2 Simultaneous equations model

Hill's *U/G Econometrics* ch 14 and EViews ch 23. The latter covers several models and is quite technical.

- The models considered so far have consisted of a single relationship with one response variable y ; the response could take different forms, including ordinary regression, probit and Tobit.

- Simultaneous equations models consist of at least two relationships and contain at least two response variables.

- Once again least squares is not appropriate in these models ...

The econometricians of the 1940/70s developed statistical techniques for *systems* because the basic economic models of the time were *systems* of equations:

- Micro: partial equilibrium market models, consisting of supply and demand equations.

- Macro: Keynesian models, consisting of consumption function, investment function, money demand equation etc.

The *Simultaneous Equations Model* (covering a specification and a set of estimation and testing methods) was developed by a research group the *Cowles Commission* in Chicago as part of a general review of how statistical theory should be applied to economics.

Haavelmo's Probability Approach in Econometrics, *Econometrica* (1944) is the classic statement of the Cowles approach.

This emphasised

- The need for economic theory in empirical work

- The need for a stochastic model to underpin the statistical analysis

- The centrality of a particular type of stochastic model: the SEM

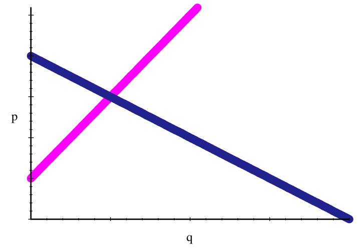
Cowles had a great influence on economists' conceptions of the place of empirical work in economics and for decades the simultaneous equations model was the central model in econometrics.

Supply and Demand model (14.2).

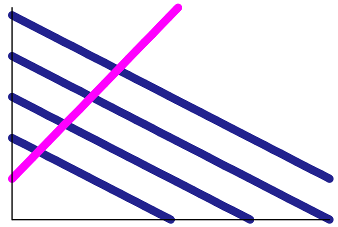
The simplest textbook model describing the working of a market consists of demand and supply equations

$$q_t = \alpha^d p_t : q_t = \alpha^s p_t$$

(I have omitted intercepts to simplify the algebra below.)



texbook diagram


income changing

The second diagram shows the model extended to include a variable y (income) which affects demand. The demand curve shifts as income changes:

$$q_t = \alpha^d p_t + \beta^d y_t.$$

Random disturbances to demand (ε_t^d) would also shift the original line, to produce

$$q_t = \alpha^d p_t + \varepsilon_t^d$$

Putting both the external variable and the disturbance into the demand function and introducing supply disturbances (ε_t^s) and making the usual assumptions about the disturbances we get a full simultaneous equations model

$$
\begin{aligned}
q_t &= \alpha^d p_t + \beta^d y_t + \varepsilon_t^d \\
q_t &= \alpha^s p_t + \varepsilon_t^s \\
E\varepsilon_t^d &= 0;\ E\varepsilon_t^s = 0; \\
var\varepsilon_t^d &= \sigma_d^2;\ var\varepsilon_t^s = \sigma_s^2.
\end{aligned}
$$

We will assume there is no serial correlation or heteroscedasticity in the errors. (These complications can be introduced....)

- The logic of the model is that y and the two errors ε^d and ε^s are given from outside the market and then the market works to determine p and q.

- In terms of observable variables there are 2 dependent variables (p and q) and one independent variable y. In SEM terminology dependent variables are called *endogenous* and independent called *exogenous*.

- In the regression model an equation contains *only one* independent variable. There is *no* reason to suppose that applying ordinary regression methods to the supply and demand equations will produce sensible answers.

- Before we start looking for *good* ways of estimating the parameters $\alpha^d, \alpha^s, \beta^d$, etc. we should stop and ask whether they *can* be estimated. This is called the *identification problem*.

2.1 The Identification Problem

Hill's *U/G Econometrics* ch 14.5.

There is a difficulty with estimating the supply and demand model given above. It can be seen in several ways.

- Consider the "dream" situation where the disturbances are always zero. If we have observations we expect to estimate the slopes without error. The 2nd diagram above illustrates what happens. We *can* estimate α^s without error because the shifting demand line traces out the supply line. However there is nothing to tell us about the slope of the demand lines. We cannot estimate α^d.

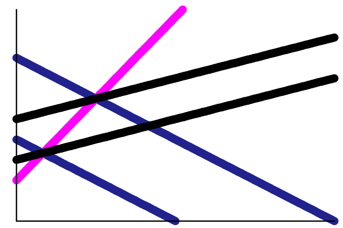
- Staying with the dream but switching to algebra, we observe solutions to the equations

$$
\begin{aligned}
q_t &= \alpha^d p_t + \beta^d y_t \\
q_t &= \alpha^s p_t
\end{aligned}
$$

  Suppose now we construct a new system where the first equation is formed by combining these 2 equations–e.g. averaging them–but the second equation is left unchanged

$$
\begin{aligned}
q_t &= \frac{1}{2}(\alpha^d + \alpha^s) p_t + \frac{1}{2}\beta^d y_t \\
q_t &= \alpha^s p_t
\end{aligned}
$$

  This new system will have the *same* solution as the old system, although the slopes in the demand equation are *different*. In the picture the black lines are the averages:



true is black or blue?

