THE CLASSICAL SIMULTANEOUS-EQUATIONS MODEL

Consider the system

$$y_1 = y_2\gamma_{21} + \varepsilon_1 : \qquad \text{The Demand Equation,} \tag{1}$$

$$y_2 = y_1\gamma_{12} + x\beta + \varepsilon_2 : \qquad \text{The Supply Equation.} \tag{2}$$

Here

y_1 represents the quantity of popcorn consumed and produced
y_2 represents the price of popcorn, and
x represents the cost of maize.

These variables, which are deviations from mean values, have expected values of zero. The effect of taking deviations is to simplify the algebra; for the intercept terms are thereby eliminated from the equations.

Another feature to take note of is the use of indices. The subscripts on the parameter γ_{21}, for example, indicate a mapping from y_2, which is the dependent variable of the second equation, to y_1, which is the dependent variable of the first equation. We shall assume that the disturbances ε_1 and ε_2 are independent of the variable x, which is described as an exogenous variable to indicate that it is generated in a context which lies outside the model.

The notion which lies behind this model is that the consumers of popcorn, whose behaviour is represented by the demand equation, respond to the price of popcorn, whereas the producers, whose behaviour is represented by the supply equation, set the price in view of the demand for their product and in view of their costs of production. The market is in a state of equilibrium where the quantity produced is equal to the quantity consumed.

Although the cost of maize is not the only cost of production, we shall assume, for the moment, that it is the only one which varies. The other costs, which are fixed, will have an effect which is subsumed in an intercept term which has been eliminated. The factors, other than price, which determine the demand for popcorn are likewise assumed to be constant and are subsumed in another intercept.

There are markets where output is ostensibly determined by supply factors and where the price adjusts to ensure that all of the output is sold. Some markets for agricultural produce are examples. In such cases, we might wish to place y_2 on the LHS of equation (1) and y_1 on the LHS of equation (2). However, there is no need to adapt the equations; for, in a situation of equilibrium, where both sides of the market are reconciled, it cannot be said that either is peculiarly responsible for the price of the item or for the quantity produced.

The economist Alfred Marshall, who may be credited with formulating much of modern microeconomic theory, likened the supply and demand equations of a market in equilibrium to the blades of scissors. It is no more appropriate to ask which of the equations determines the price and which of them determines the quantity than it is to ask which of blades is cutting a sheet of paper.

It follows that, given a state of equilibrium, the choice of dependent variables in equations (1) and (2) is arbitrary. Nevertheless, the choice should reflect our understanding of how the two parties might behave in the process of achieving the equilibrium.

Now let us consider using a method of moments in estimating the parameters of the system. This entails finding expressions for the parameters which are in terms of the moments of the observable variables. Once these expressions have been found, we may consider replacing the theoretical moments by the empirical counterparts to derive the estimating equations.

To find an expression for γ_{21}, we multiply the demand equation (1) by x and we take expectations. This gives

$$E(xy_1) = E(xy_2)\gamma_{21}, \tag{3}$$

from which we see that

$$\gamma_{21} = \frac{E(xy_1)}{E(xy_2)}. \tag{4}$$

When we attempt to apply the same method to the supply equation (2), we find that there is not sufficient information to determine the two remaining parameters. Multiplying the equation by x and taking expectations leads to

$$E(xy_2) = E(xy_1)\gamma_{12} + E(x^2)\beta. \tag{5}$$

If we seek another equation by multiplying equation (2) by y_1 and by taking expectations, then we shall introduce another unknown quantity which is the nonzero moment $E(y_1\varepsilon_2)$:

$$E(y_1y_2) = E(y_1^2)\gamma_{12} + E(y_1x)\beta + E(y_1\varepsilon_2). \tag{6}$$

We have an equal lack of success in attempting to form an estimating equation by multiplying the equation (2) by y_2 and taking expectations.

In view of its role in generating estimating equations, the exogenous variable x is apt to be described as an instrumental variable. The problem of the supply equation is the impossibility of estimating two parameters γ_{12} and β when there is only one instrumental variable. The two parameters are said to be unidentified. A necessary condition for the identification of the parameters of any equation is that their number should not exceed the number of the available instrumental variables.

