PRINCIPLES OF ECONOMETRICS

5TH EDITION

ANSWERS TO ODD-NUMBERED EXERCISES IN CHAPTER 11

EXERCISE 11.1

(a) The system of two equations has two unknowns.

$$y_1 = \alpha_1 y_2 + e_1 \qquad (1)$$
$$y_2 = \alpha_2 y_1 + \beta_1 x_1 + \beta_2 x_2 + e_2 \qquad (2)$$

Substitute equation (1) into equation (2) and simplify:

$$\begin{aligned} y_2 &= \alpha_2 y_1 + \beta_1 x_1 + \beta_2 x_2 + e_2 \\ &= \alpha_2 \left(\alpha_1 y_2 + e_1\right) + \beta_1 x_1 + \beta_2 x_2 + e_2 \\ &= \alpha_1 \alpha_2 y_2 + \alpha_2 e_1 + \beta_1 x_1 + \beta_2 x_2 + e_2 \end{aligned}$$

Rearrange terms with y_2 on the left,

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

Solve for y_2:

$$y_2 = \frac{\beta_1}{\left(1-\alpha_1\alpha_2\right)} x_1 + \frac{\beta_2}{\left(1-\alpha_1\alpha_2\right)} x_2 + \frac{e_2 + \alpha_2 e_1}{\left(1-\alpha_1\alpha_2\right)} = \pi_1 x_1 + \pi_2 x_2 + v_2$$

To show the correlation

$$\begin{aligned} \operatorname{cov}\left(y_2, e_1 \mid \mathbf{x}\right) = E\left(y_2, e_1 \mid \mathbf{x}\right) &= E\left[\left(\frac{\beta_1}{\left(1-\alpha_1\alpha_2\right)} x_1 + \frac{\beta_2}{\left(1-\alpha_1\alpha_2\right)} x_2 + \frac{e_2 + \alpha_2 e_1}{\left(1-\alpha_1\alpha_2\right)}\right) e_1 \middle| \mathbf{x}\right] \\ &= E\left[\frac{\beta_1}{\left(1-\alpha_1\alpha_2\right)} x_1 e_1 \middle| \mathbf{x}\right] + E\left[\frac{\beta_2}{\left(1-\alpha_1\alpha_2\right)} x_2 e_1 \middle| \mathbf{x}\right] + E\left[\left(\frac{e_2 + \alpha_2 e_1}{\left(1-\alpha_1\alpha_2\right)}\right) e_1 \middle| \mathbf{x}\right] \\ &= 0 + 0 + E\left[\left(\frac{e_2 e_1 + \alpha_2 e_1^2}{\left(1-\alpha_1\alpha_2\right)}\right) \middle| \mathbf{x}\right] \end{aligned}$$

The first two terms are zero because the x's are exogenous and uncorrelated with the random errors. Then, assuming that the two equation errors are uncorrelated (see p. 533 in *POE5*),

$$\begin{aligned} \operatorname{cov}\left(y_2, e_1 \mid \mathbf{x}\right) &= E\left(y_2, e_1 \mid \mathbf{x}\right) \\ &= \frac{E\left(e_2 e_1 \mid \mathbf{x}\right) + \alpha_2 E\left(e_1^2 \mid \mathbf{x}\right)}{\left(1-\alpha_1\alpha_2\right)} \\ &= \frac{\alpha_2}{\left(1-\alpha_1\alpha_2\right)} \sigma_1^2 \end{aligned}$$

This is not zero unless $\alpha_2 = 0$, in which case there is no simultaneity.

(b) Neither of the structural equations. Both equations (1) and (2) have an endogenous variable on the right-hand side. OLS is biased and inconsistent. On the other hand the reduced form equation parameters can be estimated consistently using OLS because only exogenous variables appear on the right-hand side.

(c) There are $M = 2$ equations. Identification requires that $M - 1$ variables be omitted from each equation. Equation (2) is not identified. Equation (1) is identified because x_1 and x_2 are omitted. It is possible to estimate α_1 consistently.

(d) These moment conditions arise from the assumptions that the x's are exogenous. It follows that

$$E\left(x_{i1}v_{i1} \mid \mathbf{x}\right) = E\left(x_{i2}v_{i2} \mid \mathbf{x}\right) = 0$$

From part (a), the reduced form equation for y_2 is

$$y_2 = \frac{\beta_1}{\left(1-\alpha_1\alpha_2\right)}x_1 + \frac{\beta_2}{\left(1-\alpha_1\alpha_2\right)}x_2 + \frac{e_2+\alpha_2 e_1}{\left(1-\alpha_1\alpha_2\right)} = \pi_1 x_1 + \pi_2 x_2 + v_2$$

The reduced form error is uncorrelated with the x's because

$$E\left[x_{ik}\left(\frac{e_2+\alpha_2 e_1}{\left(1-\alpha_1\alpha_2\right)}\right)\middle| \mathbf{x}\right] = E\left[\frac{1}{\left(1-\alpha_1\alpha_2\right)}x_{ik}e_2\middle| \mathbf{x}\right] + E\left[\frac{\alpha_2}{\left(1-\alpha_1\alpha_2\right)}x_{ik}e_1\middle| \mathbf{x}\right] = 0+0$$

(e) The sum of squares function, omitting the subscript i for convenience, is $S\left(\pi_1,\pi_2 \mid \mathbf{y},\mathbf{x}\right) = \sum\left(y_2 - \pi_1 x_1 - \pi_2 x_2\right)^2$. The first derivatives are

$$\frac{\partial S\left(\pi_1,\pi_2 \mid \mathbf{y},\mathbf{x}\right)}{\partial \pi_1} = 2\sum\left(y_2 - \pi_1 x_1 - \pi_2 x_2\right)x_1 = 0$$
$$\frac{\partial S\left(\pi_1,\pi_2 \mid \mathbf{y},\mathbf{x}\right)}{\partial \pi_2} = 2\sum\left(y_2 - \pi_1 x_1 - \pi_2 x_2\right)x_2 = 0$$

Divide these equations by 2, and multiply the moment equations by N to see that they are equivalent.

(f) The moment conditions are

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

Multiplying these out we have

$$\sum x_{i1}y_{i2} - \pi_1\sum x_{i1}^2 - \pi_2\sum x_{i1}x_{i2} = 0$$
$$\sum x_{i2}y_{i2} - \pi_1\sum x_{i1}x_{i2} - \pi_2\sum x_{i2}^2 = 0$$

Inserting the given values, we have

$$3 - \hat{\pi}_1 = 0 \Rightarrow \hat{\pi}_1 = 3$$
$$4 - \hat{\pi}_2 = 0 \Rightarrow \hat{\pi}_2 = 4$$

(g) The first structural equation is $y_1 = \alpha_1 y_2 + e_1$, so that

$$E\left[\left(\pi_1 x_1 + \pi_2 x_2\right)e_1 \mid \mathbf{x}\right] = E\left[\left(\pi_1 x_1 + \pi_2 x_2\right)\left(y_1 - \alpha_1 y_2\right) \mid \mathbf{x}\right] = 0$$

The empirical analog of this moment condition is

$$N^{-1}\sum\left(\pi_1 x_{i1}+\pi_2 x_{i2}\right)\left(y_{i1}-\alpha_1 y_{i2}\right)=0$$

