1 Introduction

Consider the following two sets of APT regressions corresponding to two regions

$$\begin{aligned} r_{j1t} &= \alpha_{j1} + \beta_{j1} x_{1t} + \gamma_{j1} w_{1t} + \varepsilon_{j1t} \\ r_{j2t} &= \alpha_{j2} + \beta_{j2} x_{2t} + \varepsilon_{j2t}, \end{aligned}$$

for $j = 1, 2, ..., N_i$ firms in region i, and $t = 1, 2, ..., T_i$. It is assumed that there are N_i firms in region i, and the sample size, T_i, could also differ across the regions. Notice that the regressor w is missing from region 2. The set up can be generalized to more regions. The log-likelihood function for this model (assuming the idiosyncratic terms, ε_{jit}, are independently distributed across, i, j and t) can be written as

$$\begin{aligned} l(\boldsymbol{\theta}) &= \frac{-T_1}{2} \sum_{j=1}^{N_1} \ln(2\pi\sigma_{j1}^2) - \frac{1}{2} \sum_{t=1}^{T_1} \sum_{j=1}^{N_1} \frac{\left(r_{j1t} - \alpha_{j1} - \beta_{j1} x_{1t} - \gamma_{j1} w_{1t}\right)^2}{\sigma_{j1}^2} \\ & \frac{-T_2}{2} \sum_{j=1}^{N_2} \ln(2\pi\sigma_{j2}^2) - \frac{1}{2} \sum_{t=1}^{T_2} \sum_{j=1}^{N_2} \frac{\left(r_{j2t} - \alpha_{j2} - \beta_{j2} x_{2t}\right)^2}{\sigma_{j2}^2} \end{aligned}$$

All the various estimators we are considering can be computed by maximization of $l(\boldsymbol{\theta})$, with respect to $\boldsymbol{\theta}$, which denotes the unknown parameters under the particular pooling or no pooling assumptions.

The MG estimators that we are using does not impose any homogeneity restrictions.

Pooled OLS imposes the highest degree of homogeneity, namely

$$\begin{aligned} \alpha_{j1} &= \alpha_{j2} = \alpha \\ \beta_{j1} &= \beta_{j2} = \beta \\ \gamma_{j1} &= \gamma \end{aligned}$$

and

$$\sigma_{j1}^2 = \sigma_{j2}^2 = \sigma^2.$$

The fixed effects estimators impose

$$\begin{aligned} \beta_{j1} &= \beta_{j2} = \beta \\ \gamma_{j1} &= \gamma \end{aligned}$$

and

$$\sigma_{j1}^2 = \sigma_{j2}^2 = \sigma^2.$$

A (cross-sectionally) heteroscedastic version of the FE estimator can also be computed by imposing

$$\begin{aligned} \beta_{j1} &= \beta_{j2} = \beta \\ \gamma_{j1} &= \gamma, \end{aligned}$$

without requiring that $\sigma_{j1}^2 = \sigma_{j2}^2 = \sigma^2$. Clearly, other intermediate cases can also be considered.

2 Pooled and Fixed Effects Estimators

To deal with the case where different regressors are included in the APT regressions of different regions one could simply set the values of the missing regressors equal to zero under pooling (but not when doing individual return regressions).

In general we could now write (with the missing regressors set equal to zero)

$$l(\boldsymbol{\theta}) = \sum_{i=1}^{R}\sum_{j=1}^{N_i} \frac{-T_i}{2} \ln(2\pi\sigma_{ji}^2) - \frac{1}{2}\sum_{i=1}^{R}\sum_{j=1}^{N_i}\sum_{t=1}^{T_i} \frac{\left(r_{jit} - \alpha_{ji} - \boldsymbol{\beta}_{ji}'\mathbf{x}_{it}\right)^2}{\sigma_{ji}^2},$$

where R is the number of regions, and $\mathbf{x}_{it}$ is the vector of the regressors (with dimension $k \times 1$) with the missing ones set to zero.

Pooled estimator

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\boldsymbol{\beta}} \end{pmatrix}_{pooled} = \left(\sum_{i=1}^{R}\sum_{j=1}^{N_i}\sum_{t=1}^{T_i} \mathbf{z}_{jit}\mathbf{z}_{jit}'\right)^{-1} \left(\sum_{i=1}^{R}\sum_{j=1}^{N_i}\sum_{t=1}^{T_i} \mathbf{z}_{jit} r_{jit}\right),$$

where $\mathbf{z}_{jit} = (1, \mathbf{x}_{it}')'$.

$$\widehat{Var}\begin{pmatrix} \hat{\alpha} \\ \hat{\boldsymbol{\beta}} \end{pmatrix}_{pooled} = \hat{\sigma}_{pooled}^2 \left(\sum_{i=1}^{R}\sum_{j=1}^{N_i}\sum_{t=1}^{T_i} \mathbf{z}_{jit}\mathbf{z}_{jit}'\right)^{-1},$$

where

$$\hat{\sigma}_{pooled}^2 = \frac{\sum_{i=1}^{R}\sum_{j=1}^{N_i}\sum_{t=1}^{T_i}\left(r_{jit} - \hat{\alpha}_{pooled} - \hat{\boldsymbol{\beta}}_{pooled}'\mathbf{x}_{it}\right)^2}{\sum_{i=1}^{R} N_i T_i - k - 1}.$$

