

## 15.6

Using the NLS panel data on  $N = 716$  young women, we consider only years 1987 and 1988. We are interested in the relationship between  $\ln(WAGE)$  and experience, its square, and indicator variables for living in the south and union membership. Some estimation results are in Table 15.10.

TABLE 15.10 Estimation Results for Exercise 15.6					
	(1) OLS 1987	(2) OLS 1988	(3) FE	(4) FE Robust	(5) RE
$C$	0.9348 (0.2010)	0.8993 (0.2407)	1.5468 (0.2522)	1.5468 (0.2688)	1.1497 (0.1597)
$EXPER$	0.1270 (0.0295)	0.1265 (0.0323)	0.0575 (0.0330)	0.0575 (0.0328)	0.0986 (0.0220)
$EXPER^2$	-0.0033 (0.0011)	-0.0031 (0.0011)	-0.0012 (0.0011)	-0.0012 (0.0011)	-0.0023 (0.0007)
$SOUTH$	-0.2128 (0.0338)	-0.2384 (0.0344)	-0.3261 (0.1258)	-0.3261 (0.2495)	-0.2326 (0.0317)
$UNION$	0.1445 (0.0382)	0.1102 (0.0387)	0.0822 (0.0312)	0.0822 (0.0367)	0.1027 (0.0245)
$N$	716	716	1432	1432	1432

(standard errors in parentheses)

- a. The OLS estimates of the  $\ln(WAGE)$  model for each of the years 1987 and 1988 are reported in columns (1) and (2). How do the results compare? For these individual year estimations, what are you assuming about the regression parameter values across individuals (heterogeneity)?

Ans.

兩年 OLS 的結果相當接近，各變數的係數估計值與標準誤差變化不大。

OLS 模型假設所有個體的母體參數值（包括截距）皆相同，忽略了異質性。

- b. The  $\ln(WAGE)$  equation specified as a panel data regression model is

$$\ln(WAGE_{it}) = \beta_1 + \beta_2 EXPER_{it} + \beta_3 EXPER_{it}^2 + \beta_4 SOUTH_{it} + \beta_5 UNION_{it} + (u_i + e_{it}) \quad (XR15.6)$$

Explain any differences in assumptions between this model and the models in part (a).

Ans.

變數上加上個體與時間的下標  $it$ ，有 2 個誤差項  $(u_i, e_{it})$ 。

$e_{it}$  表示在個體與時間之間變動的誤差項； $u_i$  是不隨時間改變的誤差項，代表各個體之間的不可觀察異質性。

- c. Column (3) contains the estimated fixed effects model specified in part (b). Compare these estimates with the OLS estimates. Which coefficients, apart from the intercepts, show the most difference?

Ans.

Fixed Effects 的 95% 信賴區間：EXPER  $[-0.0085, 0.1235]$ 、EXPER<sup>2</sup>  $[-0.0034, 0.001]$ 、SOUTH  $[-0.5777, -0.0745]$ 、UNION  $[0.0198, 0.1446]$

EXPER 的 OLS 的估計值 0.1270 沒有落在 Fixed Effects 的信賴區間  $[-0.0085, 0.1235]$ ，OLS 和 Fixed Effects 模型的結果有顯著差異。

其他變數 (EXPER<sup>2</sup>、SOUTH、UNION) 的 OLS 的估計值都有落在 Fixed Effects 的信賴區間，這幾個變數差異不大，兩種方法的結果大致一致。

- d. The  $F$ -statistic for the null hypothesis that there are no individual differences, equation (15.20), is 11.68. What are the degrees of freedom of the  $F$ -distribution if the null hypothesis (15.19) is true? What is the 1% level of significance critical value for the test? What do you conclude about the null hypothesis.

Ans.

$$H_0 : \beta_{11} = \beta_{12} = \beta_{13} = \dots = \beta_{1N}$$

$$H_1 : \text{the } \beta_{1i} \text{ are not all equal}$$

$$F = \frac{(SSE_R - SSE_U)/(N - 1)}{SSE_U/(NT - N - K)}$$

> qf(p = 0.99, df1 = 715, df2 = 712)  
[1] 1.190496

$N=716$  是個體數、 $T=2$  是時間期數、 $NT=1432$  是總觀測值、 $K=5$  是 XR15.6 式中的解釋變數個數（不含截距）。分子自由度 =  $N-1 = 716-1 = 715$ 、分母自由度 =  $NT-N-K = 1432-716-4 = 712$

因為  $F = 11.68 >>$  臨界值 1.19，所以我們拒絕虛無假設  $H_0$ ：即拒絕「沒有個體差異」的假設。

這表示固定效果是重要的，存在顯著的個體異質性。

- e. Column (4) contains the fixed effects estimates with cluster-robust standard errors. In the context of this sample, explain the different assumptions you are making when you estimate with and without cluster-robust standard errors. Compare the standard errors with those in column (3). Which ones are substantially different? Are the robust ones larger or smaller?

Ans.

在 within transformation 後，隨機誤差會變成  $\tilde{e}_{it} = e_{it} - \bar{e}_i$ 。雖然原本的誤差  $e_{it}$  是不相關的，但經過變換後，新的誤差項在同一個體內會變得「時間上相關」，即有自我相關性。

透過使用 cluster robust SE，我們容許誤差項  $e_{it}$  在「不同個體間」、「不同時間點」存在異質變異或時間上的自相關。

比較表格中 FE (欄3) 和 FE Robust (欄4) 的標準誤

EXPER、EXPER<sup>2</sup> → Robust 跟傳統差不多，代表估計穩定。

SOUTH → Robust 大很多，代表傳統標準誤嚴重低估，使用 robust 更可靠。

UNION → Robust大一點，代表傳統標準誤也有高估，但影響較小。

FE Robust 與 FE 的估計值相同，但標準誤考慮了 Clustered Errors 與異質變異數問題，使得 SOUTH 與 UNION 的標準誤變大，檢定更可靠。