Example. An attempt to estimate equation (1) by the ordinary method of regression would lead to a biased estimator. The method is inappropriate because the disturbance term ε_1 is correlated with the variable y_2 which is the putative regressor. This correlation is evident in the fact that we can trace a connection running from ε_1 to y_1, within equation (1) and thence from y_1 to y_2 through equation (2). To find an expression for the covariance of y_2 and ε_1, we may substitute equation (1) into equation (2) to give

$$y_2 = (y_2\gamma_{21} + \varepsilon_1)\gamma_{12} + x\beta + \varepsilon_2. \tag{7}$$

Rearranging this gives

$$y_2 = \frac{x\beta}{1-\gamma_{21}\gamma_{12}} + \frac{\varepsilon_1\gamma_{12} + \varepsilon_2}{1-\gamma_{21}\gamma_{12}}. \tag{8}$$

Therefore

$$C(y_2,\varepsilon_1)=\frac{V(\varepsilon_1)\gamma_{12}+C(\varepsilon_1,\varepsilon_2)}{1-\gamma_{21}\gamma_{12}}. \tag{9}$$

Now let us consider some circumstances which would enable us to estimate both the supply equation and the demand equation. Consider the system

$$y_1=y_2\gamma_{21}+x_1\beta_{11}+\varepsilon_1: \quad \text{The Demand Equation,} \tag{10}$$

$$y_2=y_1\gamma_{12}+x_2\beta_{22}+\varepsilon_2: \quad \text{The Supply Equation.} \tag{11}$$

Compared with equation (1), the revised demand equation incorporates an extra variable x_1 which represents the price of candy floss. If candy floss and popcorn are attractive to the same people, then one may expect the demand for popcorn to fall if the price of candy floss is reduced. Given the additional instrumental variable, we can now estimate the parameters of both equations, which have an identical structure.

To derive estimating equations for the parameters of the demand equation, we multiply the latter in turn by x_1 and x_2 and we take expectations. The results are

$$E(x_1y_1)=E(x_1y_2)\gamma_{21}+E(x_1^2)\beta_{11}, \tag{12}$$

$$E(x_2y_1)=E(x_2y_2)\gamma_{21}+E(x_2x_1)\beta_{11}. \tag{13}$$

These equations serve simultaneously to determine both γ_{21} and β_{11}. Their empirical counterparts, which are derived by replacing the theoretical moments by the corresponding sample moments serve as estimating equations for the parameters. We may use exactly the same device in estimating the supply equation.

We must avoid the false impression that new variables may be introduced at will. The presence, in the demand equation, of the price of candy floss can be justified only if the latter has an active effect on the level of demand. That is to say, x_1 must vary within the sample of observations if it is to assist in identifying the parameters of the model. If this price is constant, then its effect will be subsumed, as before, in the intercept term of the demand equation. It is also required that the price of candy floss should not enter the supply equation for popcorn, which seems plausible.

Now let us consider a third possibility which puts a different construction on the problem of estimation. Consider the system

$$y_1=y_2\gamma_{21}+\varepsilon_1: \quad \text{The Demand Equation,} \tag{14}$$

$$y_2=y_1\gamma_{12}+x_1\beta_{12}+x_2\beta_{22}+\varepsilon_2: \quad \text{The Supply Equation.} \tag{15}$$

Here

x_1 represents the cost of maize, and
x_2 represents the cost of pink sugar.

The price of candy floss no longer enters the demand equation; and we might imagine that the makers of candy floss no longer occupy their stalls on the seaside promenade. There are now two instrumental variables which can serve to identify the demand equation. Thus

$$E(x_1y_1) = E(x_1y_2)\gamma_{21}, \tag{16}$$

$$E(x_2y_1) = E(x_2y_2)\gamma_{21}. \tag{17}$$

The parameter γ_{21} is said to be overidentified.

In practice, when we replace the theoretical moments by their empirical counterparts, the estimates which are generated by the two equations are liable to differ. Since both of the estimates are valid, we should attempt, in the interests of statistical efficiency, to combine them.

In order to resolve the conflict between the two estimates of γ_{21} we shall resort to a procedure which involves the errors-in-variables estimator. We begin by deriving the so-called reduced-form equations. Substituting equation (15) into equation (14) gives

$$y_1 = (y_1\gamma_{12} + x_1\beta_{12} + x_2\beta_{22})\gamma_{21} + (\varepsilon_1 + \varepsilon_2\gamma_{21}). \tag{18}$$

On rearranging this we get

$$\begin{aligned} y_1 &= \frac{(x_1\beta_{12} + x_2\beta_{22})\gamma_{21}}{1-\gamma_{12}\gamma_{21}} + \frac{\varepsilon_1 + \varepsilon_2\gamma_{21}}{1-\gamma_{12}\gamma_{21}} \\ &= x_1\pi_{11} + x_2\pi_{21} + \eta_1 \\ &= \xi_1 + \eta_1, \end{aligned} \tag{19}$$

which is the so-called reduced-form equation for y_1. Substituting equation (14) into equation (15) gives

$$y_2 = y_2\gamma_{21}\gamma_{12} + x_1\beta_{12} + x_2\beta_{22} + (\varepsilon_2 + \varepsilon_1\gamma_{12}). \tag{20}$$

