Simultaneous Equations Models

Chapter 11

Prepared by Vera Tabakova, East Carolina University

Chapter 11: Simultaneous Equations Models

- 11.1 A Supply and Demand Model
- 11.2 The Reduced Form Equations
- 11.3 The Failure of Least Squares
- 11.4 The Identification Problem
- 11.5 Two-Stage Least Squares Estimation
- 11.6 An Example of Two-Stage Least Squares Estimation
- 11.7 Supply and Demand at the Fulton Fish Market

11.1 A Supply and Demand Model

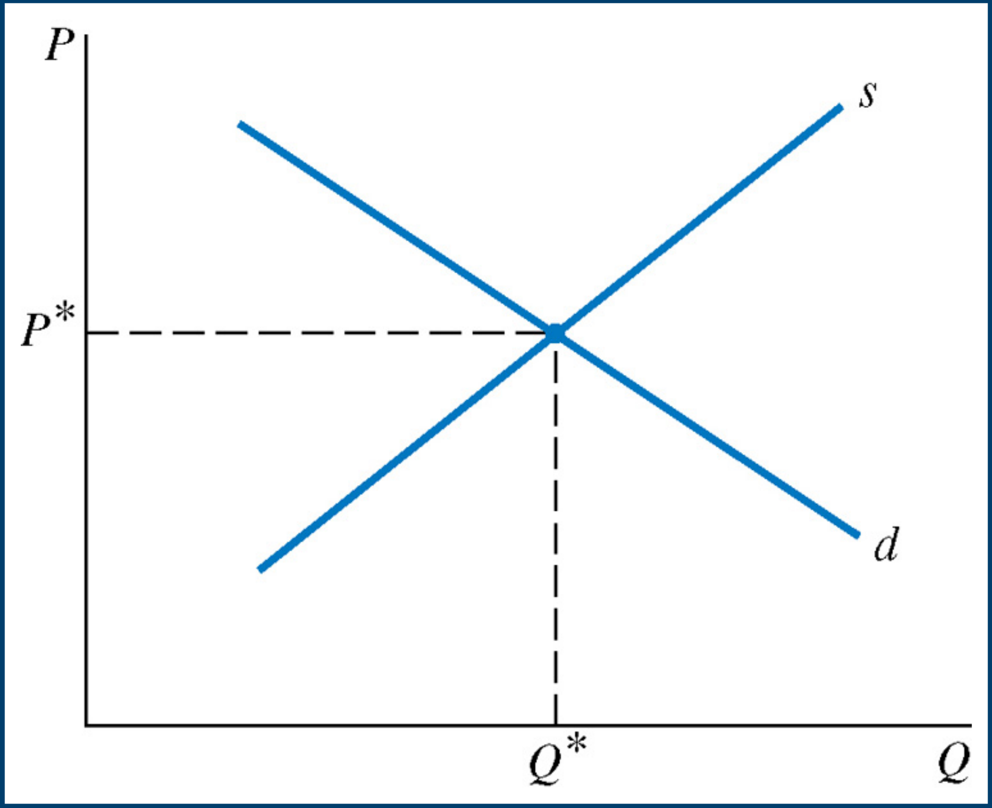

Figure 11.1 Supply and demand equilibrium

11.1 A Supply and Demand Model

$$\text{Demand:} \quad Q = \alpha_1 P + \alpha_2 X + e_d \tag{11.1}$$

$$\text{Supply:} \quad Q = \beta_1 P + e_s \tag{11.2}$$

11.1 A Supply and Demand Model