- Thus if we just know the intersections of supply and demand we can reconstruct the supply but not the demand.

- The parameter α^s is said to be *identified* but α^d and β^d are *not identified*.

- Nothing is changed if we hang errors on these equations. We do the details later.

- What *does change* the prospects for making inferences about the slopes is the appearance of a variable in the supply equation that does not appear in the demand equation.

- If we have

$$q_t = \alpha^d p_t + \beta^d y_t$$
$$q_t = \alpha^s p_t + \beta^s w_t$$

so that weather in Brazil influences the supply of coffee we cannot combine the equations to form a new system without introducing w or y where we know it does *not* belong.

2.2 Algebra–the reduced form

Hill's *U/G Econometrics* ch 14.3.

The system of equations given by "economic theory" constitute the *structural form* and the component equations (e.g. demand and supply equations) are called *structural* equations.

If we *solve* the structural equations for the endogenous variables (p and q in our example) we obtain the *reduced form*.

In the model

$$\begin{aligned} q_t &= \alpha^d p_t + \beta^d y_t + \varepsilon_t^d \\ q_t &= \alpha^s p_t + \beta^s w_t + \varepsilon_t^s \end{aligned}$$

we equate the 2 right hand side expressions to give an equation for p_t and then use it to get an expression for q_t :

$$p_t = \frac{(\varepsilon_t^s + \beta^s w_t) - (\varepsilon_t^d + \beta^d y_t)}{\alpha^d - \alpha^s}$$
$$q_t = \frac{\alpha^d(\varepsilon_t^s + \beta^s w_t) - \alpha^s(\varepsilon_t^d + \beta^d y_t)}{\alpha^d - \alpha^s}$$

The equations describe how the exogenous variables y and w and the error terms determine the endogenous variables p and q.

The "dream" situations of no errors can be illustrated

- Model without w

$$p_t = \frac{-\beta^d y_t}{\alpha^d - \alpha^s} : q_t = \frac{-\alpha^s \beta^d y_t}{\alpha^d - \alpha^s}$$

  From these 2 equations with 2 y_t coefficients we can get α^s but not α^d and β^d. The demand parameters are not identified

- Model with w

$$p_t = \frac{\beta^s w_t - \beta^d y_t}{\alpha^d - \alpha^s} : q_t = \frac{\alpha^d \beta^s w_t - \alpha^s \beta^d y_t}{\alpha^d - \alpha^s}$$

  From these 2 equations with 2 y_t coefficients plus 2 w_t coefficients we can get everything: $\alpha^d, \alpha^s, \beta^d, \beta^s$.

The *reduced form* also gives us equations we can estimate by least squares! Each equation has one endogenous variable. So we can estimate the 2 equations (expressing p and q in terms of y and w) separately using least squares. These estimators have nice properties.

Write the reduced form in a notation suitable for more than 2 endog. and 2 exog. variables:

$$p_t = \pi_{11} y_t + \pi_{12} w_t + v_{t1}$$
$$q_t = \pi_{21} y_t + \pi_{22} w_t + v_{t2}.$$

(The double subscripts suggest that the use of matrices would be useful!)

From the algebra generating the reduced form from the structural form we know that

$$\pi_{11} = \tfrac{-\beta^d}{\alpha^d - \alpha^s}; \quad \pi_{12} = \tfrac{\beta^s}{\alpha^d - \alpha^s};$$
$$\pi_{21} = \tfrac{-\alpha^s \beta^d}{\alpha^d - \alpha^s}; \quad \pi_{22} = \tfrac{\alpha^d \beta^s}{\alpha^d - \alpha^s}.$$

We can obtain estimates of our structural coefficients by first getting estimates of the $\pi's$ (call them $\widehat{\pi}'s$) and then putting the estimates into these relations.

Thus we obtain the estimates

$$\widehat{\alpha^s} = \frac{\widehat{\pi}_{21}}{\widehat{\pi}_{11}}; \qquad \widehat{\alpha^d} = \frac{\widehat{\pi}_{22}}{\widehat{\pi}_{12}};$$

$$\widehat{\beta^s} = \frac{\widehat{\pi}_{22}\left(\frac{\widehat{\pi}_{22}}{\widehat{\pi}_{12}} - \frac{\widehat{\pi}_{21}}{\widehat{\pi}_{11}}\right)}{\frac{\widehat{\pi}_{22}}{\widehat{\pi}_{12}}}; \quad \widehat{\beta^d} = \frac{\widehat{\pi}_{21}\left(\frac{\widehat{\pi}_{22}}{\widehat{\pi}_{12}} - \frac{\widehat{\pi}_{21}}{\widehat{\pi}_{11}}\right)}{\frac{\widehat{\pi}_{21}}{\widehat{\pi}_{11}}}.$$

This method of estimating the structural parameters by first estimating the reduced form parameters is called indirect least squares.