11.1 Our aim is to estimate the parameters of the simultaneous equations model

$$y_1 = \alpha_1 y_2 + e_1$$

$$y_2 = \alpha_2 y_1 + \beta_1 x_1 + \beta_2 x_2 + e_2$$

We assume that x_1 and x_2 are exogenous and uncorrelated with the error terms e_1 and e_2.

a. Solve the two structural equations for the reduced-form equation for y_2, that is, $y_2 = \pi_1 x_1 + \pi_2 x_2 + v_2$. Express the reduced-form parameters in terms of the **structural parameters** and the reduced-form error in terms of the structural parameters and e_1 and e_2. Show that y_2 is correlated with e_1.

b. Which equation parameters are consistently estimated using OLS? Explain.

c. Which parameters are "identified," in the simultaneous equations sense? Explain your reasoning.

11.1

(a) $y_2 = \alpha_2(\alpha_1 y_2 + e_1) + \beta_1 x_1 + \beta_2 x_2 + e_2$

$= \beta_1 x_1 + \beta_2 x_2 + (\alpha_1\alpha_2 y_2 + \alpha_2 e_1 + e_2)$

$(1-\alpha_1\alpha_2) y_2 = \beta_1 x_1 + \beta_2 x_2 + \alpha_2 e_1 + e_2$

$y_2 = \frac{\beta_1}{1-\alpha_1\alpha_2} x_1 + \frac{\beta_2}{1-\alpha_1\alpha_2} x_2 + \frac{(\alpha_2 e_1 + e_2)}{1-\alpha_1\alpha_2}$

$\pi_1 = \frac{\beta_1}{1-\alpha_1\alpha_2}, \quad \pi_2 = \frac{\beta_2}{1-\alpha_1\alpha_2}, \quad v_2 = \frac{\alpha_2 e_1 + e_2}{1-\alpha_1\alpha_2}$

$E[(y_2 - E(y_2))(e_1 - E(e_1))]$

$= E(y_2 e_1)$

$= \frac{\beta_1}{1-\alpha_1\alpha_2} \cancel{E(x_1 e_1)}^{0} + \frac{\beta_2}{1-\alpha_1\alpha_2} \cancel{E(x_2 e_1)}^{0} + \frac{\alpha_2}{1-\alpha_1\alpha_2} E(e_1^2) + \frac{1}{1-\alpha_1\alpha_2} \cancel{E(e_1 e_2)}^{0}$

$= \frac{\alpha_2}{1-\alpha_1\alpha_2} E(e_1^2) \neq 0$

(b) 都無法用 OLS consistently estimate
2條方程式都有 endogenous variable

(c) $M=2$ 至少要 omit $M-1=1$ 個 variable
equation 1 omit 2 個
equation 2 omit 0 個
α_1 indentified

d. To estimate the parameters of the reduced-form equation for y_2 using the method of moments (MOM), which was introduced in Section 10.3, the two moment equations are

$$N^{-1}\sum x_{i1}\left(y_2 - \pi_1 x_{i1} - \pi_2 x_{i2}\right) = 0$$

$$N^{-1}\sum x_{i2}\left(y_2 - \pi_1 x_{i1} - \pi_2 x_{i2}\right) = 0$$

Explain why these two moment conditions are a valid basis for obtaining consistent estimators of the reduced-form parameters.

e. Are the MOM estimators in part (d) the same as the OLS estimators? Form the sum of squared errors function for $y_2 = \pi_1 x_1 + \pi_2 x_2 + v_2$ and find the first derivatives. Set these to zero and show that they are equivalent to the two equations in part (d).

f. Using $\sum x_{i1}^2 = 1$, $\sum x_{i2}^2 = 1$, $\sum x_{i1}x_{i2} = 0$, $\sum x_{i1}y_{1i} = 2$, $\sum x_{i1}y_{2i} = 3$, $\sum x_{i2}y_{1i} = 3$, $\sum x_{i2}y_{2i} = 4$, and the two moment conditions in part (d) show that the MOM/OLS estimates of π_1 and π_2 are $\hat{\pi}_1 = 3$ and $\hat{\pi}_2 = 4$.

g. The fitted value $\hat{y}_2 = \hat{\pi}_1 x_1 + \hat{\pi}_2 x_2$. Explain why we can use the moment condition $\sum \hat{y}_{i2}\left(y_{i1} - \alpha_1 y_{i2}\right) = 0$ as a valid basis for consistently estimating α_1. Obtain the IV estimate of α_1.

h. Find the 2SLS estimate of α_1 by applying OLS to $y_1 = \alpha_1 \hat{y}_2 + e_1^*$. Compare your answer to that in part (g).

