Advanced Econometrics: Panel Data Econometrics Unobserved heterogeneity

PhD. Econ. Freddy Rojas Cama
frojas@usat.edu.pe
@freddyrojascama

13 de enero de 2022

Universidad Católica Santo Toribio de Mogrovejo

Contenido

1 Motivation

2 Error component model

3 The model

- Within-Group estimator
- Between-Group estimator
- Random Efects model
- Estimating in STATA

4 References

Motivation

Unobserved-Heterogeneity

- Controlling unobserved time invariant heterogeneity in cross-sectional models.
- Dynamics of cross-sectional populations through time; change in permanent component of each household or unit across time.

Unobserved-Heterogeneity

Identification problem. Considering the following

$$\begin{aligned} \text{true model} \quad &: \quad Y = X_1\beta_1^* + X_2\beta_2 + u \\ \text{estimated model} \quad &: \quad Y = X_1\beta_1 + u \end{aligned}$$

having the classical properties e.g u is $i.i.d.$ we estimate the model

$$\begin{aligned} \widehat{\beta_1} &= \left(X_1'X_1\right)^{-1} X_1'Y \\ \widehat{\beta_1} &= \left(X_1'X_1\right)^{-1} X_1' \left(X_1\beta_1^* + X_2\beta_2 + u\right) \\ \widehat{\beta_1} &= \beta_1^* + \left(X_1'X_1\right)^{-1} X_1'X_2\beta_2 + \left(X_1'X_1\right)^{-1} X_1'u \end{aligned} \tag{1}$$

the estimator is biased.

Unobserved-Heterogeneity

$$
\begin{aligned}
E\left(\widehat{\beta}_1\right) =& \beta_1^* + \left(X_1' X_1\right)^{-1} X_1' X_2 \beta_2 \\
E\left(\widehat{\beta}_1\right) =& \beta_1^* + \frac{cov\left(X_1, X_2\right)}{var\left(X_1\right)} \beta_2
\end{aligned} \tag{2}
$$

We have effects in the variance as well

$$
var\left(\widehat{\beta}_1\right) - var\left(\widehat{\beta}_1^*\right) = \sigma^{-2}\left(\left(X_1' X_2\right)\left(X_1' X_2\right)^{-1}\left(X_1' X_2\right)\right) \tag{3}
$$

above expression is positive

Unobserved-Heterogeneity

why is this important?

Unobserved-Heterogeneity

Identification problem. What if we have redundant variables?.Considering the following

$$\begin{aligned} \text{estimated model} \quad &: \quad Y = X_1\beta_1 + X_2\beta_2 + u \\ \text{true model} \quad &: \quad Y = X_1\beta_1^* + u \end{aligned}$$

having the classical properties e.g u is *i.i.d.* In this particular case $X_1 = \iota_T$ (the true model just has a constant). We estimate the model with the classical procedure of deviations from time means:

$$\begin{aligned} \widehat{\beta}_2 &= \left(\widetilde{X}_2'\widetilde{X}_2\right)^{-1}\widetilde{X}_2'\widetilde{Y} \\ \widehat{\beta}_2 &= \left(\widetilde{X}_2'\widetilde{X}_2\right)^{-1}\widetilde{X}_2'\left(\widetilde{X}_1\beta_1^* + \widetilde{u}\right) \\ \widehat{\beta}_2 &= 0 \end{aligned} \tag{4}$$

getting the constant $\beta_1^* = \overline{Y} - \overline{X}_2\widehat{\beta}_2 = \overline{Y}$

Unobserved-Heterogeneity motivation

If we have redundant variables in our regression model

- OLS estimator is unbiased and consistent

Observables and Unobservables

In some cases we would expect correlation between unobservales and observables. For instance in the demand- and supply-equation model

$$\begin{aligned} Q^d &= p'\beta + (f'\gamma + \epsilon) \\ p &= Q^{s\prime}\phi + \psi \end{aligned} \tag{5}$$

being $\beta < 0$ and $\phi > 0$. f is an unobservable variable (shocks related to preferences, culture) In equilibrium:

$$p = \frac{f'\gamma\phi}{(1-\beta)} + \frac{\epsilon\phi + \psi}{(1-\beta)}$$

So, $E(p, f) \neq 0$. In a regression model, we wonder if we have a demand or supply equation.

