ECONOMETRICS II, FALL 2021

Systems of Equations.

Consider two (in general K) equations of form

$$y_1 = \Sigma_{i=2}^{p} \alpha_i * y_i + \Sigma_{i=1}^{l} \kappa_i * x_i + u_1,$$

and

$$y_2 = \Sigma_{i=1, i \neq 2}^{p} \beta_i * y_i + \Sigma_{i=1}^{l} \lambda_i * x_i + u_2 \,,$$

where the u terms are mean zero error terms independent of "everything else," except the terms can be correlated across equations. We ignore the intercepts that you would usually have in applications. We consider the random variables here, not a sample. Some of the α, κ, β, and λ coefficients may be 0 (or occasionally, have known values other than 0), which we refer to as restrictions. In fact, without restrictions the coefficients of interest cannot be estimated without bias and we here analyse this issue.

In matrix-vector form, we have the "structural" system of equations

$$\Gamma Y = BX + u \,,$$

where Y is a vector of dependent variables, say GDP, investment, consumption, and the Fed Funds rate, and X is vector of exogenous variables that we can treat as fixed regressors. u is independent errors with variance matrix Σ. The Γ and B matrices need to have constraints (some values set in advance) or the model is not "identified."

With a suitable sample of data, we can estimate from the "reduced form"

$$Y = \Pi X + v \,.$$

without bias, because the x's are exogenous. For predicting, one can regress Y on X, but more often we want to know how the endogenous variables affect each other. For example, a policy intervention may increase the value of y_1, and we cannot predict the effect of that on other endogenous variables without knowing the structural form.

We say that the model is identified if Γ and B can be found (i.e., the unrestricted parameters) from the reduced form. I.e., if one can solve the equation(s)

$$\Gamma^{-1} B = \Pi \,,$$

for Γ and Π. However, hardly anyone does this (it would be called 'indirect least squares"), and the main issue here is to analyze the issue.

We will do the 2-equation case, with regressors in more detail. The logic of higher-order equations should then be obvious. If

$$y_1 = \alpha_1 * y_2 + b_1 * x_1 + b_2 x_2 + b_3 x_3 + u_1,$$

and

$$y_2 = \alpha_2 * y_1 + b_4 * x_1 + b_5 x_2 + b_6 x_3 + u_2 \,,$$

we have

$$\Gamma = \begin{pmatrix} 1 & -\alpha_1 \\ -\alpha_2 & 1 \end{pmatrix} .$$

so

$$\Gamma^{-1} = \frac{1}{1-\alpha_1\alpha_2} \begin{pmatrix} 1 & \alpha_1 \\ \alpha_2 & 1 \end{pmatrix} .$$

The B matrix is

$$B = \begin{pmatrix} b_1 & b_2 & b_3 \\ b_4 & b_5 & b_6 \end{pmatrix} ,$$

So,

$$\Gamma^{-1}B = \frac{1}{1-\alpha_1\alpha_2} \begin{pmatrix} b_1 + b_4\alpha_1 & b_2 + b_5\alpha_1 & b_3 + b_6\alpha_1 \\ b_4 + b_1\alpha_2 & b_5 + b_2\alpha_2 & b_6 + b_3\alpha_2 \end{pmatrix} ,$$

and $\Gamma^{-1}B = \Pi$ then, for known (estimated) Π gives us 6 equations in 8 unknowns, so you need to impose 2 constraints to get 6 equations in 6 unknowns. Consider for example, the contraints that $b_6 = 0$ and $b_5 = 0$. Then we have the 6 equations

1. $\frac{b_1+b_4\alpha_1}{1-\alpha_1\alpha_2} = \pi_{11}$
2. $\frac{b_2}{1-\alpha_1\alpha_2} = \pi_{12}$
3. $\frac{b_3}{1-\alpha_1\alpha_2} = \pi_{13}$
4. $\frac{b_4+b_1\alpha_2}{1-\alpha_1\alpha_2} = \pi_{21}$
5. $\frac{+b_2\alpha_2}{1-\alpha_1\alpha_2} = \pi_{22}$
6. $\frac{+b_3\alpha_2}{1-\alpha_1\alpha_2} = \pi_{23}$