(d) $E(X_1V_1)=0$, $E(X_2V_2)=0 \Rightarrow$ exogenous
2個未知parameter
需2個equation

(e) $SSE = \Sigma(y_2 - \pi_1 X_1 - \pi_2 X_2)^2$

$\frac{\partial SSE}{\partial \pi_1} = 2\Sigma(y_2 - \pi_1 X_1 - \pi_2 X_2)\cdot(-X_1) = 0$

$\frac{\partial SSE}{\partial \pi_2} = 2\Sigma(y_2 - \pi_1 X_1 - \pi_2 X_2)\cdot(-X_2) = 0$

與MOM相同parameter解

(f) $3-\pi_1=0$, $4-\pi_2=0$, $\hat{\pi}_1=3$, $\hat{\pi}_2=4$

(g) X_1, X_2 exogenous, y_2的IV
$\hat{y}_2$ uncorrelate with $e_1 \Rightarrow \alpha_1$ consistently estimated
$E(\hat{y}_2 e_1)=0$

$\Sigma \hat{y}_{i2}y_{i1} - \alpha_1 \Sigma \hat{y}_{i2}^2 = 0$

$\alpha_1 = \frac{\Sigma \hat{y}_{i2}y_{i1}}{\Sigma \hat{y}_{i2}^2} = \frac{3\cdot 2+4\cdot 3}{9\cdot 1+12\cdot 0+16\cdot 1} = \frac{18}{25} = 0.72$

(h) $\Sigma(y_1 - \alpha_1\hat{y}_2)^2 = SSE$

$2\Sigma(y_1-\alpha_1\hat{y}_2)\cdot(-\hat{y}_2)=0$

$\Sigma y_1\hat{y}_2 - \alpha_1\Sigma\hat{y}_2^2 = 0$

$\alpha_1 = \frac{\Sigma y_1\hat{y}_2}{\Sigma \hat{y}_2^2} = 0.72$

same as part (g)

11.16 Consider the following supply and demand model

$$\text{Demand: } Q_i = \alpha_1 + \alpha_2 P_i + e_{di}, \quad \text{Supply: } Q_i = \beta_1 + \beta_2 P_i + \beta_3 W_i + e_{si}$$

where Q is the quantity, P is the price, and W is the wage rate, which is assumed exogenous. Data on these variables are in Table 11.7.

TABLE 11.7 **Data for Exercise 11.16**

Q	P	W
4	2	2
6	4	3
9	3	1
3	5	1
8	8	3

a. Derive the algebraic form of the reduced-form equations, $Q = \theta_1 + \theta_2 W + v_2$ and $P = \pi_1 + \pi_2 W + v_1$, expressing the reduced-form parameters in terms of the structural parameters.

b. Which structural parameters can you solve for from the results in part (a)? Which equation is "identified"?

c. The estimated reduced-form equations are $\hat{Q} = 5 + 0.5W$ and $\hat{P} = 2.4 + 1W$. Solve for the identified structural parameters. This is the method of **indirect least squares**.

d. Obtain the fitted values from the reduced-form equation for P, and apply 2SLS to obtain estimates of the demand equation.

