11.30

Example 11.3 introduces Klein's Model I. Use the data file $klein$ to answer the following questions.

a. Estimate the investment function in equation (11.18) by OLS. Comment on the signs and significance of the coefficients.

Ans.

$\hat{I}_t = 10.126 + 0.480P_t + 0.333P_{t-1} - 0.112K_{t-1}$

P_t：正向且顯著，表示公司當期利潤高，有更多自有資金與投資誘因，會增加投資。

P_{t-1}：正向且顯著，前期利潤佳也會影響當期投資，反映投資決策的延遲或調整行為。

K_{t-1}：負向且顯著，符合邊際資本遞減的理論預期。前期資本存量愈高，表示企業設備可能已充足，減少新增投資需求。

所有變數在 1% 或 5% 水準下皆具統計顯著性（p-value < 0.05）。

```
Call:
lm(formula = i ~ p + plag + klag, data = klein)

Residuals:
     Min       1Q   Median       3Q      Max
-2.56562 -0.63169  0.03687  0.41542  1.49226

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 10.12579    5.46555   1.853 0.081374 .
p            0.47964    0.09711   4.939 0.000125 ***
plag         0.33304    0.10086   3.302 0.004212 **
klag        -0.11179    0.02673  -4.183 0.000624 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.009 on 17 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.9313,    Adjusted R-squared:  0.9192
F-statistic: 76.88 on 3 and 17 DF,  p-value: 4.299e-10
```

b. Estimate the reduced-form equation for profits, P_t, using all eight exogenous and predetermined variables as explanatory variables. Test the joint significance of all the variables except lagged profits, P_{t-1}, and lagged capital stock, K_{t-1}. Save the residuals, $\hat{v}_t$ and compute the fitted values, $\hat{P}_t$.

Ans.

```
Call:
lm(formula = p ~ g + w2 + tx + time + plag + klag + elag, data = klein)

Residuals:
    Min      1Q  Median      3Q     Max
-3.9067 -1.3050  0.3226  1.3613  2.8881

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 50.38442   31.63026   1.593   0.1352
g            0.43902    0.39114   1.122   0.2820
w2          -0.07961    2.53382  -0.031   0.9754
tx          -0.92310    0.43376  -2.128   0.0530 .
time         0.31941    0.77813   0.410   0.6881
plag         0.80250    0.51886   1.547   0.1459
klag        -0.21610    0.11911  -1.814   0.0928 .
elag         0.02200    0.28216   0.078   0.9390
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.183 on 13 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.8261,    Adjusted R-squared:  0.7324
F-statistic: 8.821 on 7 and 13 DF,  p-value: 0.0004481
```

$H_0: \beta_g = \beta_{w_2} = \beta_{tx} = \beta_{time} = \beta_{elag} = 0$　　H_1：至少一個不為 0（聯合顯著）

$\alpha = 0.05$　　F(5, 13) = 3.0254

F = 1.9345 小於 臨界值 3.0254，因此 無法拒絕虛無假設。

代表這些變數（g, w2, tx, time, elag）無法共同有效解釋利潤 P_t，可能存在工具變數較弱的問題。

換句話說，它們不是一組強工具變數（strong instruments）

```
> p_hat
       2        3        4        5        6        7        8        9       10
13.255556 16.577368 19.282347 20.960143 19.766509 18.238731 17.573065 19.541720 20.375101
       11       12       13       14       15       16       17       18       19
17.180415 12.705026  8.999780  9.054102 12.671263 14.421338 14.711907 19.796405 19.206691
       20       21       22
17.419605 20.305654 22.657273
```

c. The Hausman test for the presence of endogenous explanatory variables is discussed in Section 10.4.1. It is implemented by adding the reduced–form residuals to the structural equation and testing their significance, that is, using OLS estimate the model

$$I_t = \beta_1 + \beta_2 P_t + \beta_3 P_{t-1} + \beta_4 K_{t-1} + \delta \hat{v}_t + e_{2t}$$

Use a t–test for the null hypothesis $H_0: \delta = 0$ versus $H_1: \delta \neq 0$ at the 5% level of significance. By rejecting the null hypothesis, we conclude that P_t is endogenous. What do we conclude from the test? In the context of this simultaneous equations model what result should we find?

Ans.

```
Call:
lm(formula = i ~ p + plag + klag + v_hat, data = klein)

Residuals:
     Min       1Q   Median       3Q      Max
-1.04645 -0.56030  0.06189  0.25348  1.36700

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 20.27821    4.70179   4.313 0.000536 ***
p            0.15022    0.10798   1.391 0.183222
plag         0.61594    0.10147   6.070 1.62e-05 ***
klag        -0.15779    0.02252  -7.007 2.96e-06 ***
v_hat        0.57451    0.14261   4.029 0.000972 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7331 on 16 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.9659,    Adjusted R-squared:  0.9574
F-statistic: 113.4 on 4 and 16 DF,  p-value: 1.588e-11
```

v_hat 的係數為 0.57451，t 值 = 4.029，p 值 = 0.000972 < 0.05。

在 5% 顯著水準下，拒絕虛無假設 $H_0: \delta = 0$。

我們有統計證據顯示殘差 $\hat{v}_t$ 對投資 I_t 有顯著影響。根據 Hausman 檢定，這表示利潤 P_t 是一個 內生變數，與結構方程誤差項 e_{2t} 有相關。因此，使用 OLS 估計投資方程將產生 偏誤且不一致的估計值。應採用 2SLS 或其他工具變數法對模型進行修正。

d. Obtain the 2SLS estimates of the investment equation using all eight exogenous and predetermined variables as IVs and software designed for 2SLS. Compare the estimates to the OLS estimates in part (a). Do you find any important differences?