If we knew π_1 and π_2 we could solve this moment condition for an estimator of α_1. While we do not know these parameters we can consistently estimate them from the reduced form equations. In large samples the consistent estimators converge to the true parameter values,

$$\text{plim}\ \hat{\pi}_1=\pi_1 \ \text{ and }\ \text{plim}\ \hat{\pi}_2=\pi_2$$

In a sense, having consistent estimators of parameters is "just as good as" knowing the parameter values. Replacing the unknowns by their estimates in the empirical moment condition we have

$$\sum\left(\hat{\pi}_1 x_{i1}+\hat{\pi}_2 x_{i2}\right)\left(y_{i1}-\alpha_1 y_{i2}\right)=\sum\hat{y}_{i2}\left(y_{i1}-\alpha_1 y_{i2}\right)=0$$

So that

$$\sum\hat{y}_{i2}y_{i1}-\alpha_1\sum\hat{y}_{i2}y_{i2}=0\Rightarrow\hat{\alpha}_{1,IV}=\frac{\sum\hat{y}_{i2}y_{i1}}{\sum\hat{y}_{i2}y_{i2}}$$

Inserting the values, we find

$$\hat{\alpha}_{1,IV}=\frac{\sum\hat{y}_{i2}y_{i1}}{\sum\hat{y}_{i2}y_{i2}}=\frac{\sum\left(\hat{\pi}_1 x_{i1}+\hat{\pi}_2 x_{i2}\right)y_{i1}}{\sum\left(\hat{\pi}_1 x_{i1}+\hat{\pi}_2 x_{i2}\right)y_{i2}}=\frac{\hat{\pi}_1\sum x_{i1}y_{i1}+\hat{\pi}_2\sum x_{i2}y_{i1}}{\hat{\pi}_1\sum x_{i1}y_{i2}+\hat{\pi}_2\sum x_{i2}y_{i2}}=\frac{18}{25}$$

(h) The least squares estimator of the simple regression model with no intercept is given in Exercise 2.4. Applying that result here, and substituting $\hat{y}_2$ for *x* and y_1 for *y*, we have

$$\hat{\alpha}_{1,2\text{SLS}}=\frac{\sum\hat{y}_{i2}y_{i1}}{\sum\hat{y}_{i2}^2}$$

To show that the equations are equivalent, recall that $\hat{v}_2=y_2-\hat{y}_2\Rightarrow\hat{y}_2=y_2-\hat{v}_2$ and therefore

$$\sum\hat{y}_{i2}^2=\sum\hat{y}_{i2}\left(y_2-\hat{v}_2\right)=\sum\hat{y}_{i2}y_2-\sum\hat{y}_{i2}\hat{v}_2=\sum\hat{y}_{i2}y_2$$

The term

$$\sum\hat{y}_{i2}\hat{v}_{i2}=\sum\left(\hat{\pi}_1 x_{i1}+\hat{\pi}_2 x_{i2}\right)\hat{v}_{i2}=\hat{\pi}_1\sum x_{i1}\hat{v}_{i2}+\hat{\pi}_2\sum x_{i2}\hat{v}_{i2}=0$$

because $\sum x_{i1}\hat{v}_{i2}=0$ and $\sum x_{i2}\hat{v}_{i2}=0$. This is a fundamental property of OLS that is illustrated in Exercises 2.3(f) and 2.4(g).

EXERCISE 11.3

(a) Multiply demand by Γ_{12} and supply by $-\Gamma_{11}$ and add.

$$\left(\Gamma_{12}\Gamma_{21}-\Gamma_{11}\Gamma_{22}\right)p+\left(\Gamma_{12}\text{B}_{11}-\Gamma_{11}\text{B}_{12}\right)+\left(\Gamma_{12}\text{B}_{21}-\Gamma_{11}\text{B}_{22}\right)x+\left(\Gamma_{12}E_1-\Gamma_{11}E_2\right)=0$$

Then solve for *p*,

$$p = \frac{\left(\Gamma_{12}\mathrm{B}_{11} - \Gamma_{11}\mathrm{B}_{12}\right)}{\left(\Gamma_{11}\Gamma_{22} - \Gamma_{12}\Gamma_{21}\right)} + \frac{\left(\Gamma_{12}\mathrm{B}_{21} - \Gamma_{11}\mathrm{B}_{22}\right)}{\left(\Gamma_{11}\Gamma_{22} - \Gamma_{12}\Gamma_{21}\right)}x + \frac{\left(\Gamma_{12}E_1 - \Gamma_{11}E_2\right)}{\left(\Gamma_{11}\Gamma_{22} - \Gamma_{12}\Gamma_{21}\right)}$$
$$= \pi_1 + \pi_2 x + v$$

(b) The reduced form equation will be $p = \pi_1^* + \pi_2^* x + v^*$. Express π_1^* and π_2^* in terms of parameters Γ_{ij} and B_{ij}.

$$p = \frac{\left(\Gamma_{12}\mathrm{B}'_{11} - \Gamma'_{11}\mathrm{B}_{12}\right)}{\left(\Gamma'_{11}\Gamma_{22} - \Gamma_{12}\Gamma'_{21}\right)} + \frac{\left(\Gamma_{12}\mathrm{B}'_{21} - \Gamma'_{11}\mathrm{B}_{22}\right)}{\left(\Gamma'_{11}\Gamma_{22} - \Gamma_{12}\Gamma'_{21}\right)}x + \frac{\left(\Gamma_{12}E_1 - \Gamma'_{11}E_2\right)}{\left(\Gamma'_{11}\Gamma_{22} - \Gamma_{12}\Gamma'_{21}\right)}$$

Then,

$$\pi_1^* = \frac{\left(\Gamma_{12}\left(3\mathrm{B}_{11} + 2\mathrm{B}_{12}\right) - \left(3\Gamma_{11} + 2\Gamma_{12}\right)\mathrm{B}_{12}\right)}{\left(\left(3\Gamma_{11} + 2\Gamma_{12}\right)\Gamma_{22} - \Gamma_{12}\left(3\Gamma_{21} + 2\Gamma_{22}\right)\right)}$$
$$= \frac{3\Gamma_{12}\mathrm{B}_{11} + 2\Gamma_{12}\mathrm{B}_{12} - 3\Gamma_{11}\mathrm{B}_{12} - 2\Gamma_{12}\mathrm{B}_{12}}{3\Gamma_{11}\Gamma_{22} + 2\Gamma_{12}\Gamma_{22} - 3\Gamma_{12}\Gamma_{21} - 2\Gamma_{12}\Gamma_{22}}$$
$$= \frac{3\left(\Gamma_{12}\mathrm{B}_{11} - \Gamma_{11}\mathrm{B}_{12}\right)}{3\left(\Gamma_{11}\Gamma_{22} - \Gamma_{12}\Gamma_{21}\right)} = \pi_1$$

