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.

$$\text{Demand: } Q_i = \alpha_1 + \alpha_2 P_i + \alpha_3 PS_i + \alpha_4 DI_i + e_{di} \qquad (11.11)$$

$$\text{Supply: } Q_i = \beta_1 + \beta_2 P_i + \beta_3 PF_i + e_{si} \qquad (11.12)$$

By Demand: $\alpha_2 P_i = Q_i - (\alpha_1 + \alpha_3 PS_i + \alpha_4 DI_i + e_{di})$

$$\Rightarrow P_i = \frac{-\alpha_1}{\alpha_2} + \frac{1}{\alpha_2} Q_i - \frac{\alpha_3}{\alpha_2} PS_i - \frac{\alpha_4}{\alpha_2} DI_i - \frac{1}{\alpha_2} e_{di} = \gamma_1 + \gamma_2 Q_i + \gamma_3 PS_i + \gamma_4 DI_i + u$$

If $\gamma_2 < 0$, law of demand

$\gamma_3 > 0$, substitute goods

$\gamma_4 > 0$, truffles are normal goods

By Supply: $\beta_2 P_i = Q_i - (\beta_1 + \beta_3 PF_i + e_{si})$

$$\Rightarrow P_i = \frac{-\beta_1}{\beta_2} + \frac{1}{\beta_2} Q_i - \frac{\beta_3}{\beta_2} PF_i + \frac{1}{\beta_2} e_{si} = \delta_1 + \delta_2 Q_i + \delta_3 PF_i + \varepsilon$$

If $\delta_2 > 0$, supply curve slope > 0

$\delta_3 > 0$, cost $\uparrow \longrightarrow P \uparrow$

b.

```
Call:
ivreg(formula = p ~ q + ps + di | ps + di + pf, data = truffles)

Residuals:
    Min      1Q  Median      3Q     Max
-39.661  -6.781   2.410   8.320  20.251

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -11.428     13.592  -0.841  0.40810
q             -2.671      1.175  -2.273  0.03154
ps             3.461      1.116   3.103  0.00458
di            13.390      2.747   4.875 4.68e-05

(Intercept)
q           *
ps          **
di          ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13.17 on 26 degrees of freedom
Multiple R-Squared: 0.5567,     Adjusted R-squared: 0.5056
Wald test: 17.37 on 3 and 26 DF,  p-value: 2.137e-06
```

```
Call:
ivreg(formula = p ~ q + pf | pf + ps + di, data = truffles)

Residuals:
    Min      1Q  Median      3Q     Max
-9.7983 -2.3440 -0.6281  2.4350 11.1600

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -58.7982     5.8592  -10.04 1.32e-10 ***
q             2.9367     0.2158   13.61 1.32e-13 ***
pf            2.9585     0.1560   18.97  < 2e-16 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.399 on 27 degrees of freedom
Multiple R-Squared: 0.9486,     Adjusted R-squared: 0.9448
Wald test: 232.7 on 2 and 27 DF,  p-value: < 2.2e-16
```

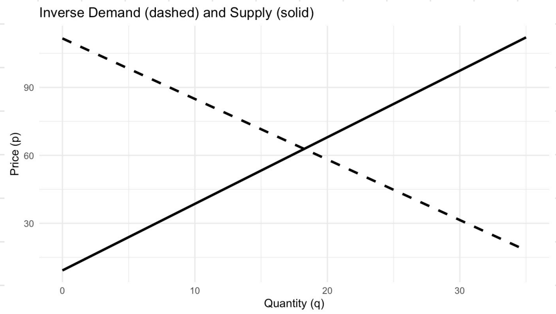
Using ivreg() (2SLS) to estimate the demand and supply equations for truffles, the signs of all key variables are consistent with economic theory, including a downward-sloping demand curve, an upward-sloping supply curve, and the expected effects of substitute price, income, and production cost on price.

Moreover, except for the intercept term in the demand equation, all main explanatory variables are statistically significant at the 5% level (and even at the 1% or 0.1% level), indicating strong statistical significance in the model estimates.

c.

```
> with(truffles, {
+   Pbar <- mean(p);   Qbar <- mean(q)
+   alpha2 <- 1 / coef(demand.iv)["q"]
+   elas  <- alpha2 * Pbar / Qbar
+   round(elas, 3)
+ })
     q
-1.272
```

d.



e.

```
> c(p_star = round(p_star, 2), q_star = round(q_star, 2))
p_star.(Intercept) q_star.(Intercept)
             62.84              18.25
```

The reduced-form model predicts a quantity of 18.26 and a price of 62.815. Based on this comparison, we believe the results from both approaches are highly consistent.