On rearranging this we get

$$\begin{aligned} y_2 &= \frac{x_1\beta_{12} + x_2\beta_{22}}{1-\gamma_{21}\gamma_{12}} + \frac{\varepsilon_2 + \varepsilon_1\gamma_{12}}{1-\gamma_{21}\gamma_{12}} \\ &= x_1\pi_{12} + x_2\pi_{22} + \eta_2 \\ &= \xi_2 + \eta_2, \end{aligned} \tag{21}$$

which is the reduced-form equation for y_2. On comparing equations (19) and (21), it can be seen that

$$y_1 - \eta_1 = (y_2 - \eta_2)\gamma_{21}. \tag{22}$$

This is the equation of an errors-in-variables model wherein one of the parameters has been normalised with a value of -1.

We can use the errors-in-variables estimator for γ_{21} provided that we can find values for the variances and covariances for the errors η_1 and η_2 which are, in fact, the disturbances of the reduced-form regression equations. Let

$$\begin{aligned} &h_{1t} = y_{1t} - x_{1t}\hat{\pi}_{11} - x_{2t}\hat{\pi}_{21} \quad \text{and} \\ &h_{2t} = y_{2t} - x_{1t}\hat{\pi}_{12} - x_{2t}\hat{\pi}_{22} \end{aligned} \tag{23}$$

be the residuals from using the ordinary method of regression in fitting the reduced-from equations to a sample of T observations. Then the moments of the reduced-form disturbances may be estimated as follows:

$$\begin{aligned} \hat{\omega}_{11} &= \frac{1}{T}\sum_{t=1}^{T} h_{1t}^2, \\ \hat{\omega}_{22} &= \frac{1}{T}\sum_{t=1}^{T} h_{2t}^2, \\ \hat{\omega}_{12} &= \frac{1}{T}\sum_{t=1}^{T} h_{1t}h_{2t}. \end{aligned} \tag{24}$$

Using equation (15) as a model, we can now construct an estimating equation for γ_{21} in the form of

$$\left\{ \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} - \lambda \begin{bmatrix} \hat{\omega}_{11} & \hat{\omega}_{12} \\ \hat{\omega}_{21} & \hat{\omega}_{22} \end{bmatrix} \right\} \begin{bmatrix} -1 \\ \gamma_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{25}$$

This gives

$$\hat{\gamma}_{21} = \frac{s_{11} - \lambda\hat{\omega}_{11}}{s_{12} - \lambda\hat{\omega}_{12}} = \frac{s_{21} - \lambda\hat{\omega}_{21}}{s_{22} - \lambda\hat{\omega}_{22}}. \tag{26}$$

The value of λ which guarantees the equality above, is found by solving the determinental equation

$$0 = \text{Det}\begin{bmatrix} s_{11} - \lambda\hat{\omega}_{11} & s_{12} - \lambda\hat{\omega}_{12} \\ s_{21} - \lambda\hat{\omega}_{21} & s_{22} - \lambda\hat{\omega}_{22} \end{bmatrix}, \tag{27}$$

which is a quadratic equation. The root which is closest to unity is taken. As the various empirical moments tend to their true values, so λ will tend to unity.

There are other ways of estimating the parameter which become virtually equivalent to the errors-in-variables method when the sample size is large. One possibility is to use a system which is modelled on equation (23):

$$\left\{ \begin{bmatrix} s_{11} - \hat{\omega}_{11} & s_{12} - \hat{\omega}_{12} \\ s_{21} - \hat{\omega}_{21} & s_{22} - \hat{\omega}_{22} \end{bmatrix} - \mu \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\} \begin{bmatrix} -1 \\ \gamma_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{28}$$

In comparison with equation (25), it can be seen that λ has been set to unity. Since λ is no longer available for the purpose of rendering the equations algebraically consistent, a new factor μ has been introduced to perform this task. Whereas λ will tend to unity with the convergence of the moments to the true values, the value of μ will to zero.

The solution of equation (28) is

$$\hat{\gamma}_{21} = \frac{s_{21} - \hat{\omega}_{21}}{s_{22} - \hat{\omega}_{22}}. \tag{29}$$

This is, in fact, the so-called two-stage least-squares estimator of γ_{21}; and it differs from the ordinary least-squares estimator by virtue of the adjustments which are made to the moments s_{21} and s_{22}.

Two-Stage Least Squares and Limited-Information Maximum Likelihood.

The estimator of the demand equation which we have derived under the guise of the errors-in-variables model was originally derived as the limited-information maximum-likelihood (LIML) estimator by Anderson and Rubin in 1949, when they were members of the Cowles Commission for Research in Economics. The Commission consisted of a group of statisticians and economists whose research was funded by the American industrialist Alfred Cowles. It is arguable the era of modern econometrics began with the work of the Commission.