Observables and Unobservables

- Classical approach to pinpoint demand equation is to regress quantity on price, because of correlation between price and unobservable the classical strategy is useless.
- Strategies for identifying demand equation is available in econometric literature (see Gujarati). For example we suppose that supply equation is given by:

$$p = \rho p_{-1} + Q^{s\prime}\lambda + \xi$$

so a valid instrument should be $Z = p_{-1}$, thus the staregy of identification would be

$$Q^d = Z'\beta + \zeta$$

because of $E(Z, f) = 0$.

Observables and Unobservables

- Earnings equations:

$$wage_{it} = \alpha + \alpha_s \cdot schooling_{it} + \alpha_u Union_{it} + \varepsilon_{it}$$

thus estimator of $\alpha' s$ is $E\left(X_{it}{}'\varepsilon_{it}\right) = 0$. But, ability is correlated with schooling, so makes sense to have a model including a individual specific variable or endownment

$$wage_{it} = \alpha + \alpha_s \cdot schooling_{it} + \alpha_u Union_{it} + \mu_i + \varepsilon_{it}$$

- Cost or production function (firm data). In a general sense μ_i measures a *"managerial ability"* or *"entrepreneurial"* ability or firm efficiency.
- Demand equation or consumption functions. Differences across states, countries or individuals.

Error component model

Error component model

- Fixed effects or within-grup estimator
- Random effects estimator
  - Feasible GLS estimation
  - Maximum Likelihood Estimation.

Error component model

In the model

$$y_{it} = \alpha + \beta' X_{it} + u_{it} \tag{6}$$
$$u_{it} = \mu_i + v_{it}$$

- In Fixed effects or within-grup estimator the moments to match are $E\left(X_{it}' v_{it}\right) = 0$, there is no restriction on $E\left(X_{it}' \mu_i\right)$.
- In Random effects estimator, the moments to match are $E\left(X_{it}' v_{it}\right) = 0$ and $E\left(X_{it}' \mu_i\right) = 0$.

The model

The model

- The following model in the first observation:

$$y_{i1} = \beta x_i + \mu_i + v_{i1} \tag{7}$$

and

$$E\left(v_{i1} | x_{i1}, \mu_i\right) = 0$$

The model

- The following model in the first observation:

$$y_{i1} = \beta x_i + \mu_i + v_{i1} \tag{7}$$

and

$$E\left(v_{i1} | x_{i1}, \mu_i\right) = 0$$

- Why do we require μ_i to be observed?

The model

- The following model in the first observation:

$$y_{i1} = \beta x_i + \mu_i + v_{i1} \tag{7}$$

and

$$E\left(v_{i1} | x_{i1}, \mu_i\right) = 0$$

- Why do we require μ_i to be observed?
  - Fixed effects (invariant in time)

The model

- The following model in the first observation:

$$y_{i1} = \beta x_i + \mu_i + v_{i1} \tag{7}$$

and

$$E\left(v_{i1} | x_{i1}, \mu_i\right) = 0$$

- Why do we require μ_i to be observed?
  - Fixed effects (invariant in time)
  - Idiosincratic effects.

The model

- The following model in the first observation:

$$y_{i1} = \beta x_i + \mu_i + v_{i1} \tag{7}$$

and

$$E\left(v_{i1} | x_{i1}, \mu_i\right) = 0$$

- Why do we require μ_i to be observed?
  - Fixed effects (invariant in time)
  - Idiosincratic effects.
- In the model (7) β is identified even if we do not observe μ_i but you require that they (x_i and μ_i) are not correlated:

$$cov\left(x_{i1}, \mu_i\right) = 0 \qquad \beta = \frac{cov\left(x_{i1}, y_{i1}\right)}{var\left(x_{i1}\right)} \tag{8}$$

The model

- The following model in the first observation:

$$y_{i1} = \beta x_i + \mu_i + v_{i1} \tag{7}$$

and

$$E\left(v_{i1} | x_{i1}, \mu_i\right) = 0$$

- Why do we require μ_i to be observed?
  - Fixed effects (invariant in time)
  - Idiosincratic effects.
- In the model (7) β is identified even if we do not observe μ_i but you require that they (x_i and μ_i) are not correlated:

$$cov\left(x_{i1}, \mu_i\right) = 0 \qquad \beta = \frac{cov\left(x_{i1}, y_{i1}\right)}{var\left(x_{i1}\right)} \tag{8}$$