$$E(e_d) = 0, \quad \operatorname{var}(e_d) = \sigma_d^2$$

$$E(e_s) = 0, \quad \operatorname{var}(e_s) = \sigma_s^2 \tag{11.3}$$

$$\operatorname{cov}(e_d, e_s) = 0$$

11.1 A Supply and Demand Model

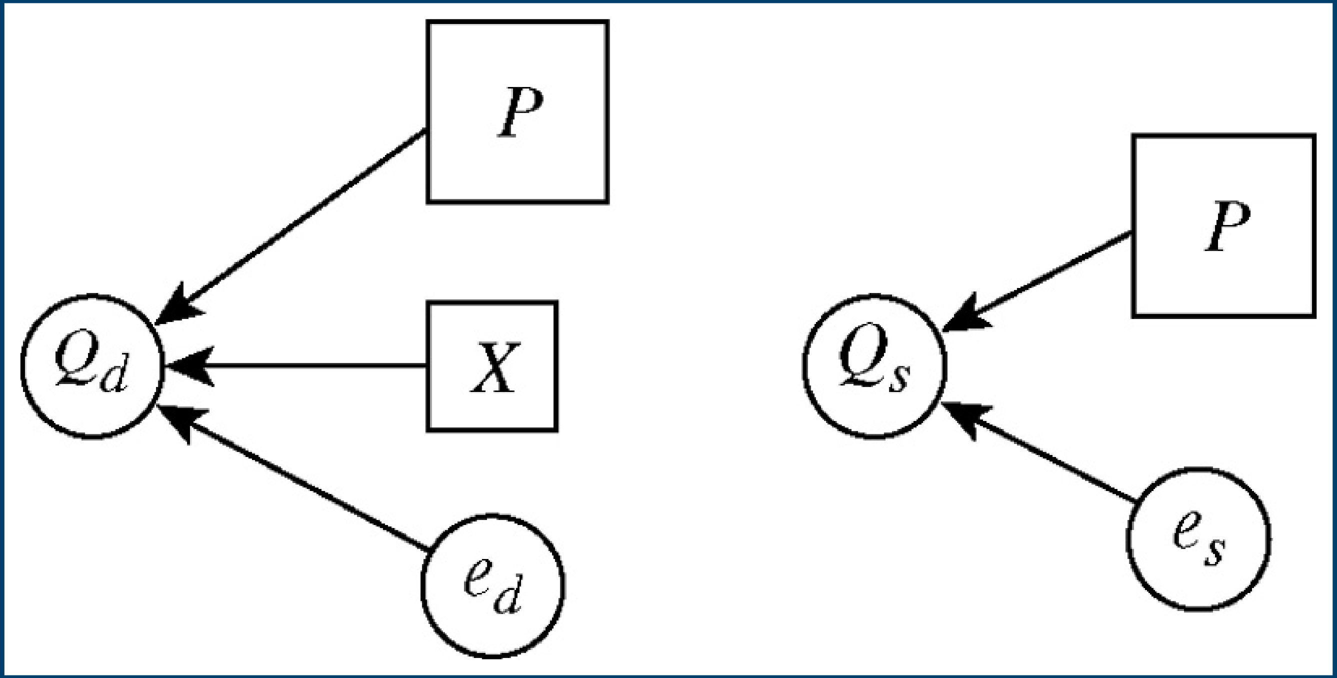

Figure 11.2 Influence diagrams for two regression models

11.1 A Supply and Demand Model

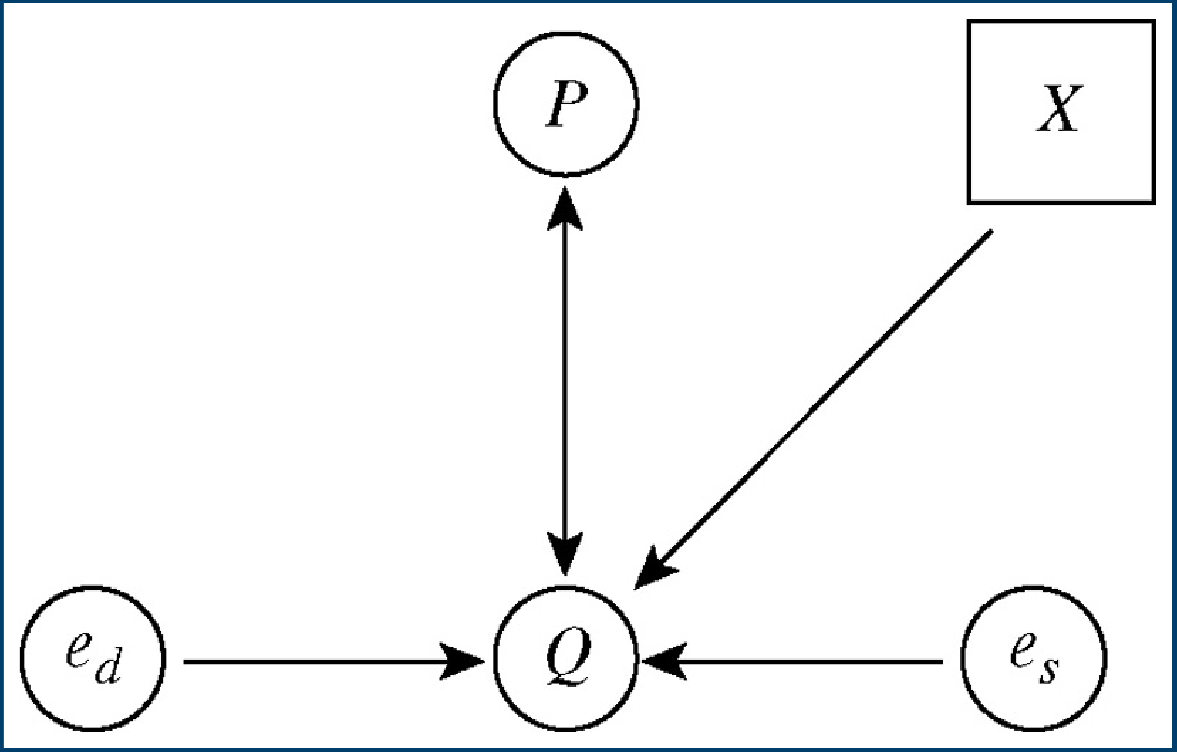

Figure 11.3 Influence diagram for a simultaneous equations model

11.2 The Reduced Form Equations

$$\beta_1 P + e_s = \alpha_1 P + \alpha_2 X + e_d$$

$$P = \frac{\alpha_2}{(\beta_1 - \alpha_1)} X + \frac{e_d - e_s}{(\beta_1 - \alpha_1)} \qquad (11.4)$$

$$= \pi_1 X + v_1$$

11.2 The Reduced Form Equations

$$
\begin{aligned}
Q &= \beta_1 P + e_s \\
&= \beta_1 \left[\frac{\alpha_2}{(\beta_1 - \alpha_1)} X + \frac{e_d - e_s}{(\beta_1 - \alpha_1)} \right] + e_s \\
&= \frac{\beta_1 \alpha_2}{(\beta_1 - \alpha_1)} X + \frac{\beta_1 e_d - \alpha_1 e_s}{(\beta_1 - \alpha_1)} \\
&= \pi_2 X + v_2
\end{aligned}
\tag{11.5}
$$

11.3 The Failure of Least Squares

The least squares estimator of parameters in a structural simultaneous equation is biased and inconsistent because of the correlation between the random error and the endogenous variables on the right-hand side of the equation.

11.4 The Identification Problem

In the supply and demand model given by (11.1) and (11.2)

- the parameters of the demand equation, α_1 and α_2, <u>cannot</u> be consistently estimated by *any* estimation method, but

- the slope of the supply equation, β_1, can be consistently estimated.