Similarly,

$$\pi_2^* = \frac{\left(\Gamma_{12}\mathrm{B}'_{21} - \Gamma'_{11}\mathrm{B}_{22}\right)}{\left(\Gamma'_{11}\Gamma_{22} - \Gamma_{12}\Gamma'_{21}\right)}$$
$$= \frac{\left(\Gamma_{12}\left(3\mathrm{B}_{11} + 2\mathrm{B}_{12}\right) - \left(3\Gamma_{11} + 2\Gamma_{12}\right)\mathrm{B}_{22}\right)}{\left(\left(3\Gamma_{11} + 2\Gamma_{12}\right)\Gamma_{22} - \Gamma_{12}\left(3\Gamma_{21} + 2\Gamma_{22}\right)\right)}$$
$$= \frac{3\Gamma_{12}\mathrm{B}_{11} + 2\Gamma_{12}\mathrm{B}_{12} - 3\Gamma_{11}\mathrm{B}_{22} - 2\Gamma_{12}\mathrm{B}_{22}}{3\Gamma_{11}\Gamma_{22} + 2\Gamma_{12}\Gamma_{22} - 3\Gamma_{12}\Gamma_{21} - 2\Gamma_{12}\Gamma_{22}}$$
$$= \frac{3\left(\Gamma_{12}\mathrm{B}_{11} - \Gamma_{11}\mathrm{B}_{22}\right)}{3\left(\Gamma_{11}\Gamma_{22} - \Gamma_{12}\Gamma_{21}\right)} = \pi_2$$

What is the point of this long, tedious exercise? Based on the reduced form coefficients we cannot determine the true demand or supply equation from an arbitrary mixture, linear combination, of the original equations. This is unfortunate because we can consistently estimate the reduced form coefficients, but without further information about the structural equations we cannot tell anything about the structural parameters.

EXERCISE 11.5

(a) For q:

$$q = -p + 3 + 2x + e_1 = -\left(q + 1 + e_2\right) + 3 + 2x + e_1$$
$$\Rightarrow \quad q = 1 + x + \frac{\left(e_1 - e_2\right)}{2}$$

For p:

$$p = \left[1 + x + \frac{(e_1 - e_2)}{2}\right] + 1 + e_2 = 2 + x + \frac{(e_1 + e_2)}{2}$$

(b) For q:

$$q = -5\left[q + 1 + e_2\right] + 11 + 6x + e_1^* = 1 + x + \frac{(e_1 - e_2)}{2}$$

This is the same reduced form equation as in part (a) for q. Therefore, the reduced form for p stays the same as well.

$$p = 2 + x + \frac{(e_1 + e_2)}{2}$$

(c)

$$q = -p + 3 + 2x + e_1 \quad \Rightarrow \quad -q - p + 3 + 2x + e_1 = 0$$

$$p = q + 1 + e_2 \quad \Rightarrow \quad q - p + 1 + e_2 = 0$$

Then

$$\begin{aligned}
&3(-q - p + 3 + 2x + e_1) + 2(q - p + 1 + e_2) = 0 \\
&\Rightarrow \quad -q - 5p + 11 + 6x + (3e_1 + 2e_2) = 0 \\
&\Rightarrow \quad q = -5p + 11 + 6x + (3e_1 + 2e_2)
\end{aligned}$$

(d) No. There is an endogenous variable on the right-hand side. The supply equation is $p = q + 1 + e_2$ and the reduced form equation for q is $q = 1 + x + (e_1 - e_2)/2$, which is a function of e_2. Assuming $E(e_2 \mid \mathbf{x}) = 0$ and $E(xe_2 \mid \mathbf{x}) = 0$ because x is exogenous, and assuming $\operatorname{cov}(e_1, e_2 \mid \mathbf{x}) = 0$, and that the errors are conditionally homoskedastic, $E(e_2^2 \mid \mathbf{x}) = \sigma_2^2$, we have

$$\operatorname{cov}(q, e_2 \mid \mathbf{x}) = E(qe_2 \mid \mathbf{x}) = E\left[\left(1 + x + \frac{(e_1 - e_2)}{2}\right)e_2 \middle| \mathbf{x}\right]$$

$$= E\left[\left(e_2 + xe_2 + \frac{(e_1e_2 - e_2^2)}{2}\right)\middle| \mathbf{x}\right] = 0 + 0 + 0 - \frac{\sigma_2^2}{2}$$

Thus, the OLS estimator of the supply equation is biased and inconsistent.

(e) Yes. The reduced form equations are

$$q = 1 + x + \frac{(e_1 - e_2)}{2} \qquad p = 2 + x + \frac{(e_1 + e_2)}{2}$$

The right-hand side variable x is exogenous and uncorrelated with the random errors e_1 and e_2, so that the OLS estimator is consistent.

(f) No. Presumably x is "income", which affects demand and not supply. So, the given equation could not be a supply equation based on economic theory.

(g) There are M = 2 equations, so to be identified at least $M-1=1$ variable must be omitted from an equation. The demand equation includes all the variables and is not identified. Supply omits x and satisfies the necessary condition for identification.

EXERCISE 11.7

(a) First the reduced form for q.

$$q=\alpha_1 p+\alpha_2 x+\alpha_3 w+e_1=\alpha_1\left(\beta_1 q+e_2\right)+\alpha_2 x+\alpha_3 w+e_1$$

So that

$$q=\frac{\alpha_2}{1-\alpha_1\beta_1}x+\frac{\alpha_3}{1-\alpha_1\beta_1}w+\left(\alpha_1 e_2+e_1\right)=\pi_{11}x+\pi_{21}w+v_1$$

Then the reduced form for p is

$$\begin{aligned}p&=\beta_1\left(\frac{\alpha_2}{1-\alpha_1\beta_1}x+\frac{\alpha_3}{1-\alpha_1\beta_1}w+\left(\alpha_1 e_2+e_1\right)\right)+e_2\\&=\frac{\beta_1\alpha_2}{1-\alpha_1\beta_1}x+\frac{\beta_1\alpha_3}{1-\alpha_1\beta_1}w+\left[\beta_1\left(\alpha_1 e_2+e_1\right)+e_2\right]\\&=\pi_{12}x+\pi_{22}w+v_2\end{aligned}$$

(b) Only supply is identified. $\beta_1=\frac{\pi_{12}}{\pi_{11}}=\frac{\pi_{22}}{\pi_{21}}=\frac{2/5}{1/5}=2$. In this system there are two ways to estimate β_1, and if we know the true reduced form parameters each approach gives the same solution. If we did not know the true reduced form parameters, then we would wind up with two different estimates of β_1. This is an example of an overidentified equation. There are M = 2 equations and we must omit at least $M-1=1$ to identify it. The supply equation omits 2 variables.