Instrumental variables for identification

Or the ability to have an instrument

$$cov\left(Z_i, \mu_i\right) = 0 \tag{9}$$

$$\beta = \frac{cov\left(Z_i, y_{i1}\right)}{cov\left(Z_i, x_{i1}\right)}$$

If μ_i is correlated with x_i?. So it is true that:

$$\begin{aligned} cov\left(Z_i, x_i\right) &\neq 0 \\ cov\left(Z_i, \mu_i\right) &= 0 \\ cov\left(Z_i, v_i\right) &= 0 \end{aligned} \tag{10}$$

What we need is the second one for regression strategy to work. we should replace x_i with some variable we know is not correlated with μ_i but correlated with x_i

We need something else!

- Suppose that neither of these two above options are available, so a valid strategy is:

$$y_{i2} = \beta x_{i2} + \mu_i + v_{i2}$$

satisfying

$$E\left(v_{it} | x_{i2}, x_{i1}, \mu_i\right) = 0$$

taking first differences (difference-in-difference estimator)

$$\beta = \frac{cov\left(\Delta x_{i2}, \Delta y_{i2}\right)}{var\left(\Delta x_{i2}\right)} \tag{11}$$

- **Conclusion**: Panel data help us for identifying structural parameters.

We need something else!

- Suppose that neither of these two above options are available, so a valid strategy is:

$$y_{i2} = \beta x_{i2} + \mu_i + v_{i2}$$

satisfying

$$E\left(v_{it} | x_{i2}, x_{i1}, \mu_i\right) = 0$$

taking first differences (difference-in-difference estimator)

$$\beta = \frac{cov\left(\Delta x_{i2}, \Delta y_{i2}\right)}{var\left(\Delta x_{i2}\right)} \tag{11}$$

- **Conclusion**: Panel data help us for identifying structural parameters.

The model

Within-Group estimator

Within-Group (WG) estimator

A simplistic homogeneous model

$$y_{it} = \alpha + X_{it}'\beta + u_{it} \tag{12}$$
$$u_{it} = \mu_i + v_{it}$$

heterogeneity given by μ_i.
Model (12) y_{it} could be a earning equation. So, the dependent variable could be the log of wage. X gathers experience, experience^2,schooling, sex, race, union, marital status and so on., μ_i is the individual specific ability.v_{it} is the classical disturbance *iid*. Also, we should take notice that we have the same β for all individuals.

Within-Group (WG) estimator

Arellano (2003) states that the name "withing-group"was originated in the context of data with a group structure. Panel data can be regarded as a special case of this type of data in which the group is formed by the time series observations from a given individual.

Within-Group estimator and some matrix algebra

In order to get the WG estimator we need some matrix algebra. the folowingmodel (another way to present the fixed effects model):

$$\underset{NT\times 1}{y} = \alpha \underset{NT\times 1}{\iota_{NT}} + \underset{NT\times K}{X} \underset{K\times 1}{\beta} + \underset{NT\times 1}{u}$$

$$u = Z_\mu \mu + \nu$$

where $Z_\mu = I_N \otimes \iota_T$. Z_μ is a matrix of ones and zeros. Baltagi calls Z_μ the "matrix of individual dummies".(*auditorium: we did not get it!*).

Within-Group estimator and some matrix algebra

The matrix Z_μ and for $T = 3$.

$$Z_\mu = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

Within-Group estimator and some matrix algebra

some curiosities (it is worth getting them!):

$$Z_\mu Z_\mu' = I_N \otimes \mathcal{J}_T \tag{13}$$

where $\mathcal{J}_T$ is a matrix of ones of dimension T. How did you get expression (13)?.

$$Z_\mu Z_\mu' = (I_N \otimes \iota_T)\,(I_N \otimes \iota_T)'$$
$$Z_\mu Z_\mu' = I_N \otimes \iota_T \iota_T'$$
$$Z_\mu Z_\mu' = I_N \otimes \mathcal{J}_T$$

