

Q18

(a)

```
> mean(mroz_lfp$parentcoll)
[1] 0.1869159
```

約 18.69%的家庭，其父母皆有大學學歷

(b)

mothereduc 和 fathereduc 是連續變數，資訊量較多，但可能與誤差項相關

mothercoll 與 fathercoll 是二元的變數，更可能是外生

```
> cor_matrix
```

	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)

Call:

```
ivreg(formula = log(wage) ~ exper + I(exper^2) + educ | exper +  
      I(exper^2) + mothercoll, data = subset(mroz_lfp, wage > 0))
```

Residuals:

	Min	1Q	Median	3Q	Max
	-3.08719	-0.32444	0.04147	0.36634	2.35621

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.1327561	0.4965325	-0.267	0.78932
exper	0.0433444	0.0134135	3.231	0.00133 **
I(exper^2)	-0.0008711	0.0004017	-2.169	0.03066 *
educ	0.0760180	0.0394077	1.929	0.05440 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6703 on 424 degrees of freedom

Multiple R-Squared: 0.147, Adjusted R-squared: 0.1409

Wald test: 8.2 on 3 and 424 DF, p-value: 2.569e-05

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

(d)

```
call:
lm(formula = educ ~ exper + I(exper^2) + mothercoll, data = mroz_lfp)
```

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 值 > 10，MOTHERCOLL 屬於強工具變數

```
> anova(first_stage)
```

Analysis of Variance Table

Response: educ

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
exper	1	0.52	0.516	0.1133	0.7366
I(exper^2)	1	10.46	10.464	2.2990	0.1302
mothercoll	1	289.32	289.317	63.5631	1.455e-14 ***
Residuals	424	1929.90	4.552		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(e)

```
Call:
ivreg(formula = log(wage) ~ exper + I(exper^2) + educ | exper +
      I(exper^2) + mothercoll + fathercoll, data = subset(mroz_lfp,
      wage > 0))
```

```
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
exper         0.0426761   0.0132950   3.210   0.00143 **
I(exper^2)   -0.0008486   0.0003976  -2.135   0.03337 *
educ         0.0878477   0.0307808   2.854   0.00453 **
```

```
---
```

```
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
```

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

信心區間變窄

(f)

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

```
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
I(exper^2)   -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 + I(exper^2) + 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

MOTHERCOLL 和 FATHERCOLL 都是強工具變數

(g)

Call:

```
lm(formula = resid_iv ~ mothercoll + fathercoll, data = mroz_lfp)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-3.07857	-0.31995	0.03656	0.37417	2.36122

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.000603	0.035313	0.017	0.986
mothercoll	-0.043871	0.105537	-0.416	0.678
fathercoll	0.040464	0.107342	0.377	0.706

Residual standard error: 0.667 on 425 degrees of freedom

Multiple R-squared: 0.0005481, Adjusted R-squared: -0.004155

F-statistic: 0.1165 on 2 and 425 DF, p-value: 0.89

```
> S <- nrow(mroz_lfp) * summary(sargan_test)$r.squared
```

```
> (p_value <- 1 - pchisq(S, df = 1))
```

```
[1] 0.6281333
```

H0: IV 為合適的工具變數，p-value = 0.6281，不拒絕 H0

there is no overidentification.

Q20

(a)

```
Call:
lm(formula = y ~ x, data = data_capm)

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
x            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
```

Beta 約為 1.2，風險略高於大盤

(b)

```
Call:
lm(formula = mkt_excess ~ 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
```

Rank 和 market excess returns 正相關且顯著，F 統計量也高於 10，Rank 屬於強工具變數

(c)

```

Call:
lm(formula = y ~ mkt_excess + 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
mkt_excess   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

```

v_{hat} 的係數在 99%顯著水準下不顯著，無法拒絕 market_excess 為外生變數的假設。但在 95%水準下是顯著，可以拒絕 market_excess 為外生變數

