Appendix

Panel Analysis: Fixed and Random Effect Model

One of the major benefit of panel data over cross section data is that panel data set will provide the investigator abundant flexibility in modeling behavior differences across individuals. The regression model is of the form

$$y_{it} = x'_{it}\beta + z'_i\alpha + \varepsilon_{it} \qquad \ldots 1$$

$$= x'_{it}\beta + c_i + \varepsilon_i \qquad \ldots 2$$

Let there are *K* regressors in x_{it}, without a constant term. The heterogeneity, or individual effect is $z'_i\alpha$ where z_i contains a constant term and a set of individual or group specific variables, which may be observed and c_i is unobserved. The main objective of the analysis will be consistent and efficient estimation of the partial effects,

$$\beta = \partial E[y_{it} \mid x_{it}]/\partial x_{it} \qquad \ldots 3$$

Now foe estimating the Panel regression analysis one can consider either

I) Fixed Effect Model

Or

II) Random Effect Model

Fixed Effects: If z_i is unobserved, but correlated with x_{it}, then the least squares estimator of β is biased and inconsistent as a consequence of an omitted variable. In case pf fixed effect model one can consider the following model

$$y_{it} = x'_{it}\beta + \alpha_i + \varepsilon_{it} \qquad \ldots 4$$

Where $\alpha_i = z'_i\alpha$, captures all the observable effects and specifies an estimable conditional mean. For fixed effects model, α_i stands for a group-specific constant term. It may be pointed out that "fixed" denotes the correlation between c_i and x_{it} in equation 2, note that c_i is nonstochastic.

Random Effects: When unobserved individual heterogeneity, can be assumed to be uncorrelated with the included variables, then one can consider the following model

$$y_{it} = x'_{it}\beta + E[z'_i\alpha] + \{z'_i\alpha - E[z'_i\alpha]\} + \varepsilon_{it} \qquad \ldots 5$$

$$= x'_{it}\beta + \alpha + u_i + \varepsilon_{it} \qquad \ldots 6$$

That is, as a linear regression model with a compound disturbance that could be consistently, albeit inefficiently, estimated by least squares. The random effects approach specifies u_i to be a group-specific random element. The fundamental difference between random and fixed effects is whether the unobserved individual effect represents elements which are associated with the regressors or not.

Hausman's test for the Random Effects Model

For estimating the model under Panel set-up one can consider two types of specification, i) Fixed Effect Model or ii) Random Effect Model. Now to test which specification is better one have to consider the Hausman's test which says that the random effects are free of the right hand side variables. The test is on the basis of the conjecture that under the hypothesis of zero correlation between the right hand side variables and the random effects are consistent estimators but fixed effects is inefficient.

The test is on the basis of the following Wald statistic:

$W = [\boldsymbol{\beta}_{FE} - \boldsymbol{\beta}_{\mathbf{RE}}]' \boldsymbol{\Psi}^{-1}[\boldsymbol{\beta}_{FE} - \boldsymbol{\beta}_{\mathbf{RE}}]$

where $Var[\boldsymbol{\beta}_{FE} - \boldsymbol{\beta}_{\mathbf{RE}}] = Var[\boldsymbol{\beta}_{FE}] - Var[\boldsymbol{\beta}_{RE}] = \boldsymbol{\Psi}$

W is distributed as X^2 with (K-1) degrees of freedom where K is the number of parameters in the model. If W is larger than the tabulated value, then the null hypothesis is rejected i.e. of "no correlation between the right hand side variables and the 'random effects' and in this case the fixed effects model turned out to be the better one.

Seemingly Unrelated Regression (SUR)

SUR is appropriate when all the right hand side regressors X are assumed to be exogenous, and the errors are heteroscedastic and contemporaneously correlated so that the error variance matrix is given by $V = \sum \otimes I_T$.

Zellner's SUR estimator of β takes the following form:

$$b_{SUR} = (X^{/}\ (\hat{\sum} \otimes I_T)^{-1}\ X)^{-1} X^{/}\ (\hat{\sum} \otimes I_T)^{-1} y$$

Where $\hat{\sum}$ is a consistent estimate of $\sum$ with typical element s_{ij}, for all *i* and *j*.

If autoregressive terms are incorporated in the equation, then the equation as below is estimated:

$$y_{jt} = X_{jt}\beta_j + \left[\sum_{r=1}^{p_j} p_{jr}\left(y_{j(t-r)} - X_{j(t-r)}\right)\right] + \epsilon_{jt}$$

where ε_j is assumed to be serially independent, but maybe contemporaneously correlated across equations. Now, generalized least squares (GLS) specifications can be estimated which accounts for several patterns of correlation amongst the residuals.

In the present chapter contemporaneous covariance is considered.

Cross Section SUR or Contemporaneous Covariances: This class of covariance structures permits for conditional correlation between contemporaneous residuals for cross section i and j, but confines residuals in different times to be uncorrelated, specifically that:

$E(\in_{it}\in_{jt}/\ x_t^*) = \ \sigma_{ij}$, $E(\in_{is}\in_{jt}/\ x_t^*) = \ 0$ for all i, j , s and t with s≠t.

The error terms may be thought of as cross-sectionally correlated. Alternatively, this error structure is at times denoted as clustered by period since observations for a given period are correlated. Using the period specific residual vectors one may rewrite this assumption as follows:

$E(\in_t\in_t^{/}/\ x_t^*) = \ \Omega_M$

For all t, where

$$\Omega_M = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1M} \\ \vdots & \ddots & \vdots \\ \sigma_{M1} & \cdots & \sigma_{MM} \end{bmatrix}$$

This is a cross section specification because it encompasses covariances across cross section as in a SUR type framework. Cross section SUR GLS on this specification is merely the feasible systems GLS estimator where the residuals are both cross sectionally heteroskedastic and contemporaneously correlated.

Eviews employs residual from stage 1 estimates to form an estimate of Ω_M. In stage 2, they perform feasible GLS.

White Cross-Section or Cross Section Heteroscedasticity: The White Cross-Section method is based on the assumption of contemporaneously (Cross-Sectionally) correlated (Period Clustered) errors. The method considers pool regressions as a multivariate regression (with an equation for each cross section) and calculates robust standard errors for the equations system. The coefficient covariance estimator is as follows:

$$\left(\frac{N^*}{N^* - K^*}\right)\left(\sum_t X_t^{/} X_t\right)^{-1}\left(\sum_t X_t^{/} \hat{\epsilon}_t \hat{\epsilon}_t^{/} X_t\right)\left(\sum_t X_t^{/} X_t\right)^{-1}$$

Where the leading term is a degrees of freedom adjustment contingent on the total observations in the stacked data, N^* is the total stacked observations and K^* is the total estimated parameters.

Wald Test

The Wald test statistic is $\lambda_w = \frac{(RRSS - URSS)}{URSS/(N-K)}$ which follows χ^2 distribution with k degree of freedom.

Where, k= Number of restrictions in the model under H_0, RSS_R= Residual sum of squares of the model under H_0, RSS_{UR}= Residual sum of squares of the model under H_1, N=Number of observations, K=Number of parameters in the model under H_1.