EXERCISE 11.9

(a)
$$y_{i1}=\alpha_1 y_{i2}+\alpha_2 x_{i1}+e_{i1}=\alpha_1\left(\alpha_2 y_{i1}+\beta_1 x_{i1}+e_{i2}\right)+\alpha_2 x_{i1}+e_{i1}$$

$$\Rightarrow\quad y_{1i}=\frac{\left(\alpha_1\beta_1+\alpha_2\right)}{1-\alpha_1\beta_1}x_{i1}+\frac{\alpha_1 e_{i2}+e_{i1}}{1-\alpha_1\beta_1}$$

(b)
$$\begin{aligned}\operatorname{cov}\left(y_{i1},e_{i2}\mid\mathbf{x}_1\right)&=E\left(y_{i1}e_{i2}\mid\mathbf{x}_1\right)=E\left(\left(\frac{\left(\alpha_1\beta_1+\alpha_2\right)}{1-\alpha_1\beta_1}x_{i1}+\frac{\alpha_1 e_{i2}+e_{i1}}{1-\alpha_1\beta_1}\right)e_{i2}\middle|\mathbf{x}_1\right)\\&=E\left(\frac{\left(\alpha_1\beta_1+\alpha_2\right)}{1-\alpha_1\beta_1}x_{i1}e_{i2}\middle|\mathbf{x}_1\right)+E\left(\frac{\alpha_1 e_{i2}^2+e_{i1}e_{i2}}{1-\alpha_1\beta_1}\middle|\mathbf{x}_1\right)\\&=\frac{\alpha_1\sigma_2^2+\sigma_{12}}{1-\alpha_1\beta_1}\end{aligned}$$

(c) $$\operatorname{cov}\left(y_{i1}, e_{i2} \mid \mathbf{x}_1\right) = \frac{\alpha_1\sigma_2^2 + \sigma_{12}}{1-\alpha_1\beta_1} = \frac{0\sigma_2^2 + 0}{1-0\beta_1} = \frac{0}{1} = 0$$

(d) Under the conditions in (c) the two equations are

$$y_{i1} = \alpha_2 x_{i1} + e_{i1}$$
$$y_{i2} = \alpha_2 y_{i1} + \beta_1 x_{i1} + e_{i2}$$

The first equation can be consistently estimated by OLS because it has a single exogenous variable on the right-hand side.

(e) Yes, because under the conditions in (c), y_{i2} is not correlated with e_{i2} and is thus not endogenous. OLS can be used to estimate the second equation consistently.

EXERCISE 11.11

(a) Demand is not identified. There are $M = 2$ equations, so the necessary condition is that at least $M-1=1$ variables need to be omitted from an equation for it to be identified. Demand has no variables omitted. Supply is identified, as it excludes *PS* and *DI*.

(b) To be adequately strong in the first stage regression $P = \pi_1 + \pi_2 PS + \pi_3 DI + v$ we must reject the null hypothesis $H_0: \pi_2 = 0, \pi_3 = 0$ with a large F-value. The rule of thumb threshold to reject the notion that the instruments are weak is $F > 10$. Here the overall $F = 54.21$, which is large; we conclude the instruments are not weak. Furthermore, the F-test statistic for $H_0: \pi_2 = 0$ is $F = 10.45$, and the F-test statistic for $H_0: \pi_3 = 0$ is $F = 48.34$, so each of the demand shift variables is strongly significant.

(c) $$\hat{Q} = 8.6455 + 0.1564\bar{P} = 8.6455 + 0.1564(62.724) = 18.458333 = \bar{Q}$$

(d) $$\hat{\varepsilon} = b_2\left(\bar{P}/\bar{Q}\right) = 0.1564(62.724/18.458) = 0.532$$

Using the estimates from Table 11.3b

$$\hat{\varepsilon} = b_2\left(\bar{P}/\bar{Q}\right) = 0.338(62.724/18.458) = 1.149$$

With the revised supply equation, the estimated price elasticity is only 0.532, compared to 1.149 from the full model.

(e) *PF* should be included. In the full model the reduced form for *PRICE* is

$$P = \pi_{12} + \pi_{22}PS + \pi_{32}DI + \pi_{42}PF + v_2$$

The F-test of $H_0: \pi_{22} = 0, \pi_{32} = 0$ yields an $F = 41.49$, so that *PS* and *DI* are strong instruments. The t-value for *PF* in the estimated supply equation is $t = -12.13$. We reject the null hypothesis that $\beta_3 = 0$. The 95% interval estimate is $[-1.17, -0.83]$ which is reasonably narrow.

EXERCISE 11.13

This exercise is adapted from "Econometric Models, Techniques and Applications, Second Edition," by Michael D. Intriligator, Ronald G. Bodkin and Cheng Hsiao (1996) Prentice-Hall, pp. 468-469. The reported results are from Teigen, R. (1964) "Demand and Supply Functions for Money in the United States: Some Structural Estimates," *Econometrica*, 32, 476-509. Any errors are ours.

(a) M_{t-1} is an exogenous variable in the demand equation with a positive and significant coefficient. If $\Delta M_{t-1} = 1$ then the demand curve shifts right (up) indicating an increase in demand. On the other hand, $\Delta M_{t-1} > 0$ has no effect on the supply curve.

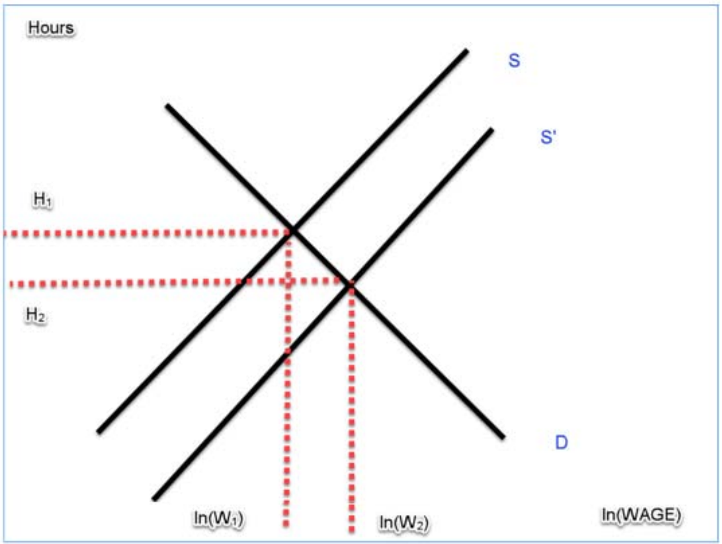
(b) $\partial \hat{M}_t / \partial GNP_t = 0.0618 - 0.0025R$. If the short-term interest rate is less than 24.72%, then the increase in *GNP* will increase demand, shifting the demand curve to the right (up). A change in *GNP* does not affect supply.

(c) $\partial \hat{M}_t / \partial R_{d,t} = -0.0751 M_t^*$. Since $M_t^* > 0$, the increase in the discount rate reduces the money supply, shifting the supply curve to the left (down).

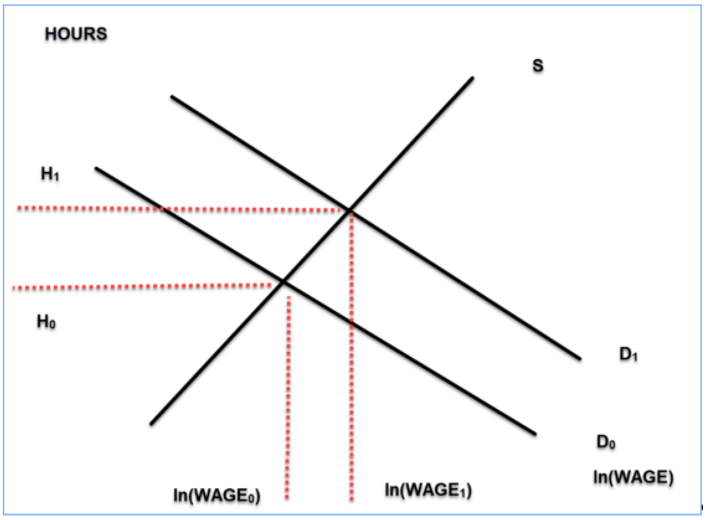
(d) Identification requires that at least one variable that appears elsewhere in the system be omitted from each of the two equations. The demand equation omits $R_{d,t}$, the discount rate. Changes in $R_{d,t}$ shift the supply curve relative to demand. GNP and M_{t-1} are omitted from supply. Changes in these variables shift demand relative to supply. Thus, both equations satisfy the necessary condition for identification.