(*auditorium: Fred, why you are showing this? where are we going to?*)

Within-Group estimator and some matrix algebra

a new matrix (an idempotent one!):

$$P_\mu = Z_\mu \left(Z_\mu' Z_\mu \right)^{-1} Z_\mu'$$

P is an inetersting matrix because if you pre multiply this matrix with a row vector, you will have a vector of meaned series across time. (*auditorium: How do you get this result?*). Let's see, we know that

$$Z_\mu' Z_\mu = \left(I_N \otimes \iota_T' \right) \left(I_N \otimes \iota_T \right)$$
$$Z_\mu' Z_\mu = I_N \otimes T$$
$$Z_\mu' Z_\mu = T \cdot I_N$$

Within-Group estimator and some matrix algebra

considering expression for P_μ (projection matrix on Z_μ) and incorporating on it the above results:

$$P = T^{-1}\left(I_N \otimes \mathcal{J}_T\right)$$
$$P = I_N \otimes \overline{\mathcal{J}}_T$$

where $\overline{\mathcal{J}}_T = \mathcal{J}_T \cdot T^{-1}$. So, it is clear that

- $\underset{NT\times 1}{Py}$ which has a typical element $\overline{y}_i = T^{-1}\sum_{t=1}^{T} y_{it}$

Within-Group estimator and some matrix algebra

considering expression for P_μ (projection matrix on Z_μ) and incorporating on it the above results:

$$P = T^{-1}\left(I_N \otimes \mathcal{J}_T\right)$$
$$P = I_N \otimes \overline{\mathcal{J}}_T$$

where $\overline{\mathcal{J}}_T = \mathcal{J}_T \cdot T^{-1}$. So, it is clear that

- $\underset{NT\times 1}{Py}$ which has a typical element $\overline{y}_i = T^{-1}\sum_{t=1}^{T} y_{it}$
- $\underset{NT\times 1}{Qy}$ which has a typical element $\widetilde{y}_{it} = y_{it} - T^{-1}\sum_{t=1}^{T} y_{it}$. $Q = I - P_\mu$.

Within-Group estimator

the matrix $\underset{NT\times NT}{Q}$ is known as the deviations-from time-means or withing group estimator because it transforms the originals eries into deviations from time means

$$\widetilde{y}_i = Qy_i$$

again $\underset{NT\times 1}{Qy}$ which has a typical element $\widetilde{y}_i = y_{it} - T^{-1}\sum_{t=1}^{T} y_{it}$. $Q = I - P_\mu$.

Some properties

P+Q=I

P,Q idempotent PP=P, Q Q=Q

and PQ=0

Within-Group estimator

Considering the model:

$$\underset{NT\times 1}{y} = \alpha \underset{NT\times 1}{\iota_{NT}} + \underset{NT\times K}{X} \underset{K\times 1}{\beta} + \underset{NT\times 1}{u}$$

$$u = Z_\mu \mu + v$$

Z_μ is the "matrix of individual dummies",μ is a $N \times 1$ column vector. Setting up the transformation:

$$Qy = QX\beta + Qv$$

the estimator withing group (recall econometric review session)

$$\widehat{\beta}_{WG} = (X'QX)^{-1}(X'Qy)$$

Within-Group estimator

the estimator withing group:

$$\widehat{\beta}_{WG} = (X'QX)^{-1}(X'Qy) \tag{14}$$

the individual effects are wiped out. Vector β_{WG} gathers parameters of interest, incidental parameters are not included because they are individual effects. Expression (14) is the withing-group or fixed effects estimator This estimator is very popular and most of econometric softwares have implement it. STATA and Eviews do this procedure and they construct the dummies series (for each member of panel) but they do not show the dummies.