11.4 The Identification Problem

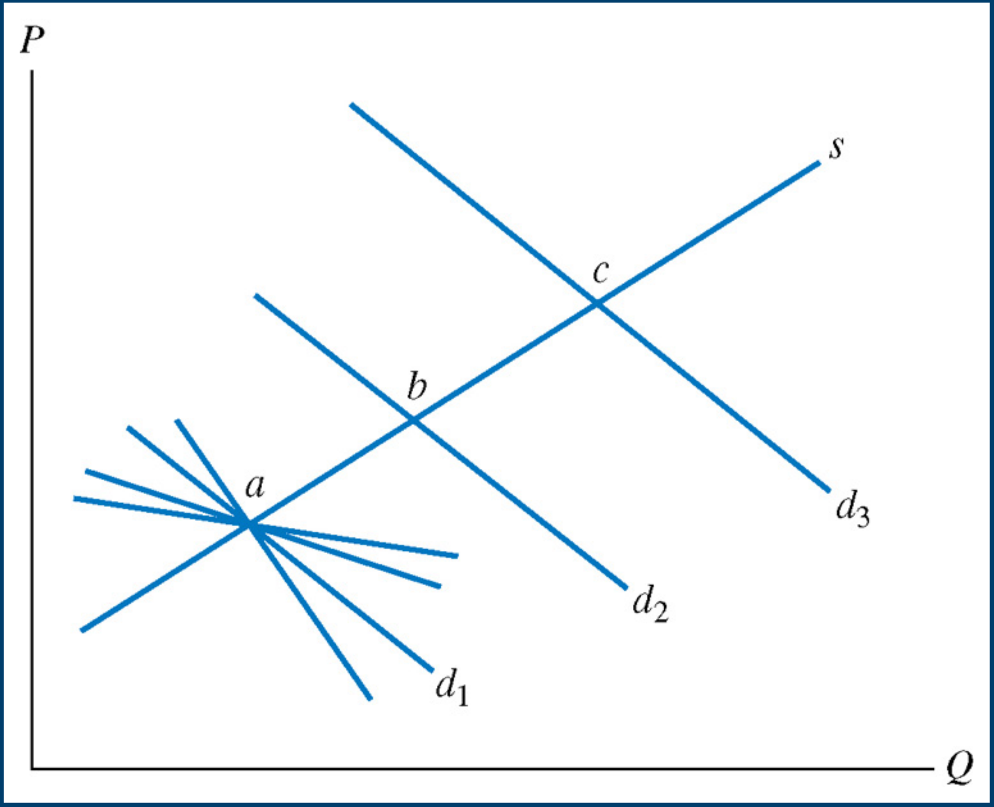

Figure 11.4 The effect of changing income

11.4 The Identification Problem

A Necessary Condition for Identification: In a system of M simultaneous equations, which jointly determine the values of M endogenous variables, at least $M–1$ variables must be absent from an equation for estimation of its parameters to be possible. When estimation of an equation's parameters is possible, then the equation is said to be *identified*, and its parameters can be estimated consistently. If less than $M–1$ variables are omitted from an equation, then it is said to be *unidentified* and its parameters can not be consistently estimated.

11.4 The Identification Problem

Remark: The two-stage least squares estimation procedure is developed in Chapter 10 and shown to be an instrumental variables estimator. The number of instrumental variables required for estimation of an equation within a simultaneous equations model is equal to the number of right-hand-side endogenous variables. Consequently, identification requires that the number of excluded exogenous variables in an equation be at least as large as the number of included right-hand-side endogenous variables. This ensures an adequate number of instrumental variables.

11.5 Two-Stage Least Squares Estimation

$$P = E(P) + v_1 = \pi_1 X + v_1 \qquad (11.6)$$

$$\begin{aligned} Q &= \beta_1 \left[E(P) + v_1 \right] + e_s \\ &= \beta_1 E(P) + \left(\beta_1 v_1 + e_s \right) \\ &= \beta_1 E(P) + e_* \end{aligned} \qquad (11.7)$$

11.5 Two-Stage Least Squares Estimation

$$\hat{P} = \hat{\pi}_1 X$$

$$Q = \beta_1 \hat{P} + \hat{e}_* \tag{11.8}$$

11.5 Two-Stage Least Squares Estimation

Estimating the (11.8) by least squares generates the so-called **two-stage least squares** estimator of β_1, which is consistent and asymptotically normal. To summarize, the *two stages* of the estimation procedure are:

- Least squares estimation of the reduced form equation for P and the calculation of its predicted value $\hat{P}$.
- Least squares estimation of the structural equation in which the right-hand side endogenous variable P is replaced by its predicted value $\hat{P}$.