EXERCISE 11.15

(a) The signs of β_5 and β_6 are likely to be negative. An increase in the number of children will reduce supply, shifting it down (right). This will reduce the number of hours worked and increase the market wage rate, as illustrated in the figure below.



(b) An increase in experience will increase demand, shifting the demand relation upwards (right), resulting in higher equilibrium wages and hours, as illustrated in the figure below.



(c) In a system of $M = 2$ equations at least $M - 1 = 1$ variable must be omitted from an equation to identify it. The supply equation omits experience and its square, the necessary condition holds. The reduced form equation is

$$HOURS = \pi_1 + \pi_2 EDUC + \pi_3 AGE + \pi_4 KIDSL6 + \pi_5 KIDS618$$
$$+ \pi_6 NWIFEINC + \pi_7 EXPER + \pi_8 EXPER^2 + v_1$$

To test instrument strength, we test the joint null hypothesis $H_0 : \pi_7 = 0, \pi_8 = 0$.

(d) The demand equation omits *AGE*, *KIDSL6*, *KIDS618*, and *NWIFEINC*. The necessary condition that one variable in the system be omitted is satisfied. The reduced form is

$$\ln(WAGE) = \gamma_1 + \gamma_2 EDUC + \gamma_3 AGE + \gamma_4 KIDSL6 + \gamma_5 KIDS618$$
$$+ \gamma_6 NWIFEINC + \gamma_7 EXPER + \gamma_8 EXPER^2 + v_2$$

Test the joint null hypothesis $H_0 : \gamma_3 = 0, \gamma_4 = 0, \gamma_5 = 0, \gamma_6 = 0$.

EXERCISE 11.17

(a) There are $M = 8$ equations requiring 7 omitted variables in each equation. There is a total of 16 variables in the system. The consumption equation includes 6 variables and omits 10. The necessary condition is satisfied. The investment equation includes 5 variables and omits 11. The necessary condition is satisfied. The private sector wage equation includes 5 variables and omits 11. The necessary condition is satisfied.

(b) The consumption equation has 2 RHS endogenous variables and excludes 5 exogenous variables. The investment and private wage equations have 1 RHS endogenous variable and omit 5 exogenous variables.

(c) $$W_{1t} = \pi_1 + \pi_2 G_t + \pi_3 W_{2t} + \pi_4 TX_t + \pi_5 TIME_t + \pi_6 P_{t-1} + \pi_7 K_{t-1} + \pi_8 E_{t-1} + v$$

(d) Obtain fitted values $\hat{W}_{1t}$ from the estimated reduced form equation in part (c) and similarly obtain $\hat{P}_t$. Create $W_t^* = \hat{W}_{1t} + W_{2t}$. Regress CN_t on W_t^*, $\hat{P}_t$ and P_{t-1} plus a constant by OLS.

(e) The coefficient estimates will be the same. The *t*-values will not be because the standard errors in part (d) are not correct 2SLS standard errors.

EXERCISE 11.19

(a) The summary statistics are in the following tables. We observe that women not in the labor force are on average slightly older, have more kids under 6 years of age, and a higher income from their partners.

Summary statistics for women not in the labor force, *LFP* = 0

variable	*N*	mean	Std. dev.	min	max
AGE	325	43.28308	8.467796	30	60
KIDSL6	325	.3661538	.6368995	0	3
KIDS618	325	1.356923	1.327065	0	7
NWIFEINC	325	21698.05	12728.15	1500	96000

Summary statistics for women in the labor force, *LFP* = 1

variable	*N*	mean	Std. dev.	min	max
AGE	428	41.97196	7.721084	30	60
KIDSL6	428	.1401869	.3919231	0	2
KIDS618	428	1.350467	1.315935	0	8
NWIFEINC	428	18937.48	10591.35	-29.0575	91000

(b) $\beta_2 > 0$: A higher wage leads to an increased quantity of labor supplied.

β_3: The effect of an increase in education is unclear.

β_4: This sample has been taken for working women between the ages of 30 and 60. It is not certain whether hours worked increases or decreases over this age group.

$\beta_5 < 0$, $\beta_6 < 0$: The presence of children in the household reduces the number of hours worked because they demand time from their mother.

$\beta_7 < 0$: As income from other sources increases, it becomes less necessary for the woman to work.

NWIFEINC measures the sum of all family income excluding the wife's income.

(c) The estimates are in column (1) of Table XR11.19. We certainly did not expect the coefficient of $\ln(WAGE)$ to be negative and insignificant. It is meant to be a supply equation; the coefficient should be positive and significant.

Table XR11.19

	(1) OLS	(2) Red Form	(3) 2SLS
C	2114.6973	-0.1620	2478.4349
	(6.22)	(-0.54)	(3.78)
LWAGE	-17.4078		1772.3233
	(-0.32)		(2.98)
EDUC	-14.4449	0.1011	-201.1870
	(-0.80)	(6.69)	(-2.88)
AGE	-7.7300	-0.0055	-11.2289
	(-1.40)	(-1.04)	(-1.07)
KIDSL6	-342.5048	-0.0697	-191.6588
	(-3.42)	(-0.79)	(-0.98)
KIDS618	-115.0205	-0.0207	-37.7325
	(-3.73)	(-0.74)	(-0.59)
NWIFEINC	-0.0042	0.0000	-0.0100
	(-1.16)	(1.80)	(-1.39)
EXPER		0.0175	
		(3.60)	
N	428	428	428
R^2	0.0670	0.1572	.
RMSE	755.1606	0.6686	1430.5254

t statistics in parentheses

(d) The estimates are in column (2) of Table XR11.19. The coefficient of education implies an additional year of education will increase wages by approximately 10.11%, holding other factors constant. The estimate is statistically significant at the 1% level.

(e) For the supply equation to be identified the coefficient on *EXPER* should be very significant, with an *F*-test statistic value of at least 10. The *t*-value is 3.60 which translates into $F = 12.96$, satisfying the rule of thumb threshold for an instrument that is not weak.

(f) These estimates are in column (3) of Table XR11.19. We see that $\ln(WAGE)$ now has a positive and statistically significant coefficient. The coefficient of *EDUC* is negative and significant, but the rest of the coefficients are insignificant.