Within-Group estimator

we need an estimate of V

$$\widehat{V} = \frac{1}{N}\widetilde{X}'\widehat{\Xi}\,\widetilde{X}$$

where

$$\widehat{\Xi} = \begin{bmatrix} \underset{(T\times 1)(1\times T)}{\widehat{v}_1.\widehat{v}_{1.}'} & \mathbf{0} & \cdot & \mathbf{0} \\ \mathbf{0} & \underset{(T\times 1)(1\times T)}{\widehat{v}_2.\widehat{v}_{2.}'} & \cdot & \mathbf{0} \\ \cdot & \cdot & \cdot & \cdot \\ \mathbf{0} & \mathbf{0} & \cdot & \underset{(T\times 1)(1\times T)}{\widehat{v}_N.\widehat{v}_{N.}'} \end{bmatrix}_{\dim\, NT\times NT}$$

where $\widehat{v}_i = \widetilde{y}_i - \widetilde{X}_i\widehat{\beta}$

Within-Group estimator

Let's go to Excel!

Within-Group estimator

In Arellano's terms

$$\widehat{\beta}_{WG} = \left[\sum_{i=1}^{N} \sum_{t=1}^{T} \left(x_{it} - \overline{x}_i \right) \left(x_{it} - \overline{x}_i \right)' \right]^{-1} \left[\sum_{i=1}^{N} \sum_{t=1}^{T} \left(x_{it} - \overline{x}_i \right) \left(y_{it} - \overline{y}_i \right)' \right]$$

Within-Group estimator Vs. OLS

Biased and inconsistency parameters under a true model.

Within-Group estimator is not a Panacea

(Real) heterogeneity in panel data

Within-Group estimator's disadvantages

- N-1 dummies for those fixed effects. A lof of degrees of freedom lost but people like and want to see dummies.
- "*I wanted to see the dummy for year 2006, but I couldn't!*". Stata randomly drops one or more dummies if there is perfect collinearity.
- A lot of dummies vs a procedure which estimates just 1 parameter.
- the procedure wipes out individual time-invariant specifical variables which maybe we are interested (in earning equations variables as sex, race and religion and so on are important).

Within-Group estimator's advantages

- In terms of specification: OLS baised and inconsistent if within-group effects are important (omitted variables problem).
- Specifical variables are already interesting in our study (in earning equations variables as sex, race and religion and so on are important). But that is not a problem if we deal with LSDV (least squares dummy variables).

Within-Group estimator's Interpretation

- μ_i are fixed (Baltagi "*or random but hopelessly correlated with* X_{it}'')
- In other words if the true model does have fixed effects so OLS is biased and inconsistent.

Within-Group STATA syntax

in Stata syntax:

```
xtreg depvar [indepvars] [if] [in] [, fe FE_options]
```

an example for variables `Y1 X1 X2` and group identifier `varname`

```
xtreg Y1 X1 X2, fe
```

before you perform `xtreg` you should set up the panel structure, use `xtsset`.

See `help xtreg`.

Within-Group STATA syntax

```
. xtreg I F C, fe

Fixed-effects (within) regression               Number of obs      =       200
Group variable: fn                              Number of groups   =        10

R-sq:  within  = 0.7668                         Obs per group: min =        20
       between = 0.8194                                        avg =      20.0
       overall = 0.8060                                        max =        20

                                                F(2,188)           =    309.01
corr(u_i, Xb)  = -0.1517                        Prob > F           =    0.0000

------------------------------------------------------------------------------
           I |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           F |   .1101238   .0118567     9.29   0.000     .0867345    .1335131
           C |   .3100653   .0173545    17.87   0.000     .2758308    .3442999
       _cons |  -58.74393   12.45369    -4.72   0.000    -83.31086     -34.177
-------------+----------------------------------------------------------------
     sigma_u |  85.732501
     sigma_e |  52.767964
         rho |  .72525012   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(9, 188) =    49.18              Prob > F = 0.0000
```

Within-Group STATA output

- At the top of the output we can see an "*overall*" F test but this includes only of those terms "we are interested" (excluding incidental parameters).
- There is a F-test under the null: all $\mu_i = 0$.
- $rho = \frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_\nu^2}$
- The output shows the unrestricted correlation between μ_i and $X\prime s$.
- Under this command `xtreg` the dummies (fixed effects) are not shown.

FE options

FE_options	description
Model	
fe	use fixed-effects estimator
SE/Robust	
vce(*vcetype***)**	*vcetype* may be **conventional**, **robust**, **cluster** *clustvar*, **bootstrap**, or **jackknife**
Reporting	
level(#)	set confidence level; default is **level(95)**
display_options	control spacing and display of omitted variables and base and empty cells
+ **coeflegend**	display coefficients' legend instead of coefficient table

+ **coeflegend** does not appear in the dialog box.