(d)

```

Call:
ivreg(formula = y ~ mkt_excess | RANK, data = capm5)

Residuals:
    Min       1Q   Median       3Q      Max
-0.271625 -0.049675 -0.009693  0.037683  0.355579

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.003018   0.006044   0.499   0.618
mkt_excess   1.278318   0.128011   9.986 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.08092 on 178 degrees of freedom
Multiple R-Squared:  0.3508,    Adjusted R-squared:  0.3472
Wald test: 99.72 on 1 and 178 DF,  p-value: < 2.2e-16

```

係數稍大一點，IV 可能消除了內生性問題，使關係更強烈一些

(e)

```

Call:
lm(formula = mkt_excess ~ 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 和 POS 的聯合顯著性檢驗結果顯示顯著，所以都可以視為強工具變數

(f)

```

Call:
lm(formula = y ~ mkt_excess + v_2, 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
mkt_excess   1.283118   0.126344  10.156  <2e-16 ***
v_2          -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

```

v_2 的係數在 99% 顯著水準下不顯著，無法拒絕 market_excess 為外生變數的假設。但在 95%水準下是顯著的，可以拒絕 market_excess 為外生變數

(g)

```

Call:
ivreg(formula = y ~ mkt_excess | 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
mkt_excess   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

```

係數差異不大，R-square 也相差不多，符合預期

(h)

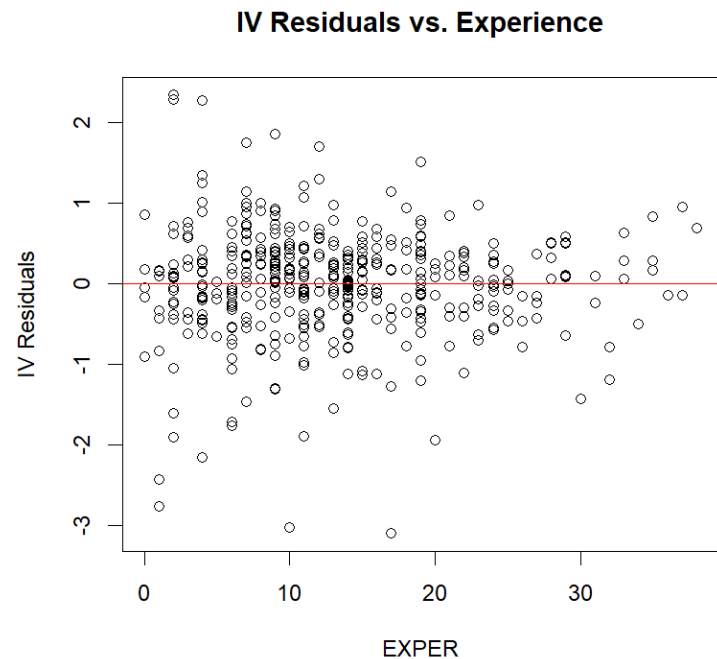
NR-square: 0.5584634

P-value: 0.45488

不拒絕 H_0 ，IV 和誤差沒有相關性

Q24

(a)



When experience is low, the variation in the residuals appears to be larger. This indicates possible heteroskedasticity.

(b)

nR2 = 7.438552 p-value = 0.006384122

Based on the result, the p-value is below 0.01, which provides strong evidence of heteroskedasticity in the model.

(c)

	Estimate	Baseline_SE	Robust_SE	Increased_SE
(Intercept)	0.04810	0.40033	0.42980	Yes
exper	0.04417	0.01343	0.01555	Yes
exper2	-0.00090	0.00040	0.00043	Yes
educ	0.06140	0.03144	0.03334	Yes

EDUC's 95% confidence interval: [-0.003947005 , 0.1267403]

Robust SE is larger than baseline model.

(d)

	Coef	Baseline_SE	Robust_SE	Bootstrap_SE	Larger_than_Baseline_SE
(Intercept)	(Intercept)	0.40033	0.42980	0.42199	Yes
exper	exper	0.01343	0.01555	0.01640	Yes
exper2	exper2	0.00040	0.00043	0.00047	Yes
educ	educ	0.03144	0.03334	0.03291	Yes
	Larger_than_Robust_SE				
(Intercept)		No			
exper		Yes			
exper2		Yes			
educ		No			

Intervals :

Level Normal

95% (-0.0055, 0.1235)