EXERCISE 11.21

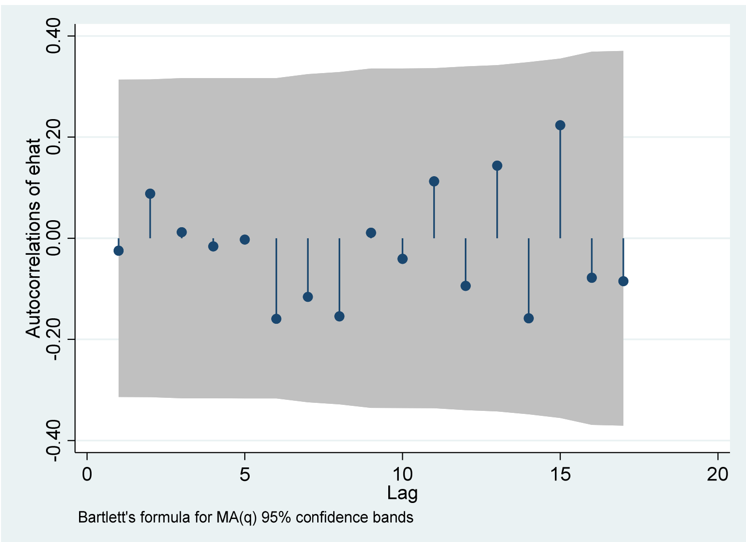
(a) *QPROD*, production, and *P*, price, are the endogenous variables. The exogenous variables are *PF*, feed price, *TIME*, and lagged log production, $\ln\left(QPROD_{t-1}\right)$. In this supply relation we expect $\beta_2 > 0, \beta_3 < 0, \beta_4 > 0, \beta_5 > 0$. A higher price should elicit more production, a higher feed price reduces supply, over time technology should improve production, and the desired level of production may not be achieved in one year, so lagged production is included to capture "partial adjustment."

(b) The estimates are in the first column of Table XR11.21. All the coefficients have the anticipated sign, but the coefficient of price is insignificant. The LM test statistic for serial correlation, AR(1), is 0.039 with a *p*-value of 0.8444. The correlogram also shows no sign of serial correlation. The lagged value of production controls for such a problem.

Table XR11.21

	(1) ols	(2) red form	(3) 2sls
C	2.154**	-5.412	2.214**
	(0.782)	(6.297)	(0.800)
LP	0.0252		0.0446
	(0.0671)		(0.123)
LPF	-0.0999*	0.177	-0.105*
	(0.0421)	(0.108)	(0.0486)
TIME	0.0113*	-0.0505*	0.0120*
	(0.00503)	(0.0216)	(0.00589)
L.LQPROD	0.727***	-0.141	0.718***
	(0.104)	(0.327)	(0.110)
LY		0.856	
		(0.630)	
LPB		0.219	
		(0.234)	
L.LEXPTS		2.322**	
		(0.709)	
POPGRO		-0.0231	
		(0.118)	
N	39	39	39
R2	0.997	0.906	0.997
RMSE	0.0304	0.0699	0.0284

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$



(c) The first stage estimates are in column (2) of Table XR11.21. The only significant external IV is the log of lagged exports. The joint F-test of the four instruments yields $F = 2.71$, which is significant at the 0.05 level, but does not reach the rule of thumb threshold of 10. We cannot reject the null hypothesis that the IV are weak using the Stock-Yogo test critical values.

(d) The 2SLS estimates are in column (3) of Table XR11.21. The results are not qualitatively different from the OLS estimates.

(e) The value of the Sargan LM test statistic is 6.5754 with $p = 0.0867$, using $\chi^2_{(3)}$. Thus at 5% we cannot reject the validity of the three surplus IV.

EXERCISE 11.23

(a) The OLS estimates of the supply equation are as follows

Dependent Variable: LQUAN
Method: Least Squares
Included observations: 111

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.500857	0.098059	86.69135	0.0000
LPRICE	-0.438081	0.194184	-2.256015	0.0261
STORMY	-0.216019	0.162994	-1.325321	0.1879
R-squared	0.092341	Mean dependent var		8.523430

We see that the coefficient of log price is negative and significant. That is quite opposite of what we expect. The coefficient of *STORMY* is insignificant. To examine possible bias in the estimation of β_2, we first derive the reduced form equation for $\ln(PRICE)$. The demand and supply equations in (11.13) and (11.14) are

$$\ln(QUAN_t) = \alpha_1 + \alpha_2 \ln(PRICE_t) + \alpha_3 MON_t + \alpha_4 TUE_t + \alpha_5 WED_t + \alpha_6 THU_t + e_{dt}$$

$$\ln(QUAN_t) = \beta_1 + \beta_2 \ln(PRICE_t) + \beta_3 STORMY_t + e_{st}$$

The reduced form for log price is obtained by setting the log quantity from supply and demand equal, in equilibrium they are

$$\begin{aligned}\beta_1 + \beta_2 \ln(PRICE_t) + \beta_3 STORMY_t + e_{st} &= \alpha_1 + \alpha_2 \ln(PRICE_t) + \alpha_3 MON_t \\ &\quad + \alpha_4 TUE_t + \alpha_5 WED_t + \alpha_6 THU_t + e_{dt}\end{aligned}$$

Solving for $\ln(PRICE)$, we have

$$\begin{aligned}\ln(PRICE_t) = {}& \frac{\alpha_1}{\beta_2 - \alpha_2} + \frac{\alpha_3}{\beta_2 - \alpha_2} MON_t + \frac{\alpha_4}{\beta_2 - \alpha_2} TUE_t \\ & + \frac{\alpha_5}{\beta_2 - \alpha_2} WED_t + \frac{\alpha_6}{\beta_2 - \alpha_2} THU_t - \frac{\beta_3}{\beta_2 - \alpha_2} STORMY_t + \frac{e_{dt} - e_{st}}{\beta_2 - \alpha_2}\end{aligned}$$

Then the covariance between $\ln(PRICE_t)$ and e_{st} is

$$E\left[\ln\left(PRICE_t\right)e_{st} \mid \mathbf{X}\right] = E\left[\left(\frac{e_{dt}-e_{st}}{\beta_2-\alpha_2}\right)e_{st}\middle|\mathbf{X}\right] = \frac{-\sigma_s^2}{\beta_2-\alpha_2}$$

The sign of the covariance is negative, because $\beta_2 > 0$ and $\alpha_2 < 0$, so that $\beta_2 - \alpha_2 > 0$. This is the same result as in Section 11.3.1 and the outcome is the same as well. The OLS estimator of the supply equation coefficient of log price is negatively biased.

(b) To find the reduced form for $\ln(PRICE)$ when *RAINY* and *COLD* are included in the demand equation, we have

$$\beta_1 + \beta_2 \ln\left(PRICE_t\right) + \beta_3 STORMY_t + e_{st} = \alpha_1 + \alpha_2 \ln\left(PRICE_t\right) + \alpha_3 MON_t$$
$$+\alpha_4 TUE_t + \alpha_5 WED_t + \alpha_6 THU_t + \alpha_7 RAINY + \alpha_8 COLD + e_{dt}$$

which yields

$$\ln\left(PRICE_t\right) = \frac{\alpha_1}{\beta_2-\alpha_2} + \frac{\alpha_3}{\beta_2-\alpha_2}MON_t + \frac{\alpha_4}{\beta_2-\alpha_2}TUE_t + \frac{\alpha_5}{\beta_2-\alpha_2}WED_t + \frac{\alpha_6}{\beta_2-\alpha_2}THU_t$$
$$+\frac{\alpha_7}{\beta_2-\alpha_2}RAINY + \frac{\alpha_8}{\beta_2-\alpha_2}COLD - \frac{\beta_3}{\beta_2-\alpha_2}STORMY_t + \frac{e_{dt}-e_{st}}{\beta_2-\alpha_2}$$

(c) These estimates are given below. We see that the coefficients of *RAINY* and *COLD* are not significant, and the joint *F*-statistic is 0.60. For a $F_{(2,103)}$ distribution, $p = 0.5519$. Thus, the coefficients of *RAINY* and *COLD* are not significant.