11.5.1 The General Two-Stage Least Squares Estimation Procedure

$$y_1 = \alpha_2 y_2 + \alpha_3 y_3 + \beta_1 x_1 + \beta_2 x_2 + e_1 \tag{11.9}$$

1. Estimate the parameters of the reduced form equations

$$y_2 = \pi_{12} x_1 + \pi_{22} x_2 + \cdots + \pi_{K2} x_K + v_2$$

$$y_3 = \pi_{13} x_1 + \pi_{23} x_2 + \cdots + \pi_{K3} x_K + v_3$$

by least squares and obtain the predicted values.

11.5.1 The General Two-Stage Least Squares Estimation Procedure

$$\begin{aligned}\hat{y}_2 &= \hat{\pi}_{12}x_1 + \hat{\pi}_{22}x_2 + \cdots + \hat{\pi}_{K2}x_K \\ \hat{y}_3 &= \hat{\pi}_{13}x_1 + \hat{\pi}_{23}x_2 + \cdots + \hat{\pi}_{K3}x_K\end{aligned} \tag{11.10}$$

11.5.1 The General Two-Stage Least Squares Estimation Procedure

2. Replace the endogenous variables, y_2 and y_3, on the right-hand side of the structural (11.9) by their predicted values from (11.10)

$$y_1 = \alpha_2 \hat{y}_2 + \alpha_3 y_3 + \beta_1 x_1 + \beta_2 x_2 + e_1^*$$

Estimate the parameters of this equation by least squares.

11.5.2 The Properties of the Two-Stage Least Squares Estimator

- The 2*SLS* estimator is a biased estimator, but it is consistent.
- In large samples the 2*SLS* estimator is approximately normally distributed.

11.5.2 The Properties of the Two-Stage Least Squares Estimator

- The variances and covariances of the 2*SLS* estimator are unknown in small samples, but for large samples we have expressions for them which we can use as approximations. These formulas are built into econometric software packages, which report standard errors, and *t*-values, just like an ordinary least squares regression program.

11.5.2 The Properties of the Two-Stage Least Squares Estimator

- If you obtain 2*SLS* estimates by applying two least squares regressions using ordinary least squares regression software, the standard errors and *t*-values reported in the *second* regression are *not* correct for the 2*SLS* estimator. Always use specialized 2*SLS* or instrumental variables software when obtaining estimates of structural equations.

11.6 An Example of Two-Stage Least Squares Estimation

$$\text{Demand: } Q_i = \alpha_1 + \alpha_2 P_i + \alpha_3 PS_i + \alpha_4 DI_i + e_i^d \quad (11.11)$$

$$\text{Supply: } Q_i = \beta_1 + \beta_2 P_i + \beta_3 PF_i + e_i^s \quad (11.12)$$

11.6.1 Identification

The rule for identifying an equation is that in a system of M equations at least $M - 1$ variables must be omitted from each equation in order for it to be identified. In the demand equation the variable PF is not included and thus the necessary $M - 1 = 1$ variable is omitted. In the supply equation both PS and DI are absent; more than enough to satisfy the identification condition.

11.6.2 The Reduced Form Equations

$$Q_i = \pi_{11} + \pi_{21} PS_i + \pi_{31} DI_i + \pi_{41} PF_i + v_{i1}$$

$$P_i = \pi_{12} + \pi_{22} PS_i + \pi_{32} DI_i + \pi_{42} PF_i + v_{i2}$$

11.6.2 The Reduced Form Equations

Table 11.1 Representative Truffle Data

OBS	*P*	*Q*	*PS*	*DI*	*PF*
1	29.64	19.89	19.97	2.103	10.52
2	40.23	13.04	18.04	2.043	19.67
3	34.71	19.61	22.36	1.870	13.74
4	41.43	17.13	20.87	1.525	17.95
5	53.37	22.55	19.79	2.709	13.71
		Summary statistics			
Mean	62.72	18.46	22.02	3.53	22.75
Std. Dev.	18.72	4.61	4.08	1.04	5.33

11.6.2 The Reduced Form Equations

Table 11.2a Reduced Form for Quantity of Truffles (Q)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.8951	3.2434	2.4342	0.0221
PS	0.6564	0.1425	4.6051	0.0001
DI	2.1672	0.7005	3.0938	0.0047
PF	−0.5070	0.1213	−4.1809	0.0003