f.

```
Call:
lm(formula = p ~ q + ps + di, data = truffles)

Residuals:
     Min       1Q   Median       3Q      Max 
-25.0753  -2.7742  -0.4097   4.7079  17.4979 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -13.6195     9.0872  -1.499   0.1460    
q             0.1512     0.4988   0.303   0.7642    
ps            1.3607     0.5940   2.291   0.0303 *  
di           12.3582     1.8254   6.770 3.48e-07 ***
---
Signif. codes:  
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 8.814 on 26 degrees of freedom
Multiple R-squared:  0.8013,	Adjusted R-squared:  0.7784 
F-statistic: 34.95 on 3 and 26 DF,  p-value: 2.842e-09
```

```
Call:
lm(formula = p ~ q + pf, data = truffles)

Residuals:
    Min      1Q  Median      3Q     Max 
-8.4721 -3.3287  0.1861  2.0785 10.7513 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -52.8763     5.0238  -10.53 4.68e-11 ***
q             2.6613     0.1712   15.54 5.42e-15 ***
pf            2.9217     0.1482   19.71  < 2e-16 ***
---
Signif. codes:  
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.202 on 27 degrees of freedom
Multiple R-squared:  0.9531,	Adjusted R-squared:  0.9496 
F-statistic: 274.4 on 2 and 27 DF,  p-value: < 2.2e-16
```

With the exception of the demand equation estimated using OLS—where the sign of the quantity coefficient is incorrect—all other coefficient signs are as expected. Furthermore, aside from the intercept and the coefficient on quantity in the OLS demand model, all other coefficients are statistically significant at conventional levels.

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.

```
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

(Intercept) .  
p           ***
plag        ** 
klag        ***
---
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
```

Both current profits (p) and lagged profits (plag) have positive effects on investment, which aligns with the idea that firms with strong profits and internal funding tend to invest more.

In contrast, the negative coefficient for lagged capital stock (klag) suggests that when firms already have a large capital base, they are less likely to need additional investment—a result consistent with the accelerator model.

Importantly, all three key variables are statistically significant, with p-values below 5%.

b.

```
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
```

```
Model 1: restricted model
Model 2: p ~ g + w2 + tx + time + plag + klag + elag

  Res.Df RSS Df Sum of Sq    F Pr(>F)
1     18 108
2     13  62  5      46.1 1.93   0.16
```

```
> cat("F statistic =", round(F_stat, 3), "\n")
F statistic = 1.934
> cat("Critical F(5,13;0.95) =", round(F_crit, 3), "\n")
Critical F(5,13;0.95) = 3.025
```

```
> df$phat
 [1] 13.255556 16.577368 19.282347 20.960143 19.766509 18.238731 17.573065
 [8] 19.541720 20.375101 17.180415 12.705026  8.999780  9.054102 12.671263
[15] 14.421338 14.711907 19.796405 19.206691 17.419605 20.305654 22.657273
```

The F-test statistic is 1.93, which is less than the critical value of 3.025. Therefore, we cannot reject the null hypothesis that all the coefficients of these variables are equal to zero.

c.

```
(c)  Hausman test (t-stat of vhat):
> print(coeftest(hausman, vcov. = vcovHC(hausman, type = "HC1"))["vhat", ])
   Estimate  Std. Error     t value    Pr(>|t|)
0.574510285 0.180749980 3.178480497 0.005835657
```

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

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 ***
vhat         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
Multiple R-squared:  0.9659,	Adjusted R-squared:  0.9574
F-statistic: 113.4 on 4 and 16 DF,  p-value: 1.588e-11
```

Because the estimated values hat(v) are significant at the 0.001 level, we conclude that price P is endogenous, which aligns with the expectations of the simultaneous equations model.

d.

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

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
```

```
OLS vs 2SLS coefficients:
> print(compare_slopes, n = Inf)
# A tibble: 8 × 6
  model term        estimate std.error statistic  p.value
  <chr> <chr>          <dbl>     <dbl>     <dbl>    <dbl>
1 OLS   (Intercept)   10.1      5.47        1.85 0.0814
2 OLS   p              0.480    0.0971      4.94 0.000125
3 OLS   plag           0.333    0.101       3.30 0.00421
4 OLS   klag          -0.112    0.0267     -4.18 0.000624
5 2SLS  (Intercept)   20.3      8.38        2.42 0.0271
6 2SLS  p              0.150    0.193       0.780 0.446
7 2SLS  plag           0.616    0.181       3.40 0.00338
8 2SLS  klag          -0.158    0.0402     -3.93 0.00108
```

The strong effect observed in the OLS estimation is likely due to simultaneity bias: years with high investment also tend to be years with high profits, which causes the OLS slope estimate to be overstated. On the other hand, the 2SLS method reduces this bias but loses some precision—especially for the price variable p—because it depends on variation from the instruments rather than from the endogenous regressor itself.

e.

```
Call:
lm(formula = i ~ phat + plag + klag, data = df)

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 .
phat         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
Multiple R-squared:  0.837,	Adjusted R-squared:  0.8082
F-statistic: 29.09 on 3 and 17 DF,  p-value: 6.393e-07
```

All the slope coefficients retain their original signs and sizes; the only difference is that their standard errors have been adjusted.

f.

```
(f)  Sargan test:
> cat("   TR^2  =", round(TR2, 3), "\n")
   TR^2  = 1.815
> cat("   χ²_0.95(df=4) =", round(crit95, 3), "\n")
   χ²_0.95(df=4) = 9.488
> if (TR2 < crit95) {
+   cat("   → Fail to reject H0 : surplus instruments appear valid.\n")
+ } else {
+   cat("   → Reject H0 : at least one surplus instrument may be invalid.\n")
+ }
   → Fail to reject H0 : surplus instruments appear valid.
```