Ans.

內生變數 P_t 的差異很大，OLS 認為 P_t 對投資有顯著正向影響；但 2SLS 結果顯示 P_t 的估計值大幅縮小且不顯著，顯示 OLS 存在上偏（upward bias）。

此差異與 11.30(c) 的 Hausman test 相符，說明 OLS 對 P_t 的估計有偏誤。

其他變數（如P_{t-1}、K_{t-1}）的估計值雖略有變動，但方向一致且皆顯著，顯示解釋力穩定。

```
Call:
ivreg(formula = i ~ p + plag + klag | g + w2 + tx + time + elag +
    plag + klag, data = klein)

Residuals:
    Min      1Q  Median      3Q     Max
-3.2909 -0.8069  0.1423  0.8601  1.7956

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 20.27821    8.38325   2.419  0.02707 *
p            0.15022    0.19253   0.780  0.44598
plag         0.61594    0.18093   3.404  0.00338 **
klag        -0.15779    0.04015  -3.930  0.00108 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.307 on 17 degrees of freedom
Multiple R-Squared: 0.8849,     Adjusted R-squared: 0.8646
Wald test:  41.2 on 3 and 17 DF,  p-value: 5.148e-08
```

```
> print(comparison_table)
            OLS_Estimate OLS_SE   OLS_t     OLS_p IV_Estimate  IV_SE    IV_t     IV_p
(Intercept)      10.1258 5.4655  1.8527 0.0813700     20.2782 8.3832  2.4189 0.027070
p                 0.4796 0.0971  4.9389 0.0001246      0.1502 0.1925  0.7802 0.446000
plag              0.3330 0.1009  3.3020 0.0042120      0.6159 0.1809  3.4044 0.003375
klag             -0.1118 0.0267 -4.1827 0.0006244     -0.1578 0.0402 -3.9298 0.001080
```

e. **Estimate the second-stage model $I_t = \beta_1 + \beta_2 \hat{P}_t + \beta_3 P_{t-1} + \beta_4 K_{t-1} + e_{2t}$ by OLS. Compare the estimates and standard errors from this estimation to those in part (d). What differences are there?**

Ans.

```
Call:
lm(formula = i ~ p_hat + plag + klag, data = klein)

Residuals:
    Min      1Q  Median      3Q     Max
-3.8778 -1.0029  0.3058  0.7275  2.1831

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 20.27821    9.97663   2.033  0.05802 .
p_hat        0.15022    0.22913   0.656  0.52084
plag         0.61594    0.21531   2.861  0.01083 *
klag        -0.15779    0.04778  -3.302  0.00421 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.556 on 17 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.837,     Adjusted R-squared:  0.8082
F-statistic: 29.09 on 3 and 17 DF,  p-value: 6.393e-07
```

估計值完全一致。

標準誤與 p 值不同：ivreg() 正確地考慮了 第一階段的不確定性（標準誤擴大）；手動 OLS 僅是單純替代變數法 → 低估了變異數，標準誤偏小，p 值偏低

f. **Let the 2SLS residuals from part (e) be $\hat{e}_{2t}$. Regress these residuals on all the exogenous and predetermined variables. If these instruments are valid, then the R^2 from this regression should be low, and none of the variables are statistically significant. The Sargan test for instrument validity is discussed in Section 10.4.3. The test statistic TR^2 has a chi-square distribution with degrees of freedom equal to the number of "surplus" IVs if the surplus instruments are valid. The investment equation includes three exogenous and/or predetermined variables out of the total of eight possible. There are $L = 5$ external instruments and $B = 1$ right-hand side endogenous variables. Compare the value of the test statistic to the 95th percentile value from the $\chi^2_{(4)}$ distribution. What do we conclude about the validity of the surplus instruments in this case?**

Ans.

H_0：所有的工具變數（IVs）都是外生的（有效的）　　H_1：至少有一個工具變數是內生的（無效的）

Sargan 統計量（TR^2）：1.3425 < 臨界值（95%）$\chi^2_{(4)}$ = 9.4877

根據 Sargan 檢定，我們沒有證據拒絕工具變數外生性的假設，因此可接受這些工具變數作為有效的 IV。換言之，過度識別條件成立，2SLS 的估計結果可信。

```
Call:
lm(formula = e2_hat ~ g + w2 + tx + time + elag + plag + klag,
    data = klein)

Residuals:
    Min      1Q  Median      3Q     Max
-3.4087 -0.8799  0.2702  1.0011  2.4987

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.671103  24.976416   0.307    0.764
g            0.034277   0.308861   0.111    0.913
w2          -0.704649   2.000800  -0.352    0.730
tx          -0.022846   0.342512  -0.067    0.948
time         0.283921   0.614439   0.462    0.652
elag        -0.116046   0.222807  -0.521    0.611
plag         0.189896   0.409708   0.463    0.651
klag        -0.002262   0.094056  -0.024    0.981

Residual standard error: 1.724 on 13 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.06102,   Adjusted R-squared:  -0.4446
F-statistic: 0.1207 on 7 and 13 DF,  p-value: 0.9953
```

```
Sargan test statistic (TR^2): 1.342544
 df = 4
 p-value: 0.8541172
 Critical value (chi^2(4), 95%):  9.487729
```