educ ~ exper + exper2 + MOTHERCOLL + FATHERCOLL

  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1    425 2219.2
2    423 1748.3  2    470.88 56.963 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(g) p-value is 0.6281333，instrument variables do not have any issues in the regression model and there are no overidentifying restrictions

Q10.20

(a)

```
Call:
lm(formula = msft_rp ~ mkt_rp, data = combined_data)

Residuals:
     Min       1Q   Median       3Q      Max
-0.27424 -0.04744 -0.00820  0.03869  0.35801

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.003250   0.006036   0.538    0.591
mkt_rp      1.201840   0.122152   9.839   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.08083 on 178 degrees of freedom
Multiple R-squared:  0.3523,    Adjusted R-squared:  0.3486
F-statistic:  96.8 on 1 and 178 DF,  p-value: < 2.2e-16
```

(b)

```
Call:
lm(formula = mkt_rp ~ RANK, data = combined_data)

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
```

(c)

```
Call:
lm(formula = msft_rp ~ mkt_rp + v_hat, data = combined_data)

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
mkt_rp       1.278318   0.126749  10.085   <2e-16 ***
v_hat       -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

>
> p_value <- coef(summary(mod_aux))["v_hat", "Pr(>|t|)"]
> if(p_value < 0.01){
+   cat("We reject H0, it means market risk premium is endogeneity.\n")
+ } else{
+   cat("We fail to reject H0, it means market risk premium might be exogeneity.\n")
+ }
We fail to reject H0, it means market risk premium might be exogeneity.
```

(d)

```
  Model      Beta        SE         R2
1   OLS 1.201840 0.1221516 0.3522665
2  2SLS 1.278318 0.1280109 0.3508400
```

(e)

```
Linear hypothesis test:
RANK = 0
POS = 0

Model 1: restricted model
Model 2: mkt_rp ~ RANK + POS

  Res.Df      RSS Df Sum of Sq        F    Pr(>F)
1    179 0.43784
2    177 0.03727  2   0.40057 951.26 < 2.2e-16 ***

Call:
lm(formula = mkt_rp ~ RANK + POS, data = combined_data)

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
```

(f)

```
Call:
lm(formula = msft_rp ~ mkt_rp + v_hat, data = combined_data)

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
mkt_rp       1.278318   0.126749  10.085   <2e-16 ***
v_hat       -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
```

(g)

```
Call:
lm(formula = msft_rp ~ mkt_rp, data = combined_data)

Residuals:
     Min       1Q   Median       3Q      Max
-0.27424 -0.04744 -0.00820  0.03869  0.35801

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.003250   0.006036   0.538   0.591
mkt_rp      1.201840   0.122152   9.839   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.08083 on 178 degrees of freedom
Multiple R-squared:  0.3523,    Adjusted R-squared:  0.3486
F-statistic:  96.8 on 1 and 178 DF,  p-value: < 2.2e-16
```

```
Call:
ivreg(formula = msft_rp ~ mkt_rp | RANK + POS, data = combined_data)

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
mkt_rp      1.283118   0.127866  10.035   <2e-16 ***
---
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
```
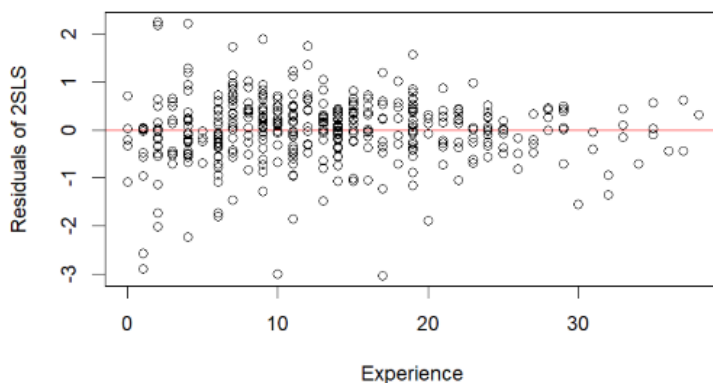
(h)

```
> sargan_stat
[1] 0.5584634
> p_value_sargan
[1] 0.45488
```

Q10.24

(a)

(b)

```
> nr2_test
[1] 8.40036
> p_value
[1] 0.003751468
```

(c)

```
            Estimate Baseline_SE Robust_SE Increased_SE
(Intercept) 0.14784     0.40221   0.42922          Yes
educ        0.06639     0.03125   0.03358          Yes
exper       0.01549     0.00406   0.00414          Yes
> cat("Conclusion:\nRobust SEs are larger, indicating heteroskedasticity.\n")
Conclusion:
Robust SEs are larger, indicating heteroskedasticity.
> cat(sprintf("95%% Robust CI for EDUC: [%.4f, %.4f]\n", ci[1], ci[2]))
95% Robust CI for EDUC: [0.0006, 0.1322]
```

(d)

95% CI for EDUC (bootstrap) = [-0.002 , 0.1248]