You can estimate α_2 easily, by dividing equation 5. by equation 2. (or equation 6. by equation 3.). Then you can find $b_1 = \pi_{11}(1 - \alpha_1\alpha_2) + b_4\alpha_1$. Substitute into equation 4. and you get

$$\frac{b_4 + \alpha_2 * (\pi_{11}(1 - \alpha_1\alpha_2) - b_4\alpha_1)}{1 - \alpha_1\alpha_2} = \pi_{21}$$

or

$$\frac{b_4 * (1 - \alpha_1\alpha_2) + \alpha_2 * \pi_{11} * (1 - \alpha_1\alpha_2)}{1 - \alpha_1\alpha_2} = \pi_{21}$$

or

$$b_4 = -\alpha_2 * \pi_{11} + \pi_{21} \,,$$

where the right hand side is known. You will however, not be able to solve for the coefficients of the y_2 equation. One can write down various conditions for when the system is identified, but those are not very useful in practice. The general rule that applied econometricians keep in mind is the following

- The parameters of the equation for y_k are identified if each of the included endogenous variables are a function of an exogenous variable that does not enter in the y_k equation. These have to be different for each endogenous variable.

- For equations that are identified, the left-out variables can be used for obtaining consistent IV-estimators. This follows from standard IV-theory: we need at least one instrument for each endogenous variable and these instruments have to satisfy the exclusion restrictions, that they do not directly affect y_k

In the example, we can estimate the equation $y_2 = \alpha_2 y_1 + b_4 x_1$ using x_2 and or x_3 as instruments. Notice, that we could find two different solutions for α_2. Those correspond to each of the instruments? So which should we use? If the model is true, each IV is consistent, so in very large samples, it doesn't matter. In practical situations, you would almost always use both instruments and you get an average of the two that could be obtained by single instruments.(There are a number of subtle points, some we will cover later under the heading of "weak instruments," and some not totally resolved, on what to do in various special cases, for example if the number of potential instruments are large.)

Doing IV, using the left-out exogenous variables as instruments for the endogenous variables is know as *Two-Stage Least Squares* (2SLS), and this is used a lot.

If the equations are correlated, the most efficient (at least asymptotically) would be to use SURE estimation (an example of feasible GLS). This is called *Three-Stage Least Squares* (3SLS): 1) estimate the equations one-by-one using 2SLS, 2) use the residuals to find the variance and covariances of the residuals the usual way and put then in matrix Σ, and 3) use the SURE estimator.

Finally, you may sometimes hear about "k-class estimators"

$$\hat{\beta}_k = (X'(I - kM_Z)X)^{-1} X'(I - kM_Z)Y \,,$$

which is the OLS estimator when $k = 0$ and 2SLS when $k = 1$.

Sometimes, "Limited Information Least Squares" (LIML) is used, where k is the smallest eigenvalue, λ, of the matrix $(Y'(I - P_{Z1})Y)\,(Y'(I - P_Z)Y)^{-1}$, where Y is all endogenous variables, Z is all exogenous variables, and $Z1$ are the exogenous variables included in the equation of interest. Lately, this estimator has received more attention as theoretical econometricians have found that it often has good properties. Different textbooks give different (but equivalent) formulas for finding the LIML value of k, see for example, Bruce Hansen's textbook. It will take us a little too far afield to derive the LIML estimator in this class and the proofs that I have seen are not that intuitive.

A variation of LIML, where we find the smallest eigenvalue, λ, as described, and then use $k = \lambda - \frac{1}{N-K}$, is called the "Fuller" estimator. The last term is a correction, which makes the estimator behave better in small samples (it obviously disappear when T becomes very large). We will not prove any of this, but only check it out in a computer homework. You will of course be expected to recognize everything that has been in a homework.