11.16

(a)

$\alpha_1 + \alpha_2 P_i + e_{di} = \beta_1 + \beta_2 P_i + \beta_3 W_i + e_{si}$

$(\alpha_2 - \beta_2) P_i = (\beta_1 - \alpha_1) + \beta_3 W_i + (e_{si} - e_{di})$

$P_i = \left(\frac{\beta_1 - \alpha_1}{\alpha_2 - \beta_2}\right) + \left(\frac{\beta_3}{\alpha_2 - \beta_2}\right) W_i + \frac{(e_{si} - e_{di})}{\alpha_2 - \beta_2}$

$Q_i = \alpha_1 + \alpha_2 P_i + e_{di}$

$= \left(\alpha_1 + \frac{\alpha_2(\beta_1 - \alpha_1)}{\alpha_2 - \beta_2}\right) + \left(\frac{\alpha_2 \beta_3}{\alpha_2 - \beta_2}\right) W_i$

$+ \left(\frac{\alpha_2(e_{si} - e_{di})}{\alpha_2 - \beta_2} + e_{di}\right)$

$\theta_1 = \frac{\alpha_2 \beta_1 - \alpha_1 \beta_2}{\alpha_2 - \beta_2}, \quad \theta_2 = \frac{\alpha_2 \beta_3}{\alpha_2 - \beta_2}$

$V_2 = \frac{\alpha_2 e_{si} - \beta_2 e_{di}}{\alpha_2 - \beta_2}$

$\pi_1 = \frac{\beta_1 - \alpha_1}{\alpha_2 - \beta_2}, \quad \pi_2 = \frac{\beta_3}{\alpha_2 - \beta_2}$

$V_1 = \frac{e_{si} - e_{di}}{\alpha_2 - \beta_2}$

(b) $M = 2$

Demand equation is identified

omit 1 variable W.

(c) $5 + 0.5W = \alpha_1 + 2.4\alpha_2 + \alpha_2 W$

$\alpha_2 = 0.5$

$\alpha_1 = 5 - 1.2 = 3.8$

(d) $Q = \alpha_1 + \alpha_2 \hat{P} + e_i$

	$\hat{P}$	$\hat{Q}$	$\hat{P}^2$	$Q\hat{P}$
	4.4	6	19.36	26.4
	5.4	6.5	29.16	35.1
	3.4	5.5	11.56	18.7
	3.4	5.5	11.56	18.7
	5.4	6.5	29.16	35.1
sum	22	30	100.8	134

$30 - 5\alpha_1 - 22\alpha_2 = 0$

$134 - 22\alpha_1 - 100.8\alpha_2 = 0$

$660 - 110\alpha_1 - 484\alpha_2 = 0$

$670 - 110\alpha_1 - 504\alpha_2 = 0$

$10 + 20\alpha_2 = 0 \Rightarrow \alpha_2 = 0.5$

$\alpha_1 = \frac{19}{5} = 3.8$

11.17 Example 11.3 introduces Klein's Model I.

- **a.** Do we have an adequate number of IVs to estimate each equation? Check the necessary condition for the identification of each equation. The necessary condition for identification is that in a system of M equations at least $M - 1$ variables must be omitted from each equation.
- **b.** An equivalent identification condition is that the number of excluded exogenous variables from the equation must be at least as large as the number of included right-hand side endogenous variables. Check that this condition is satisfied for each equation.
- **c.** Write down in econometric notation the first-stage equation, the reduced form, for W_{1t}, wages of workers earned in the private sector. Call the parameters $\pi_1, \pi_2, \ldots$

11.17

(a)

$CN_t = \alpha_1 + \alpha_2 (W_{1t} + W_{2t}) + \alpha_3 P_t + \alpha_4 P_{t-1} + e_{1t}$

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

$W_{1t} = \gamma_1 + \gamma_2 E_t + \gamma_3 E_{t-1} + \gamma_4 TIME_t + e_{3t}$

$M = 3$ at least omit 2 explanatory variables

Equation 1 omits 4 ⇒ identified

Equation 2 omits 3 ⇒ identified

Equation 3 omits 1 ⇒ underidentified

(b)

Equation 1: 4 ≥ 4 ⇒ equivalent identification

Equation 2: 3 ≥ 2 ⇒ equivalent identification

Equation 3: 1 ≥ 0 ⇒ equivalent identification

(c) $W_{1t} = \pi_1 + \pi_2 K_{t-1} + \pi_3 E_t + \pi_4 E_{t-1} + \pi_5 TIME_t$

(d) ① find as for endogenous variables

② plug in the endogenous estimation back to original equation

③ conduct second OLS

(e) Yes, the estimates would be the same. but the t-value would be different.