11.6.2 The Reduced Form Equations

Table 11.2b Reduced Form for Price of Truffles (*P*)

Variable	Coefficient	Std. Error	*t*-Statistic	Prob.
C	−32.5124	7.9842	−4.0721	0.0004
PS	1.7081	0.3509	4.8682	0.0000
DI	7.6025	1.7243	4.4089	0.0002
PF	1.3539	0.2985	4.5356	0.0001

11.6.3 The Structural Equations

$$\hat{P}_i = \hat{\pi}_{12} + \pi_{22}PS_i + \hat{\pi}_{32}DI_i + \pi_{42}PF_i$$

$$= -32.512 + 1.708PS_i + 7.602DI_i + 1.354PF_i$$

11.6.3 The Structural Equations

Table 11.3a *2SLS* **Estimates for Truffle Demand**

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	−4.2795	5.5439	−0.7719	0.4471
P	−0.3745	0.1648	−2.2729	0.0315
PS	1.2960	0.3552	3.6488	0.0012
DI	5.0140	2.2836	2.1957	0.0372

11.6.3 The Structural Equations

Table 11.3b 2SLS Estimates for Truffle Supply

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	20.0328	1.2231	16.3785	0.0000
P	0.3380	0.0249	13.5629	0.0000
PF	−1.0009	0.0825	−12.1281	0.0000

11.7 Supply and Demand at the Fulton Fish Market

$$\ln\left(QUAN_t\right) = \alpha_1 + \alpha_2 \ln\left(PRICE_t\right) + \alpha_3 MON_t + \alpha_4 TUE_t + \alpha_5 WED_t + \alpha_6 THU_t + e_t^d \tag{11.13}$$

$$\ln\left(QUAN_t\right) = \beta_1 + \beta_2 \ln\left(PRICE_t\right) + \beta_3 STORMY_t + e_t^s \tag{11.14}$$

11.7.1 Identification

The necessary condition for an equation to be identified is that in this system of $M = 2$ equations, it must be true that at least $M - 1 = 1$ variable must be omitted from each equation. In the demand equation the weather variable *STORMY* is omitted, while it does appear in the supply equation. In the supply equation, the four daily dummy variables that are included in the demand equation are omitted.

11.7.2 The Reduced Form Equations

$$\ln\left(QUAN_t\right) = \pi_{11} + \pi_{21}MON_t + \pi_{31}TUE_t + \pi_{41}WED_t + \pi_{51}THU_t + \pi_{61}STORMY_t + v_{t1} \qquad (11.15)$$

$$\ln\left(PRICE_t\right) = \pi_{12} + \pi_{22}MON_t + \pi_{32}TUE_t + \pi_{42}WED_t + \pi_{52}THU_t + \pi_{62}STORMY_t + v_{t2} \qquad (11.16)$$

11.7.2 The Reduced Form Equations

Table 11.4a Reduced Form for ln(Quantity) Fish

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.8101	0.1470	59.9225	0.0000
STORMY	−0.3878	0.1437	−2.6979	0.0081
MON	0.1010	0.2065	0.4891	0.6258
TUE	−0.4847	0.2011	−2.4097	0.0177
WED	−0.5531	0.2058	−2.6876	0.0084
THU	0.0537	0.2010	0.2671	0.7899

11.7.2 The Reduced Form Equations

Table 11.4b Reduced Form for ln(Price) Fish

Variable	Coefficient	Std. Error	*t*-Statistic	Prob.
C	−0.2717	0.0764	−3.5569	0.0006
STORMY	0.3464	0.0747	4.6387	0.0000
MON	−0.1129	0.1073	−1.0525	0.2950
TUE	−0.0411	0.1045	−0.3937	0.6946
WED	−0.0118	0.1069	−0.1106	0.9122
THU	0.0496	0.1045	0.4753	0.6356

11.7.2 The Reduced Form Equations

- To identify the supply curve the daily dummy variables must be jointly significant. This implies that at least one of their coefficients is statistically different from zero, meaning that there is at least one significant shift variable in the demand equation, which permits us to reliably estimate the supply equation.
- To identify the demand curve the variable *STORMY* must be statistically significant, meaning that supply has a significant shift variable, so that we can reliably estimate the demand equation.

