ECON 546: Themes in Econometrics

Lab Exercises #9
(17 March, 2010)

In this lab. class we are going to get some practice in estimating a Seemingly Unrelated Regression Equations (SURE) model. We'll also explore an important practical situation that often arises when this model is used.

Consider our *M*-equation system, with *T* observations on all of the variables:

$$\begin{aligned} y_1 &= X_1\beta_1 + \varepsilon_1 \\ y_2 &= X_2\beta_2 + \varepsilon_2 \\ &\vdots \\ y_M &= X_M\beta_M + \varepsilon_M \end{aligned}$$

or,

$$y = X\beta + \varepsilon$$

where y and ε are $(MT \times 1)$; X is a block-diagonal $(MT \times MK)$ matrix; and β is $(MK \times 1)$; where $K = \sum_{i=1}^{M} k_i$.

Suppose that the *same k* regressors appear in every equation. In this case we know that the SURE estimator is the same as OLS. Also, in this case we can also write this system in alternative form that will be helpful to us here:

$$Y = X^* B + U \ , \tag{1}$$

where Y and U are $(T \times M)$; X^* is $(T \times k)$; and B is $(k \times M)$. A row of this matrix equation represents one observation on the whole system; and a column corresponds to all observations on one of the equations.

Suppose that, at each observation in the sample, the values of the dependent variables of the equations sum to some linear combination of the regressors. That is:

$$Y\iota = X^*\alpha \quad ; \quad \text{where } \iota = (1,1,1,1,....,1)' \text{ is } (M \times 1) \text{; and } \alpha \text{ is } (k \times 1). \tag{2}$$

For example, it may be that the dependent variables sum to unity, and the first regressor in each equation is an intercept. In this case, $Y\iota = X^*(1,0,0,0,....,0)'$. Situations of this type arise frequently in practice – for example, with consumer demand systems, where we have Engel aggregation; with market share equations; or with trade-share models. Such models are usually called “allocation models”.

Suppose that we estimate the model by GLS. Note that equations (1) and (2) imply that:

$$B\iota = \alpha \quad \text{and} \quad U\iota = 0 . \tag{3}$$

As the regressors are the same in each equation, this just amounts to OLS applied to each equation, and results imply:

$$\hat{B}\iota = (X^{*}{}'X^{*})^{-1}X^{*}{}'Y\iota = (X^{*}{}'X^{*})^{-1}X^{*}{}'X^{*}\alpha = \alpha \tag{4}$$

and

$$\hat{Y}\iota = X^{*}\hat{B}\iota = X^{*}\alpha = Y\iota$$

So, the *estimated* values of the coefficients satisfy (3), and the *predicted* values of the dependent variables *automatically* satisfy the “adding-up” restrictions. Of course, this will not happen if there are cross-equation restrictions on the coefficients (which includes the case of different regressors in different equations – the “zero restrictions” case). However, also note, from (3) that because $U\iota = 0$, the covariance matrix of the system is *singular*:

$$\Omega\iota = \frac{1}{T}E[U'U]\iota = \frac{1}{T}E[U'U\iota] = \frac{1}{T}E[0] = 0$$

So, there is a linear dependency among the *M* rows (or columns) of *Ω*, and this $(M \times M)$ matrix therefore has rank (*M*-1). It is singular! Of course, this precludes the application of GLS estimation. We can still apply OLS, equation by equation, but what if there are restrictions on the coefficients? Then, there would be some gain form using GLS, but we can’t construct this estimator.

The solution is as follows. ***We drop one equation from the model***. The system then has (*M*-1) equations, the same as the rank of *Ω*. We then estimate the model by MLE (iterative feasible GLS). *It doesn’t matter which equation we drop - we get the same results because of the invariance of MLE*! We can “recover” the coefficients of the omitted equation by using equation (4). As an aside, to get the standard errors of these recovered coefficient estimates, a quick solution is to re-estimate the model, this time dropping a *different* equation. The coefficients that originally had to be recovered, and their standard errors, will be estimated directly this time.

Now, let’s see how this works in the case of a simple application. The data that we will be using are in the Excel workbook that is titled, **S:\Social Sciences\Economics\ECON546\LAB9.XLS**. The data relate to the market shares of various web-browsers and computer operating systems. These data are monthly, for the period October 2004 to February 2007. The variable names will be familiar to you, and the source of the data is given in the workbook.

An Eviews workfile is also on the server, ready for you to use. It is called **S:\Social Sciences\Economics\ECON546\LAB9.WF1**. A 6-equation system has been created already. Each equation explains the market share of a particular web browser. For every equation, as explanatory variables include an intercept plus all of the five “named” operating systems. (That is, do **not** use the “Other_OS” variable as a regressor – you would have a singular regressor matrix as you have an intercept in each equation.)

(a) Estimate the system by OLS – see the “system” that is already in the EViews workfile.

(b) What do the 6 estimated intercept coefficients add up to? What do the 6 estimated WINXP coefficients add up to? Why is this? Why is the determinant of the residual covariance matrix equal to zero?

(c) You can also check later (to save time now) that the ***predicted*** shares, add up across the six equations to unity. (You can’t “forecast” from the system but you can create the residuals for each equation, as a matrix, and then add them to the actual dependent variables to get the forecasts.)

(d) What happens if you treat the system as a SURE model and try to estimate is by feasible one-step GLS, or iterated GLS?

(e) Now delete the OTHER_B equation from the model and re-estimate it by OLS – compare your results with your original OLS estimates.

(f) How would you get the estimated coefficients for the Other-B equation from your second estimated model?

(g) Re-estimate this second (5-equation) model by one-step GLS and by iterated GLS. Compare your results with the OLS results.

(h) Replace the OTHER_B equation, but delete the SAFARI equation, and estimate the model by one-step GLS and iterated GLS. Compare your results with those in part (g).

Now, let’s return to the full 6-equation model, but let’s simplify it to eliminate insignificant regressors:

(i) Now delete the WIN98 regressor from ***all*** of the equations; delete the WINXP variable from the NETSCAPE, OPERA and SAFARI equations; delete the WIN2000 variable from the NETSCAPE and OPERA equations; delete the MACOS variable from the OPERA equation; and estimate this new system by OLS.

(j) Re-estimate the simplified (restricted) 6-equation model by one-step GLS and by iterated GLS. What do you notice about the significance of the MACOS regressor in the FIREFOX equation as you use different estimators?

(k) Has your iterated feasible GLS estimator converged to the MLE?

(l) Test the hypothesis that the WIN2000 regressor has the same coefficient in both the SAFARI and OTHER_B equations.

(m) If this hypothesis cannot be rejected, re-estimate the model imposing this cross-equation restriction on the coefficients.