11.28 Supply and demand curves as traditionally drawn in economics principles classes have price (P) on the vertical axis and quantity (Q) on the horizontal axis.

a. Rewrite the truffle demand and supply equations in (11.11) and (11.12) with price P on the left-hand side. What are the anticipated signs of the parameters in this rewritten system of equations?

b. Using the data in the file *truffles*, estimate the supply and demand equations that you have formulated in (a) using two-stage least squares. Are the signs correct? Are the estimated coefficients significantly different from zero?

c. Estimate the price elasticity of demand "at the means" using the results from (b).

d. Accurately sketch the supply and demand equations, with P on the vertical axis and Q on the horizontal axis, using the estimates from part (b). For these sketches set the values of the exogenous variables DI, PS, and PF to be $DI^* = 3.5$, $PF^* = 23$, and $PS^* = 22$.

e. What are the equilibrium values of P and Q obtained in part (d)? Calculate the predicted equilibrium values of P and Q using the estimated reduced-form equations from Table 11.2, using the same values of the exogenous variables. How well do they agree?

f. Estimate the supply and demand equations that you have formulated in (a) using OLS. Are the signs correct? Are the estimated coefficients significantly different from zero? Compare the results to those in part (b).

a.

11.28.

Demand

$$Q_i = \alpha_1 + \alpha_2 P_i + \alpha_3 PS_i + \alpha_4 DI_i + e_{di}$$

$$P_i = -\frac{\alpha_1}{\alpha_2} + \frac{1}{\alpha_2} Q_i - \frac{\alpha_3}{\alpha_2} PS_i - \frac{\alpha_4}{\alpha_2} DI_i - \frac{e_{di}}{\alpha_2}$$

Supply

$$Q_i = \beta_1 + \beta_2 P_i + \beta_3 PF_i + e_{si}$$

$$P_i = -\frac{\beta_1}{\beta_2} + \frac{1}{\beta_2} Q_i - \frac{\beta_3}{\beta_2} PF_i - \frac{e_{si}}{\beta_2}$$

Anticipated sign:

Demand equation: q(-), ps(+), di(+)

Supply equation: q(+), pf(+)

b.

```
2SLS estimates for 'eq1' (equation 1)
Model Formula: p ~ q + ps + di
Instruments: ~ps + di + pf

              Estimate Std. Error  t value   Pr(>|t|)
(Intercept) -11.42841   13.59161 -0.84084  0.4081026
q            -2.67052    1.17495 -2.27287  0.0315350 *
ps            3.46108    1.11557  3.10252  0.0045822 **
di           13.38992    2.74671  4.87490 4.6752e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
2SLS estimates for 'eq2' (equation 2)
Model Formula: p ~ q + pf
Instruments: ~ps + di + pf

             Estimate Std. Error  t value   Pr(>|t|)
(Intercept) -58.798223   5.859161 -10.0353 1.3165e-10 ***
q             2.936711   0.215772  13.6103 1.3212e-13 ***
pf            2.958486   0.155964  18.9690 < 2.22e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Signs of coefficients are same as anticipated, while intercept coefficient is weird.

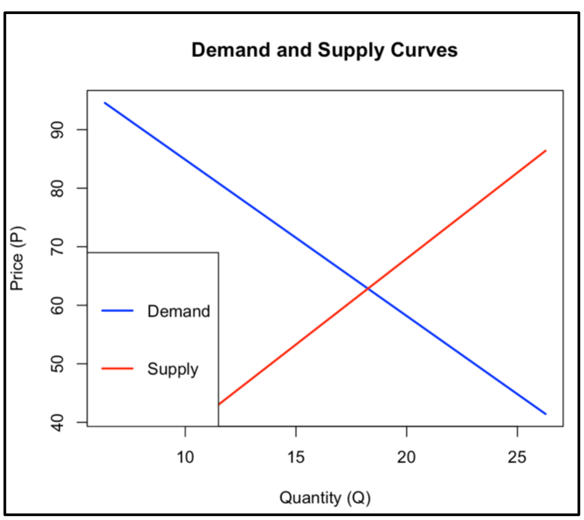
They are all significant different from 0, except demand intercept.

c. -0.7858767