11.7.2 The Reduced Form Equations

$$\widehat{\ln}\left(PRICE_t\right) = \hat{\pi}_{12} + \pi_{22}MON_t + \hat{\pi}_{32}TUE_t + \pi_{42}WED_t + \hat{\pi}_{52}THU_t + \pi_{62}STORMY_t$$

$$\widehat{\ln}\left(PRICE_t\right) = \hat{\pi}_{12} + \pi_{62}STORMY_t$$

11.7.3 Two-Stage Least Squares Estimation of Fish Demand

Table 11.5 *2SLS* Estimates for Fish Demand

Variable	Coefficient	Std. Error	*t*-Statistic	Prob.
C	8.5059	0.1662	51.1890	0.0000
ln(*PRICE*)	−1.1194	0.4286	−2.6115	0.0103
MON	−0.0254	0.2148	−0.1183	0.9061
TUE	−0.5308	0.2080	−2.5518	0.0122
WED	−0.5664	0.2128	−2.6620	0.0090
THU	0.1093	0.2088	0.5233	0.6018

Keywords

- endogenous variables
- exogenous variables
- Fulton Fish Market
- identification
- reduced form equation
- reduced form errors
- reduced form parameters
- simultaneous equations
- structural parameters
- two-stage least squares

Chapter 11 Appendix

- Appendix 11A An Algebraic Explanation of the Failure of Least Squares

Appendix 11A
An Algebraic Explanation of the Failure of Least Squares

$$\text{cov}(P, e_s) = E\left[P - E(P)\right]\left[e_s - E(e_s)\right]$$

$$= E(Pe_s) \qquad \text{[since } E(e_s) = 0\text{]}$$

$$= E\left[\pi_1 X + v_1\right] e_s \qquad \text{[substitute for } P\text{]}$$

$$= E\left[\frac{e_d - e_s}{\beta_1 - \alpha_1}\right] e_s \qquad \text{[since } \pi_1 X \text{ is exogenous]} \qquad (11A.1)$$

$$= \frac{-E(e_s^2)}{\beta_1 - \alpha_1} \qquad \text{[since } e_d, e_s \text{ assumed uncorrelated]}$$

$$= \frac{-\sigma_s^2}{\beta_1 - \alpha_1} < 0$$

Appendix 11A
An Algebraic Explanation of the Failure of Least Squares

$$b_1 = \frac{\sum P_i Q_i}{\sum P_i^2} \qquad (11A.2)$$

$$b_1 = \frac{\sum P_i \left(\beta_1 P_i + e_{si}\right)}{\sum P_i^2} = \beta_1 + \sum \left(\frac{P_i}{\sum P_i^2} \right) e_{si} = \beta_1 + \sum h_i e_{si} \qquad (11A.3)$$

$$E\left(b_1\right) = \beta_1 + \sum E\left(h_i e_{si}\right) \neq \beta_1$$

Appendix 11A
An Algebraic Explanation of the Failure of Least Squares

$$PQ = \beta_1 P^2 + Pe_s$$

$$E\left(PQ\right) = \beta_1 E\left(P^2\right) + E\left(Pe_s\right)$$

$$\beta_1 = \frac{E\left(PQ\right)}{E\left(P^2\right)} - \frac{E\left(Pe_s\right)}{E\left(P^2\right)}$$

$$b_1 = \frac{\sum Q_i P_i / N}{\sum P_i^2 / N} \rightarrow \frac{E\left(PQ\right)}{E\left(P^2\right)} = \beta_1 + \frac{E\left(Pe_s\right)}{E\left(P^2\right)} = \beta_1 - \frac{\sigma_s^2 \big/ \left(\beta_1 - \alpha_1\right)}{E\left(P^2\right)} < \beta_1$$