# Chapter 10

18.

(a) MOTHERCOLL = 12.1495 (%)

FATHERCOLL = 11.6822 (%)

(b)

```
                educ MOTHERCOLL FATHERCOLL
educ       1.0000000  0.3594705  0.3984962
MOTHERCOLL 0.3594705  1.0000000  0.3545709
FATHERCOLL 0.3984962  0.3545709  1.0000000
```

→ It shows positive correlation between each variable with medium correlation.

→ It can effectively reduce the measurement error.

(c) 95% C.I. = ( -0.001219763, 0.153255678 )

(d)

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

```
Linear hypothesis test:
MOTHERCOLL = 0

Model 1: restricted model
Model 2: educ ~ MOTHERCOLL + exper + I(exper^2)

  Res.Df    RSS Df Sum of Sq      F    Pr(>F)    
1    425 2219.2                                  
2    424 1929.9  1    289.32 63.563 1.455e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

→ $\widehat{EDUC} = 12.0791 + 0.0562\,EXPER - 0.0020\,EXPER^2 + 2.5171\,MOTHERCOLL + v$

→ ∵ F = 63.563 > $F_{crit}$ = 10

∴ it has enough evidence to show MOTHERCOLL is a strong I.V.

(e) 95% C.I. = ( 0.02751845, 0.14817686 )

→ It is narrower then answer of part (c).

(f)

```
Call:
lm(formula = educ ~ MOTHERCOLL + FATHERCOLL + exper + I(exper^2), 
    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 ***
MOTHERCOLL   1.749947   0.322347   5.429 9.58e-08 ***
FATHERCOLL   2.186612   0.329917   6.628 1.04e-10 ***
exper        0.049149   0.040133   1.225    0.221    
I(exper^2)  -0.001449   0.001199  -1.209    0.227    
---
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
```

```
Linear hypothesis test:
MOTHERCOLL = 0
FATHERCOLL = 0

Model 1: restricted model
Model 2: educ ~ MOTHERCOLL + FATHERCOLL + exper + I(exper^2)

  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
```

→ $\widehat{EDUC} = 11.8903 + 0.0492\,EXPER - 0.0015\,EXPER^2 + 1.7500\,MOTHERCOLL + 2.18661\,FATHERCOLL + v$

→ ∵ F = 56.963 > $F_{crit}$ = 10

∴ it has enough evidence to show FATHERCOLL and MOTHERCOLL are strong IV.

(g) ∵ p-value = 0.626 > 0.05

∴ we fail to reject $H_0$, it doesn't violate the exogenous problem.

20.

(a)

```
Call:
lm(formula = er_msft ~ er_mkt, data = capm5)

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    
er_mkt      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
```

→ $r_j - r_f = 0.0033 + 1.2018(r_m - r_f)$

→ ∵ $\beta = 1.201840 > \beta_{crit} = 1$

∴ Microsoft stock is "risker" than the market portfolio over this sample period.

(b)

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

→ It satisfies IV1 but violates IV2 and IV3 since it is a deterministic monotonic transformation of the endogeneous regressor ex_mkt.

→ The coefficient of RANK is significant since p-value is extremely small.

→ $R^2 = 0.9126$, it is a very strong IV.

(c)

```
Call:
lm(formula = er_msft ~ er_mkt + fsresidual, 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    
er_mkt       1.278318   0.126749  10.085   <2e-16 ***
fsresidual  -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 = 0.0428 > p = 0.01

∴ We fail to reject $H_0$. it does NOT have sufficient evidence to show that $r_m - r_f$ is endogeneous at the 1% level.

(d)

```
Call:
ivreg(formula = er_msft ~ er_mkt | capm5$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    
er_mkt      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
```

→ $r_j - r_f = 0.0030 + 1.2783\,(r_m - r_f)$

→ Yes. $\beta$ is larger, which suggests a stronger relationship between market excess return and Microsoft's.

(e)

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

→ $\because$ $p$-value $< 2.2 \times 10^{-16} < 0.1$

$\therefore$ We have adequately strong IV.

→ $R^2 = 0.9149$

(f)

```
Call:
lm(formula = er_msft ~ er_mkt + re, 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    
er_mkt       1.283118   0.126344  10.156   <2e-16 ***
re          -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
```

$\because$ $p$-value $= 0.0287 > p = 0.01$

$\therefore$ We fail to reject $H_0$. there's no sufficient evidence to show that it is endogeneous.

(g)

```
Call:
ivreg(formula = er_msft ~ er_mkt | capm5$RANK + capm5$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    
er_mkt      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
```

→ With a higher value of $\beta$, it captures some variation that OLS does not catch.

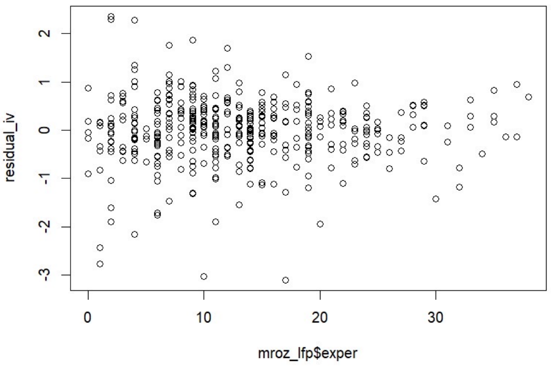
(h) $\because$ p-value $= 0.4548 > p = 0.05$

$\therefore$ fail to reject $H_0$, the extra IV does NOT violate exclusion restriction.

$\longrightarrow$ It is valid.

24.

(a)




$\longrightarrow$ It does not have any visible patterns, shows homoskedasticity holds.

(b) $\because$ p-value $= 0.0064 < p = 0.01$

$\therefore$ We reject $H_0$, the homoskedasticity exists.

(c)

```
t test of coefficients:

               Estimate   Std. Error t value Pr(>|t|)
(Intercept)  0.04810030  0.42979772  0.1119 0.910945
educ         0.06139663  0.03333859  1.8416 0.066231 .
exper        0.04417039  0.01554638  2.8412 0.004711 **
I(exper^2)  -0.00089897  0.00043008 -2.0902 0.037193 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

$\longrightarrow$ 95% baseline C.I. $= (-0.0040, 0.1230)$

95% Robust C.I. $= (0.0039, 0.1267)$

(d) 95% interval estimate $= (-0.002, 0.1248)$

$\longrightarrow$ It is larger than answer in part (c).