Topic 4: Random Effects and the Hausman Test

UNIVERSITY OF STIRLING

The Random Effects Model

Original equation

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \beta_2 x_{2it} + \ldots + \beta_k x_{kit} + \varepsilon_{it}$$

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \beta_2 x_{2it} + \ldots + \beta_k x_{kit} + \lambda_i + u_{it}$$

Remember $\varepsilon_{it} = \lambda_i + u_{it}$ λ_i now part of error term

This approach more appropriate if observations are representative of a sample rather than the whole population. Generally attractive argument.

UNIVERSITY OF STIRLING

The Variance Structure in Random Effects

In random effects, we assume the λ_i are part of the composite error term ε_{it}. To construct an efficient estimator we have to evaluate the structure of the error and then apply an appropriate generalised least squares estimator to find an efficient estimator. The assumptions must hold if the estimator is to be efficient. These are:

$$E(u_{it}) = E(\lambda_i) = 0; \qquad E(u_{it}^2) = \sigma_u^2;$$

$$E(\lambda_i^2) = \sigma_\lambda^2; \qquad E(u_{it}\lambda_i) = 0 \quad for\ all\ i, t$$

$$E(\varepsilon_{it}^2) = \sigma_u^2 + \sigma_\lambda^2 \quad t = s; \quad E(\varepsilon_{it}\varepsilon_{is}) = \sigma_\lambda^2, \quad t \neq s;$$

$$and$$

$$E(x_{kit}\lambda_i) = 0\ for\ all\ k, t, i$$

This is a crucial assumption for the RE model.
It is necessary for the consistency of the RE model, but not for FE. It can be tested with the Hausman test.

UNIVERSITY OF STIRLING

The Variance Structure in Random Effects

Derive the T by T matrix that describes the variance structure of the ε_{it} for individual *i*. Because the randomly drawn λ_i is present each period, there is a correlation between each pair of periods for this individual.

$$\varepsilon_i^{'} = (\varepsilon_{i1}, \varepsilon_{i2}, \ldots \varepsilon_{iT}); \text{ then } E(\varepsilon_i \varepsilon_i^{'}) =$$

$$\begin{bmatrix} \sigma_u^2 + \sigma_\lambda^2 & \sigma_\lambda^2 & \sigma_\lambda^2 & \sigma_\lambda^2 \\ \sigma_\lambda^2 & \sigma_u^2 + \sigma_\lambda^2 & & \sigma_\lambda^2 \\ \sigma_i^2 & & .. & .. \\ \sigma_\lambda^2 & \sigma_\lambda^2 & .. & \sigma_u^2 + \sigma_\lambda^2 \end{bmatrix} = \sigma_u^2 I + \sigma_\lambda^2 ee' = \Omega$$

$$\text{where } \quad e' = \left(1\;1\;1.....1\right) \quad \text{is a unit vector of size T}$$

Random Effects (GLS Estimation)

The Random Effects estimator has the standard generalised least squares form summed over all individuals in the dataset i.e.

$$\hat{\beta}_{RE} = \left[\sum_{i=1}^{N} (X_i' \Omega^{-1} X_i) \right]^{-1} \sum_{i=1}^{N} X_i' \Omega^{-1} y_i$$

Where, given Ω from the previous slide, it can be shown that:

$$\Omega^{-1/2} = \frac{1}{\sigma_u} \left(I_T - \frac{\theta}{T} ee' \right) \text{ where } \theta = 1 - \frac{\sigma_u}{\sqrt{T\sigma_\lambda^2 + \sigma_u^2}}$$

UNIVERSITY OF STIRLING

Fixed Effects (GLS Estimation)

The fixed effects estimator can also be written in GLS form which brings out its relationship to the RE estimator.
It is given by:

$$\hat{\beta}_{FE} = \left[\sum_{i=1}^{T} (X_i^{'} M X_i) \right]^{-1} \sum_{i=1}^{T} X_i^{'} M y_i \text{ where } M = I_T - \frac{1}{T} ee'$$

Premultiplying a data matrix, X, by M has the effect of constructing a new matrix, X* say, comprised of deviations from individual means. (This is a more elegant way mathematically to carry out the operation we described previously)
The FE estimator uses *M* as its weighting matrix rather than Ω.

UNIVERSITY OF STIRLING

Relationship between Random and Fixed Effects

The random effects estimator is a weighted combination of the “within” and “between” estimators. The “between” estimator is formed from:

$$\hat{\beta}_{RE} = \Psi \hat{\beta}_{Between} + (I_K - \Psi)\hat{\beta}_{Within}$$

Ψ depends on θ in such a way that if $\theta \to 1$ then the RE and FE estimators coincide. This occurs when the variability of the individual effects is large relative to the random errors.

$\theta \to 0$ corresponds to OLS (because the individual effects are small relative to the random error).

UNIVERSITY OF STIRLING

Random or Fixed Effects?

For random effects:

•Random effects are efficient

•Assumption one set of unobservables fixed and the other random?

•Sample information more common than from the entire population?

•Can deal with regressors that are fixed across individuals

Against random effects:

Likely to be correlation between the unobserved effects and the explanatory variables. These assumed to be zero in random effects model, but in many cases would be them to be non-zero. This implies **inconsistency** due to omitted variables in the RE model. Fixed effects is inefficient, but consistent.

The Hausman Test

If there is no correlation between regressors and effects, then FE and RE are both consistent, but FE is inefficient.
If there is correlation, FE is consistent and RE is inconsistent.

Under the null hypothesis of no correlation, there should be no differences between the estimators. To carry out the Hausman test

$$\text{Calculate} \quad \hat{\beta}_{\mathrm{RE}} - \hat{\beta}_{FE} \quad \text{and its covariance}$$

UNIVERSITY OF STIRLING

The Hausman Test

Is a test for the independence of the λ_i and the x_{kit}.

The covariance of an efficient estimator with its difference from an inefficient estimator should be zero. Under the null hypothesis we test:

$$\mathrm{W} = (\beta_{\mathrm{RE}} - \beta_{FE})'\hat{\Sigma}^{-1}(\beta_{\mathrm{RE}} - \beta_{FE}) \sim \chi^2(k)$$

If *W* is significant, we should not use the random effects estimator.

UNIVERSITY OF STIRLING