10.18 Consider the data file *mroz* on working wives. Use the 428 observations on married women who participate in the labor force. In this exercise, we examine the effectiveness of a parent's college education as an instrumental variable.

- **a.** Create two new variables. *MOTHERCOLL* is a dummy variable equaling one if *MOTHEREDUC* > 12, zero otherwise. Similarly, *FATHERCOLL* equals one if *FATHEREDUC* > 12 and zero otherwise. What percentage of parents have some college education in this sample?
- **b.** Find the correlations between *EDUC*, *MOTHERCOLL*, and *FATHERCOLL*. Are the magnitudes of these correlations important? Can you make a logical argument why *MOTHERCOLL* and *FATHERCOLL* might be better instruments than *MOTHEREDUC* and *FATHEREDUC*?
- **c.** Estimate the wage equation in Example 10.5 using *MOTHERCOLL* as the instrumental variable. What is the 95% interval estimate for the coefficient of *EDUC*?
- **d.** For the problem in part (c), estimate the first-stage equation. What is the value of the *F*-test statistic for the hypothesis that *MOTHERCOLL* has no effect on *EDUC*? Is *MOTHERCOLL* a strong instrument?
- **e.** Estimate the wage equation in Example 10.5 using *MOTHERCOLL* and *FATHERCOLL* as the instrumental variables. What is the 95% interval estimate for the coefficient of *EDUC*? Is it narrower or wider than the one in part (c)?
- **f.** For the problem in part (e), estimate the first-stage equation. Test the joint significance of *MOTHERCOLL* and *FATHERCOLL*. Do these instruments seem adequately strong?
- **g.** For the IV estimation in part (e), test the validity of the surplus instrument. What do you conclude?

(a) $MOTHERCOLL = \begin{cases} 1, & MOTHEREDUC > 12 \\ 0, & o.w. \end{cases}$

$FATHERCOLL = \begin{cases} 1, & FATHEREDUC > 12 \\ 0, & o.w. \end{cases}$

```
>
> # 計算有大學教育的父母百分比
> mean(data$MOTHERCOLL, na.rm = TRUE)
[1] 0.1009296
> mean(data$FATHERCOLL, na.rm = TRUE)
[1] 0.1075697
```

(b)

```
> #(b)
> cor(data[, c("EDUC", "MOTHERCOLL", "FATHEI
                EDUC MOTHERCOLL FATHERCOLL
EDUC       1.0000000  0.3370171  0.3193212
MOTHERCOLL 0.3370171  1.0000000  0.3674532
FATHERCOLL 0.3193212  0.3674532  1.0000000
```

corr(EDUC, MOTHERCOLL) = 0.3370171

corr(EDUC, FATHERCOLL) = 0.3193212

使用MOTHERCOLL和FATHERCOLL取代MOTHEREDUC和FATHEREDUC可使外生性更加成立，因為MOTHEREDUC反映了父母教養風格、價值觀等因素，被包含在誤差項裡，可能產生內生性。

(c)

```
Call:
ivreg(formula = log(WAGE) ~ EDUC + EXPER + EXPER2 | MOTHERCOLL +
    EXPER + EXPER2, data = data_clean)

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
EDUC         0.0760180  0.0394077   1.929  0.05440 .
EXPER        0.0433444  0.0134135   3.231  0.00133 **
EXPER2      -0.0008711  0.0004017  -2.169  0.03066 *
---
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, "EDUC", level = 0.95)
            2.5 %    97.5 %
EDUC -0.001219763 0.1532557
```

(d)

```
Call:
lm(formula = EDUC ~ MOTHERCOLL + EXPER + EXPER2, data = data)

Residuals:
     Min       1Q   Median       3Q      Max
-7.3343 -0.4752 -0.1441  0.6550  5.4277

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 11.475204   0.175367  65.435  < 2e-16 ***
MOTHERCOLL   2.140528   0.257979   8.848  < 2e-16 ***
EXPER        0.099989   0.027719   3.607  0.00033 ***
EXPER2      -0.002852   0.000896  -3.183  0.00152 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.132 on 749 degrees of freedom
Multiple R-squared:  0.1291,    Adjusted R-squared:  0.1256
F-statistic: 37.01 on 3 and 749 DF,  p-value: < 2.2e-16

> f_test <- summary(first_stage_d)$fstatistic
> cat("F-statistic for MOTHERCOLL:", f_test[1], "\n")
F-statistic for MOTHERCOLL: 37.00737
```

(e)

```
Call:
ivreg(formula = log(WAGE) ~ EDUC + EXPER + EXPER2 | MOTHERCOLL +
    FATHERCOLL + EXPER + EXPER2, data = data_clean)

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 **
EXPER2      -0.0008486  0.0003976  -2.135  0.03337 *
---
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_model2, "EDUC", level = 0.95)
            2.5 %    97.5 %
EDUC 0.02751845 0.1481769
```

narrower than (c)

(f)

```
Call:
lm(formula = EDUC ~ MOTHERCOLL + FATHERCOLL, data = data)

Residuals:
     Min       1Q   Median       3Q      Max
-6.9142 -0.9142  0.0858  0.1646  5.0858

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 11.91415    0.08261 144.223  < 2e-16 ***
MOTHERCOLL   1.92121    0.27257   7.049 4.10e-12 ***
FATHERCOLL   1.66213    0.26500   6.272 6.01e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.095 on 750 degrees of freedom
Multiple R-squared:  0.1578,    Adjusted R-squared:  0.1555
F-statistic: 70.24 on 2 and 750 DF,  p-value: < 2.2e-16

> anova(first_stage_f)  # 可用於聯合檢定
Analysis of Variance Table

Response: EDUC
             Df Sum Sq Mean Sq F value    Pr(>F)
MOTHERCOLL    1  444.1  444.10  101.14 < 2.2e-16 ***
FATHERCOLL    1  172.7  172.74   39.34 6.011e-10 ***
Residuals   750 3293.2    4.39
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

adequately strong

(g)

	df1	df2	statistic	p-value
Weak instruments	2	423	56.9628997	1.241993e-22
Wu-Hausman	1	423	0.5186358	4.718220e-01
Sargan	1	NA	0.2375851	6.259557e-01

6.259557e-01 > 0.05

do not reject H_0 (EDUC是外生)

do not reject H_0 (IV有效)