The derivation of the LIML estimator was a *tour de force.* Its complexity was due in part to the the fact that the likelihood function of a full simultaneous-equation model was taken as a starting point. An alternative derivation, which was no less complicated, was provided shortly afterward by Hood and Koopmans. The inaccessibility of both these derivations deterred econometricians from using the estimator. It was not until the alternative two-stage least-squares estimator was invented independently by Theil and Basmann in the late 1950's that the techniques of simultaneous-equation estimation began to be applied.

The affinity of the 2SLS and the LIML estimators is not evident from a comparison of the original derivations. Nor might it be clear to someone familiar with the 2SLS estimator that it corresponds to what is presented under (28) and (29). Therefore we shall give a version of the familiar derivation before showing how the equivalence may be demonstrated.

The point of departure for the original derivation of the 2SLS estimator is the recognition that, in a structural equation such as (14), the disturbance term is liable to be correlated with some of the variables on the RHS. We have already demonstrated, in an example, the correlation between y_2 and ε_1 within equation (1), which is equation (14) in a different context.

The question arising is how we might purge the variable y_2 of the component which is correlated with ε_1. An effective way, if it were available, would be to replace y_2 by the predicted value $\xi_2 = x_1\pi_{12} + x_2\pi_{22}$ which comes from the reduced-form equation. In fact, by substituting the reduced-form expression for y_2 given by (21) into the equation (14), we obtain

$$
(30)\qquad \begin{aligned} y_1 &= y_2\gamma_{21} + \varepsilon_1 \\ &= (x_1\pi_{12} + x_2\pi_{22})\gamma_{21} + (\varepsilon_1 + \eta_2\gamma_{21}) \\ &= \xi_2\gamma_{21} + \zeta_1. \end{aligned}
$$

The composite disturbance term $\zeta_1 = \varepsilon_1 + \eta_2\gamma_{21}$ is clearly uncorrelated with ξ_2 since ε_1 and η_2 are uncorrelated with x_1 and x_2. Therefore a consistent estimator of γ_{21} would be obtained from the regression of y_1 on ξ_2.

In fact, we cannot put the unknown value of ξ_2 in place of y_2, and we have to make do with its estimate $\hat{y}_2 = x_1\hat{\pi}_{12} + x_2\hat{\pi}_{22}$ which can be expected to converge to ξ_2 as the sample size increases. The resulting estimator of γ_{21} is

$$
(31)\qquad \begin{aligned} \hat{\gamma}_{21} &= \frac{\sum \hat{y}_{2t}y_{1t}}{\sum \hat{y}_{2t}^2} \\ &= \frac{\sum \hat{y}_{2t}\hat{y}_{1t}}{\sum \hat{y}_{2t}^2}. \end{aligned}
$$

The second equality depends upon the result that $\sum \hat{y}_{2t}y_{1t} = \sum \hat{y}_{2t}\hat{y}_{1t}$. The latter is due to the fact that the reduced-from disturbance h_1 within $y_1 = \hat{y}_1 + h_1$ is uncorrelated with the reduced-form regressors x_1 and x_2 and hence with $\hat{y}_2 = x_1\hat{\pi}_{12} + x_2\hat{\pi}_{22}$.

The equivalence between the expression for the 2SLS estimator under (31) and the expression under (29), follows from the identities

$$
(32)\qquad \begin{aligned} \frac{1}{T}\sum_{t=1}^{T} y_{2t}^2 &= \frac{1}{T}\sum_{t=1}^{T} \hat{y}_{2t}^2 + \frac{1}{T}\sum_{t=1}^{T} h_{2t}^2 \quad \text{and} \\ \frac{1}{T}\sum_{t=1}^{T} y_{1t}y_{2t} &= \frac{1}{T}\sum_{t=1}^{T} \hat{y}_{1t}\hat{y}_{2t} + \frac{1}{T}\sum_{t=1}^{T} h_{1t}h_{2t}. \end{aligned}
$$

Using the definitions of (14) and (24), and remembering that the varaibles are in deviation form, we see that these can be rewritten as

$$
(33)\qquad \begin{aligned} \frac{1}{T}\sum_{t=1}^{T} \hat{y}_{2t}^2 &= s_{22} - \hat{\omega}_{22} \quad \text{and} \\ \frac{1}{T}\sum_{t=1}^{T} \hat{y}_{1t}\hat{y}_{2t} &= s_{12} - \hat{\omega}_{12}; \end{aligned}
$$

and the equivalence of (29) and (31) follows immediately