Dependent Variable: LPRICE
Method: Least Squares
Included observations: 111

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.290228	0.082069	-3.536396	0.0006
MON	-0.121576	0.108589	-1.119604	0.2655
TUE	-0.056677	0.106981	-0.529786	0.5974
WED	-0.028360	0.108520	-0.261330	0.7944
THU	0.040420	0.105824	0.381961	0.7033
RAINY	-0.016733	0.093620	-0.178737	0.8585
COLD	0.080989	0.074359	1.089155	0.2786
STORMY	0.312658	0.081793	3.822553	0.0002
R-squared	0.188311	Mean dependent var		-0.193681

(d) The *F*-value for the joint hypothesis is $F = 0.61$. For a $F_{(6,103)}$ distribution, $p = 0.7229$. The variables are not jointly significant.

(e) These estimates are given below. Nothing is significant and the wrong sign on log price persists. The results from the previous parts tell us that there are no variables in the demand equation that shift it relative to supply. So, the supply equation is not identified, and the estimates are not meaningful.

Dependent Variable: LQUAN
Method: Two-Stage Least Squares
Included observations: 111
Instrument specification: STORMY MON TUE WED THU RAINY COLD
Constant added to instrument list

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.584814	0.318287	26.97190	0.0000
LPRICE	-0.148909	1.060142	-0.140462	0.8886
STORMY	-0.312967	0.386154	-0.810474	0.4195

EXERCISE 11.25

(a) Estimates for the reduced form equation for $\ln(PRICE)$ are in Table XR11.25a, column (1). The *p*-value for testing the null hypothesis that the coefficient of *STORMY* is zero is 0.0000. Since this value is less than the level of significance, 0.05, we reject the null hypothesis and conclude that this coefficient is significantly different from zero. The *F*-test value is 29.00, well above the rule of thumb threshold of 10. It is important to test for the statistical significance of *STORMY* because it is the supply equation's shift variable. It is required to be statistically significant for the demand equation to be identified. If *STORMY* is not statistically significant, then the two-stage least squares regression and the estimation procedure will be unreliable.

The Stock-Yogo test for weak instrument critical value, using the criteria of test size, is 16.38 [Table 10E.1] if we can tolerate a test with Type I error of 10% for a 5% nominal test.

Table XR11.25a: *CHANGE* = 1

	(1) red form	(2) Hausman	(3) OLS	(4) 2SLS
C	-0.482***	8.363***	8.533***	8.363***
	(0.0975)	(0.180)	(0.167)	(0.191)
LPRICE		-1.019**	-0.435*	-1.019**
		(0.317)	(0.175)	(0.336)
STORMY	0.444***			
	(0.0824)			
MON	0.0379	0.295	0.293	0.295
	(0.123)	(0.209)	(0.215)	(0.222)
TUE	0.0678	-0.345	-0.350	-0.345
	(0.123)	(0.209)	(0.215)	(0.222)
WED	0.143	-0.362	-0.408	-0.362
	(0.129)	(0.221)	(0.225)	(0.234)
THU	0.257*	0.394	0.269	0.394
	(0.127)	(0.226)	(0.224)	(0.240)
VHAT1		0.821*		
		(0.376)		
N	77	77	77	77
R^2	0.320	0.304	0.256	0.140
RMSE	0.340	0.580	0.595	0.615

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(b) These estimates are in column (2) of Table XR11.25a. The null hypothesis of this Hausman test is $H_0: \text{cov}(\ln(PRICE), e) = 0$, which is tested by testing for the significance of the coefficient of $\hat{v}_1$ in

$$\ln(QUAN) = \alpha_1 + \alpha_2 \ln(PRICE) + \alpha_3 MON + \alpha_4 TUE + \alpha_5 WED + \alpha_6 THU + \delta\hat{v}_1 + error$$

In the results above $\hat{v}_1$ is denoted *VHAT1*. The *t*-statistic and *p*-value for the null hypothesis $H_0: \delta = 0$ are 2.1847 and 0.0323 respectively. Since this *p*-value is less than the level of significance, 0.05, we reject the null hypothesis and conclude that ln(*PRICE*) is endogenous. The robust version of this test yields a *t*-statistic of 2.27, and thus our conclusion is unchanged.

(c) The 2SLS estimates are in column (4) of Table XR11.25a and the OLS estimates are in column (3). These estimates have the expected signs. The two-stage least squares and least squares estimates are very similar except for the coefficient of ln(*PRICE*). Both estimation procedures conclude that the day indicator variables are not significant at a 5% level of significance.

Compared to Table 11.5 all estimated coefficients have the same sign except for the coefficient of *MON*. Also, the intercept estimate and the coefficient estimate of ln(*PRICE*) are similar but the coefficient estimates for *TUE*, *WED* and *THU* are quite different. Furthermore, all the part (c) two-stage least-squares estimates of the weekday indicator variables are insignificant, whereas, in Table 11.5, *TUE* and *WED* are statistically significant.

(d) These estimates are in column (1) of Table XR11.25b. They are very different to those obtained in part (a). All the coefficients of the weekday indicator variables have opposite signs and the coefficient for *STORMY* is smaller. In addition, in part (a) the only variables which were not statistically significant were *MON*, *TUE* and *WED*. In part (d) all exogenous variables are statistically insignificant.

Comparing these results to Table 11.4(b), all the estimated coefficients have very different values, although the only estimated coefficient with the opposite sign is the coefficient of *THU*. All weekday indicator variables are statistically insignificant in both estimated regressions. However, *STORMY* is statistically significant in Table 11.4b and not statistically significant in the above regression.

(e) These estimates are in column (2) of Table XR11.25b. As described in part (b), the Hausman test is a test for the endogeneity of ln(*PRICE*), which is tested by testing for the significance of the coefficient of $\hat{v}_2$ in

$$\ln(QUAN) = \alpha_1 + \alpha_2 \ln(PRICE) + \alpha_3 MON + \alpha_4 TUE + \alpha_5 WED + \alpha_6 THU + \delta\hat{v}_2 + e^d$$

The variable *VHAT0* = $\hat{v}_2$. The *t*-statistic and *p*-value for the null hypothesis $H_0: \delta = 0$ are −0.0405 and 0.9680, respectively. Since this *p*-value is greater than the level of significance, 0.05, we do not reject the null hypothesis and conclude ln(*PRICE*) does not show signs of endogeneity. This is consistent with Graddy and Kennedy's expectation that when inventory changes are small, simultaneity between demand and supply does not exist.