Within-Group STATA

Other way to estimate fixed effects

areg depvar [*indepvars*] [if] [in] ,a(*varname*) [options]

an example for variables `Y1 X1 X2` and group identifier `varname`

```
areg Y1 X1 X2, a(id)
```

See `h areg`. This *absorb* command is useful because you also have an overall F test including the dummies variables

Options

areg

options	description
Model	
* **absorb(***varname***)**	categorical variable to be absorbed
SE/Robust	
vce(*vcetype***)**	*vcetype* may be **ols**, **robust**, **cluster** *clustvar*, **bootstrap**, or **jackknife**
Reporting	
level(#)	set confidence level; default is **level(95)**
display_options	control spacing and display of omitted variables and base and empty cells
+ **coeflegend**	display coefficients' legend instead of coefficient table

Output

areg

```
. areg I F C, a(fn) cluster(fn)

Linear regression, absorbing indicators                 Number of obs =     200
                                                        F(  2,     9) =   27.02
                                                        Prob > F      =  0.0002
                                                        R-squared     =  0.9441
                                                        Adj R-squared =  0.9408
                                                        Root MSE      =  52.768

                                   (Std. Err. adjusted for 10 clusters in fn)
------------------------------------------------------------------------------
             |               Robust
           I |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           F |   .1101238   .0155539     7.08   0.000     .0749383    .1453092
           C |   .3100653   .0539997     5.74   0.000     .1879096    .4322211
       _cons |  -58.74393   28.25584    -2.08   0.067    -122.6631    5.175229
-------------+----------------------------------------------------------------
          fn |   absorbed                                     (10 categories)
------------------------------------------------------------------------------
```

Least squares dummy variables (LSDV)

Incorporating the dummies in the X matrix and perform:

$$\widehat{\beta}_{LSDV} = \left(X^{+\prime} X^{+}\right)^{-1} \left(X^{+\prime} y^{+}\right) \tag{15}$$

where $X^{+} = [X \quad I_N \otimes \iota_T]$. In Stata syntax:

```
xi [, prefix(string)noomit] i. varname
```

an example for variables `Y1 X1 X2` and group identifier `varname`

STATA code

```
xi: reg Y1 X1 X2, i.varname
```

Least squares dummy variables (LSDV)

Other way to get the dummies into the regresion.
In Stata syntax:

- `[quietly] tab varname, [gen(stubname)]`
- `reg depvar [indepvar] stubname*[if] [in] [,options]`

an example for variables `Y1 X1 X2` and group identifier `varname`

STATA code

```
quietly tab id, gen(dum)
reg Y1 X1 X2 dum*
```

Let's go to STATA

Let's go to STATA. Please, Download data from `"www.wiley.co.uk/baltagi"`. Definition of variables is as follows

- I_{it} denotes real gross investment for firm i in year t.
- F_{it} is the real value of the firm (shares outstanding).
- C_{it} is the real value of the capital stock.

10 large US manufacturing firms over 20 years 1935-54.

Within-Group estimator

therefore, the robust asymptotic variance and covariance matrix of β is estimated by:

$$var\left(\beta\right) = \frac{1}{N} \cdot \left(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}\right)^{-1} N \cdot \left(\frac{1}{N}\widetilde{\mathbf{X}}'\widehat{\Xi}\,\widetilde{\mathbf{X}}\right) \cdot \left(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}\right)^{-1} N$$

$$var\left(\beta\right) = \left(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}\right)^{-1} \cdot \left(\widetilde{\mathbf{X}}'\widehat{\Xi}\,\widetilde{\mathbf{X}}\right) \cdot \left(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}\right)^{-1} \tag{16}$$

if $\widehat{\Xi} = \sigma^2 I_{NT}$ so:

$$var\left(\beta\right) = \sigma^2 \left(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}\right)^{-1}$$

Within-Group estimator, summarizing...

- OLS biased and inconsistent when the true model is a FE model. Baltagi states this result is because of an "omitted variables problem".
- Test for individual dummies (standard F test). Useful for testing Pooling models.
- Test using LR test (imposing restrictions).
- Robust Standard Errors (Arellano;1987).