Fixed effects estimator

$$\hat{\boldsymbol{\beta}}_{FE} = \left(\sum_{i=1}^{R}\sum_{j=1}^{N_i}\sum_{t=1}^{T_i} (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)'\right)^{-1} \left(\sum_{i=1}^{R}\sum_{j=1}^{N_i}\sum_{t=1}^{T_i} (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)(r_{jit} - \bar{r}_{ji})\right), \tag{1}$$

$$\bar{r}_{ji} = \frac{1}{T_i}\sum_{t=1}^{T_i} r_{jit},$$
$$\bar{\mathbf{x}}_i = \frac{1}{T_i}\sum_{t=1}^{T_i} \mathbf{x}_{it}.$$

Also

$$\hat{\alpha}_{ji} = \bar{r}_{ji} - \hat{\boldsymbol{\beta}}_{FE}'\bar{\mathbf{x}}_i,$$

and

$$\hat{\sigma}_{FE}^2 = \frac{\sum_{i=1}^{R}\sum_{j=1}^{N_i}\sum_{t=1}^{T_i}\left(r_{jit} - \hat{\alpha}_{ji} - \hat{\boldsymbol{\beta}}_{FE}'\mathbf{x}_{it}\right)^2}{\sum_{i=1}^{R} N_i(T_i - 1) - k}.$$

Finally,

$$\widehat{Var}\left(\hat{\boldsymbol{\beta}}_{FE}\right)=\hat{\sigma}_{FE}^2\left(\sum_{i=1}^{R}\sum_{j=1}^{N_i}\sum_{t=1}^{T_i}\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)'\right)^{-1}.$$

3 FE Estimators with cross-sectionaly heteroscedastic error variances

In this case the log-likelihood function will have the following form:

$$l(\boldsymbol{\theta})=\sum_{i=1}^{R}\sum_{j=1}^{N_i}\frac{-T_i}{2}\ln(2\pi\sigma_{ji}^2)-\frac{1}{2}\sum_{i=1}^{R}\sum_{j=1}^{N_i}\sum_{t=1}^{T_i}\frac{\left(r_{jit}-\alpha_{ji}-\boldsymbol{\beta}'\mathbf{x}_{it}\right)^2}{\sigma_{ji}^2}.$$

The maximum likelihood estimators for the case are given by

$$\hat{\alpha}_{ji}=\bar{r}_{ji}-\hat{\boldsymbol{\beta}}_{FEhetro}'\bar{\mathbf{x}}_i,$$

where

$$\hat{\boldsymbol{\beta}}_{FEhetro}=\left(\sum_{i=1}^{R}\sum_{j=1}^{N_i}\sum_{t=1}^{T_i}\frac{\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)'}{\hat{\sigma}_{ji}^2}\right)^{-1}\left(\sum_{i=1}^{R}\sum_{j=1}^{N_i}\sum_{t=1}^{T_i}\frac{\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)\left(r_{jit}-\bar{r}_{ji}\right)}{\hat{\sigma}_{ji}^2}\right), \tag{2}$$

and

$$\hat{\sigma}_{ji}^2=\frac{\sum_{t=1}^{T_i}\left[r_{jit}-\bar{r}_{ji}-\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)'\hat{\boldsymbol{\beta}}_{FEhetro}\right]^2}{T}. \tag{3}$$

The above system of equations are non-linear in $\hat{\boldsymbol{\beta}}_{FEhetro}$ and $\hat{\sigma}_{ji}^2$, but can be solved numerically by the method of succesive iterations. For $\boldsymbol{\beta}$ start with $\hat{\boldsymbol{\beta}}_{FE}$ given by (1), and obtain a preliminary set of estimates of $\hat{\sigma}_{ji}^2$ by using $\hat{\boldsymbol{\beta}}_{FE}$ in (2). Use the resulting estimate of $\hat{\boldsymbol{\beta}}_{FEhetro}$ in (3) to compute a new estimate of $\boldsymbol{\beta}$. Repeat this procecc until the successive estimates of $\boldsymbol{\beta}$ differ by small amounts, namely if

$$\sum_{\ell=1}^{k}\left|\beta_\ell^{(s)}-\beta_\ell^{(s-1)}\right|<\varepsilon,$$

where ε is a small positive number, and $\beta_\ell^{(s)}$ refers to the estimator of the ℓ^{th} element of $\boldsymbol{\beta}$ in the s^{th} iteration.

The variance matrix of the slope coefficients in this case are given by

$$\widehat{Var}\left(\hat{\boldsymbol{\beta}}_{FEhetro}\right)=\left(\sum_{i=1}^{R}\sum_{j=1}^{N_i}\sum_{t=1}^{T_i}\frac{\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)'}{\hat{\sigma}_{ji}^2}\right)^{-1}.$$

4 A Final Remark

In the case where the available time series data differs across firms in the same region, one needs to replace T_i in the above expressions by T_{ji}, the sample size for firm j in region i.

Hashem Pesaran
January 2004