```r
#c
coef_demand <- coef(truffle.sys$eq[[1]])
coef_supply <- coef(truffle.sys$eq[[2]])

q_mean <- mean(truffles$q)
ps_mean <- mean(truffles$ps)
di_mean <- mean(truffles$di)
#demand elasticity
m = truffle.sys$coefficients[2]
fitted_p =  coef_demand["(Intercept)"] + coef_demand["q"] * q_mean +
            coef_demand["ps"] * ps_mean + coef_demand["di"] * di_mean
demand_elasticity = m * q_mean / fitted_p
```

d.



e.

Method	Quantity	Price
Structural model (2SLS)	18.25021	62.84257
Reduced-form (OLS)	18.26040	62.81537

f.

Method	Equation	Term	Estimate	StdError	tValue	pValue
OLS	Demand	(Intercept)	-13.6195	9.0872	-1.50	0.1460
OLS	Demand	q	0.1512	0.4988	0.30	0.7642
OLS	Demand	ps	1.3607	0.5940	2.29	0.0303
OLS	Demand	di	12.3582	1.8254	6.77	0.0000
2SLS	Demand	(Intercept)	-11.4284	13.5916	-0.84	0.4081
2SLS	Demand	q	-2.6705	1.1750	-2.27	0.0315
2SLS	Demand	ps	3.4611	1.1156	3.10	0.0046
2SLS	Demand	di	13.3899	2.7467	4.87	0.0000
OLS	Supply	(Intercept)	-52.8763	5.0238	-10.53	0.0000
OLS	Supply	q	2.6613	0.1712	15.54	0.0000
OLS	Supply	pf	2.9217	0.1482	19.71	0.0000
2SLS	Supply	(Intercept)	-58.7982	5.8592	-10.04	0.0000
2SLS	Supply	q	2.9367	0.2158	13.61	0.0000
2SLS	Supply	pf	2.9585	0.1560	18.97	0.0000

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.

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, $\hat{P}_t$.

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?

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?

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?

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?

a. signs are same with anticipated

```
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
```

b.

```
reduced_P <- lm(p ~ g + w2 + tx + plag + klag + time + elag + e, data = klein)
linearHypothesis(reduced_P, c(
  "g = 0",
  "w2 = 0",
  "tx = 0",
  "time = 0",
  "elag = 0",
  "e = 0"
))
#p-value 7.29e-08
v_hat <- resid(reduced_P)
P_hat <- fitted(reduced_P)
```

c. P is not significantly endogenous

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.12886    5.17020   1.766 0.096524 .
p            0.51198    0.09310   5.499 4.85e-05 ***
plag         0.30526    0.09611   3.176 0.005865 **
klag        -0.10728    0.02526  -4.247 0.000616 ***
v_hat       -0.86267    0.48077  -1.794 0.091673 .
---
```

d. 2LSL

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.12886    5.51529   1.655 0.116228
p            0.51198    0.09931   5.155 7.93e-05 ***
plag         0.30526    0.10253   2.977 0.008453 **
klag        -0.10728    0.02695  -3.981 0.000967 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

OLS

```
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
```

No a lot of difference

e. Estimates are the same with part d.

Standard errors are less than part d, more efficient.

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.12886    5.10174   1.789 0.091387 .
P_hat        0.51198    0.09186   5.573 3.36e-05 ***
plag         0.30526    0.09484   3.219 0.005041 **
klag        -0.10728    0.02493  -4.304 0.000481 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9368 on 17 degrees of freedom
Multiple R-squared:  0.9409,    Adjusted R-squared:  0.9304
F-statistic: 90.17 on 3 and 17 DF,  p-value: 1.213e-10
```

f. Sargan test statistic: 18.78439

Critical value (chi^2(4), 95%): 9.487729

Non-reject H0, IVs seem valid.