Alternative estimator for fixed effects: difference in difference

- Withing group estimator is not the only transformation that will drop out the individual effects.
- Difference-indifference estimator makes the trick as well.
- Strategy of information only estimates parameters of ïnterest".

Alternative estimator for fixed effects: difference in difference

Which in compact form can be written as

$$Dy_i = DX_i\beta + Dv_i$$

where D is the $(T-1) \times T$ matrix "first-difference operator"

$$D = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}$$

under assumptions $E\left(Dv_i|x_i\right) = 0$ and/or $E\left(Dv_i|x_i\right) = 0$ (Arellano; 2003),

Alternative estimator for fixed effects: difference in difference

In terms of panel structure

$$\underset{\dim N \times T-1}{D_N} = I_N \otimes D$$

thus pre-multiplying this matrix to data

Alternative estimator for fixed effects: difference in difference

OLS estimates of β in this system given by

$$\widehat{\beta}_D = \left((D_N X)' (D_N X)\right)^{-1} \left((D_N X)' (D_N y)\right)$$

Arellano (2003) considers

$$\widehat{\beta}_D = \left(\sum_{i=1}^{N} \left(X_i' D' D X_i\right)\right)^{-1} \sum_{i=1}^{N} (D X_i)' (D y_i)$$

this $\widehat{\beta}_D$ is unbiased and consistent for large N. What about the variance?

Alternative estimator for fixed effects: difference in difference

the variance matrix of errors

$$var\left(D_N v|X\right) = \left(\sigma_v^2 I_{NT}\right) \cdot D_N D_N'$$

which show an autocorrelation structure

$$var\left(D_N v|X\right) = \begin{bmatrix} 2\sigma_v^2 & 0 & 0 & \dots & 0 & 0 \\ -\sigma_v^2 & 2\sigma_v^2 & 0 & \dots & 0 & 0 \\ 0 & -\sigma_v^2 & 2\sigma_v^2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & 0 & \dots & -\sigma_v^2 & 2\sigma_v^2 \end{bmatrix}$$

result: difference in difference estimator is not BLUE because it is inefficient (but unbiased). That means that Gauss Markov theorem is not hold.

Alternative estimator for fixed effects: difference in difference

Solution: generalized squares. Following the standard procedure we should use GLS. This matrix makes the trick:

$$P_{D_N} = D_N' \left(D_N' D_N\right)^{-1} D_N$$

P_{D_N} is idempotent $P_{D_N} P_{D_N}' = P_{D_N}$ or $P_{D_N} P_{D_N} = P_{D_N}$. So,

$$\widehat{\beta}_{WG} = \left(\sum_{i=1}^{N} \left(X_i' P_{D_N} X_i\right)\right)^{-1} \sum_{i=1}^{N} X_i' P_{D_N} y_i$$

Alternative estimator for fixed effects: difference in difference

such a nice surprise!. The idempotent matrix P_{D_N} also takes the form:

$$P_{D_N} = I_{NT} - I_N \otimes \bar{J}_T = Q$$

So, the difference in difference estimator is unbiased and consistent.

Alternative estimator for fixed effects: difference in difference

Orthogonal deviations. Finally, we have realized that if we are using difference and difference we need to do 2 transformations: a) to get differenced data and b) remove the moving-average serial correlation (known pattern) induced by a). The overall tranformation is:

$$A = \left(D_N D_N'\right)^{-\frac{1}{2}} D_N$$

therefore $v^* = Av$ which will have $T-1$ elements by i

$$s_{it}^* = c_t \cdot \left[s_{it} - \frac{1}{T-t} \left(s_{it+1} + \ldots + s_{iT} \right) \right]$$
$$t = 1, \ldots, T-1$$

Alternative estimator for fixed effects: difference in difference

where $c_t = \left(\frac{T-t}{T=t-1}\right)^{\frac{1}{2}}$.(see Arellano and Bover; 1995). thus, we get for each first $(T-1)$ observations we substract the mean of the remaining future observations in the sample. The weighting c_t is introduced because variances of transformed erros must be equalized. This transformation is called "*forward orthogonal deviations*".