Table XR11.25b: *CHANGE* = 0

	(1) red form	(2) Hausman	(3) OLS	(4) 2SLS
C	-0.0103	8.777***	8.776***	8.777***
	(0.112)	(0.270)	(0.264)	(0.266)
LPRICE		-0.868	-0.975*	-0.868
		(2.677)	(0.439)	(2.632)
STORMY	0.149			
	(0.167)			
MON	-0.171	-0.901	-0.912	-0.901
	(0.219)	(0.549)	(0.478)	(0.540)
TUE	-0.0322	-0.843	-0.841	-0.843
	(0.186)	(0.427)	(0.417)	(0.420)
WED	-0.190	-0.872	-0.890*	-0.872
	(0.171)	(0.607)	(0.406)	(0.597)
THU	-0.244	-0.355	-0.379	-0.355
	(0.165)	(0.720)	(0.397)	(0.707)
VHAT0		-0.110		
		(2.715)		
N	34	34	34	34
R^2	0.118	0.304	0.304	0.303
RMSE	0.336	0.806	0.792	0.792

Standard errors in parentheses, $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

(f) These estimates are in columns (3) and (4) of Table XR11.25b. All the estimates have the expected signs and are almost identical. The major difference between the two sets of estimates is that, as a consequence of the smaller least-squares standard errors, all of the least squares coefficient estimates are significantly different from zero except those for *MON*, *TUE* and *THU*, whereas none of the two-stage least squares coefficient estimates are significantly different from zero.

Comparing these values to those in part (c), we find that the coefficient estimates for ln(*PRICE*) appear to be quite similar with the exception of the least squares coefficient estimate of ln(*PRICE*) in part (c), which is likely to exhibit simultaneous equation bias. Also, the coefficient of ln(*PRICE*) is always significantly different from zero in part (c) and only significant in the least squares part (f) estimation. The estimated values of the coefficients of the weekday indicator variables are very different.

Part (c) models the demand for fish when there are large changes in inventory, and part (f) models the demand for fish for small changes in inventory. It has been postulated that sellers are more responsive to prices when more fish are sold and bought, causing large changes in inventory, and therefore endogeneity is present. On the days where there is little change in inventory, endogeneity should not be present. This is supported by our estimates which show that the two stage least squares and least squares coefficient estimates of ln(*PRICE*) are similar when *CHANGE* = 0 but very different when *CHANGE* = 1. This discrepancy suggests that a coefficient bias exists when *CHANGE* = 1 due to endogeneity. Also note that the least squares estimate of the price elasticity of demand when *CHANGE* = 0 is similar in magnitude to the two-stage least squares estimate of the price elasticity of demand when *CHANGE* = 1.

EXERCISE 11.27

The estimates are in Table XR11.27.

Table XR11.27

	(1) Demand_OLS	(2) Demand_2SLS	(3) Supply_OLS	(4) Supply_2SLS
C	1.0910	-4.2795	20.0328***	20.0328***
	(3.7116)	(5.5439)	(1.2220)	(1.2231)
P	0.0233	-0.3745*	0.3380***	0.3380***
	(0.0768)	(0.1648)	(0.0217)	(0.0249)
PS	0.7100**	1.2960**		
	(0.2143)	(0.3552)		
DI	0.0764	5.0140*		
	(1.1909)	(2.2836)		
PF			-1.0009***	-1.0009***
			(0.0764)	(0.0825)
N	30	30	30	30

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Comparing the demand equation estimates in columns (1) and (2) we see dramatic differences. The OLS estimate of the coefficient of price is insignificant, whereas the 2SLS estimate is negative and significant, as it should be, even though it has a larger standard error. The OLS and 2SLS estimates of the price of the substitute are positive and significant, although the 2SLS coefficient and its standard error are larger. The OLS estimate of the coefficient of income is not significant whereas the 2SLS estimate is positive and significant which is what we would expect.

The OLS and 2SLS estimates for the supply equation are virtually identical and both satisfy our economic reasoning.

EXERCISE 11.29

(a) The OLS estimates are in column (1) of Table XR11.29. The coefficients of total wages, *W*, and profits, *P*, have positive signs. There are 17 degrees of freedom for this equation; $t_{(0.975,17)} = 2.11$ and $t_{(0.995,17)} = 2.898$. The wages coefficient is significant at the 0.01 level with $t = 19.93$. That for profits is significant at the 0.05 level with $t = 2.12$.

(b) The estimated equation is in column (3) of Table XR11.29. The *F*-test of joint significance of the 5 variables *KLAG*, *G*, *TX*, *TIME*, *ELAG* yields $F = 2.53$. The 95th percentile of $F_{(5,13)}$ is 3.0254. Therefore, at the 5% level we do not find these variables jointly significant.

(c) The estimated equation is in column (4) of Table XR11.29. The *F*-test of joint significance of the 5 variables *KLAG*, *G*, *TX*, *TIME*, *ELAG* yields $F = 3.35$. The 95th percentile of $F_{(5,13)}$ is 3.0254. Therefore, at the 5% level we find these variables jointly significant.

(d) The regression results are in column (5) of Table XR11.29. We see that $\hat{v}_2$ is significant at the 1% level. The joint *F*-test gives $F = 5.60$. The 95th percentile of $F_{(2,15)}$ is 3.6823.

Therefore, at the 5% level we do find these variables jointly significant. We conclude that either W_{1t} or P_t is endogenous, or both are. This is what we should find in a system of simultaneous equations. The endogenous variables are jointly determined.

(e) The 2SLS estimates are in column (2) of Table XR11.29. The primary difference is that profits, *P*, has a much smaller coefficient estimate and is no longer significant. On the other hand, lagged profits, *PLAG*, has a larger coefficient estimate and is significant at the 10% level, using the critical value $t_{(0.95,17)} = 1.74$.

(f) The estimates for this regression are in column (6) of Table XR11.29. The test statistic is $TR^2 = 21(0.4177) = 8.7715$. The chi-square critical value is 7.8147. Thus, we reject the validity of the surplus IV. The test does not indicate which IV is the problem, but government spending, *G*, is statistically significant in the regression. Valid exogenous variables, and IV, should not be related to the 2SLS residuals.

Table XR11.29

	(1) CN OLS	(2) CN 2SLS	(3) W1 RF	(4) P RF	(5) Hausman	(6) Sargan
C	16.2366^{***}	16.5548^{***}	43.4356	50.3844	16.5548^{***}	5.6846
	(12.46)	(11.28)	(1.71)	(1.59)	(15.51)	(0.40)
W	0.7962^{***}	0.8102^{***}			0.8102^{***}	
	(19.93)	(18.11)			(24.90)	
P	0.1929^{*}	0.0173			0.0173	
	(2.12)	(0.13)			(0.18)	
PLAG	0.0899	0.2162	0.8719	0.8025	0.2162^{*}	-0.1885
	(0.99)	(1.81)	(2.09)	(1.55)	(2.49)	(-0.80)
KLAG			-0.1230	-0.2161		-0.0432
			(-1.28)	(-1.81)		(-0.80)
W2			-0.4437	-0.0796		-0.2560
			(-0.22)	(-0.03)		(-0.22)
G			0.8662^{*}	0.4390		-0.5044^{*}
			(2.76)	(1.12)		(-2.84)
TX			-0.6042	-0.9231		0.1397
			(-1.73)	(-2.13)		(0.71)
TIME			0.7136	0.3194		0.1174
			(1.14)	(0.41)		(0.33)
ELAG			0.0953	0.0220		0.1525
			(0.42)	(0.08)		(1.19)
VHAT1					-0.4537	
					(-1.89)	
VHAT2					0.6899^{**}	
					(3.23)	
N	21	21	21	21	21	21

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$