Difference in difference estimator: summarizing...

- FE is more efficient than difference in difference (FD) estimator when v_{it} are *iid*.
- FD is more efficient than FE estimator when v_{it} is random walk.
- Both estimators are affected in different ways when there are measurements errors.

Efficiency of difference-in-difference estimator

- FD is more efficient than FE estimator when v_{it} is random walk ($v_{it} = v_{it-1} + \xi_{it}$). Recalling that

$$\begin{aligned} var\left(D_N v|X\right) &= D_N v v' D_N' \\ &= \xi\xi' \\ &= \sigma_\xi^2 \cdot I_{NT} \end{aligned}$$

The model

Between-Group estimator

Between-Group estimator

- This estimator between group can be derived using the projection matrix P

$$\underset{NT\times 1}{y} = \alpha \underset{NT\times 1}{\iota_{NT}} + \underset{NT\times K}{X} \underset{K\times 1}{\beta} + \underset{NT\times 1}{u}$$

$$u = Z_\mu \mu + \nu$$

premultiplying by P

$$Py = P\iota_{NT} + PX\beta + Pu \tag{17}$$

the LS estimator under (17). (Recall econometric review session)

$$\beta_{BE} = \left(X'\left(P - \overline{\mathcal{J}}_{NT}\right)X\right)^{-1}\left(X'\left(P - \overline{\mathcal{J}}_{NT}\right)y\right) \tag{18}$$

Between-Group estimator

- The estimator between group (in easy format):

$$\overline{y}_i = \alpha + \overline{X}_i\beta + \overline{u}_i \tag{19}$$

the cross-section regression is based on N observations. The STATA syntax is

```
xtreg depvar [indepvars] [if] [in] [,be BE_options]
```

an example for variables `Y1 X1 X2` and group identifier `varname`

BE options

BE_options	description
Model	
be	use between-effects estimator
wls	use weighted least squares
SE	
vce(*vcetype***)**	*vcetype* may be **conventional**, **bootstrap**, or **jackknife**
Reporting	
level(#)	set confidence level; default is **level(95)**
display_options	control spacing and display of omitted variables and base and empty cells
+ **coeflegend**	display coefficients' legend instead of coefficient table

+ **coeflegend** does not appear in the dialog box.

BE-estimator output

```
. xtreg I F C, i(fn) be

Between regression (regression on group means)  Number of obs      =       200
Group variable: fn                              Number of groups   =        10

R-sq:  within  = 0.4778                         Obs per group: min =        20
       between = 0.8578                                        avg =      20.0
       overall = 0.7551                                        max =        20

                                                F(2,7)             =     21.11
sd(u_i + avg(e_i.))=  85.02366                  Prob > F           =    0.0011
```

I	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
F	.1346461	.0287455	4.68	0.002	.0666739	.2026183
C	.0320315	.1909378	0.17	0.872	-.4194647	.4835276
_cons	-8.527114	47.51531	-0.18	0.863	-120.883	103.8287

The model

Random Efects model

Random Effect model

- Baltagi states that within-group estimation would mean a loss of degrees of freedom. That statement comes from the fact that within-group estimator needs to estimate N fixed effects or include $(N-1)$ dummies as regresors.

Random Effect model

- Baltagi states that within-group estimation would mean a loss of degrees of freedom. That statement comes from the fact that within-group estimator needs to estimate N fixed effects or include $(N-1)$ dummies as regresors.
- In this especification the individual effect is characterized as random and drawn from a sample, inference is conducted using that sample.

Random Effect model

- Baltagi states that within-group estimation would mean a loss of degrees of freedom. That statement comes from the fact that within-group estimator needs to estimate N fixed effects or include $(N-1)$ dummies as regresors.
- In this especification the individual effect is characterized as random and drawn from a sample, inference is conducted using that sample.
- Nerlove and Balestra (1996)'s estimator.

Random Effect model

- Baltagi states that within-group estimation would mean a loss of degrees of freedom. That statement comes from the fact that within-group estimator needs to estimate N fixed effects or include $(N-1)$ dummies as regresors.
- In this especification the individual effect is characterized as random and drawn from a sample, inference is conducted using that sample.
- Nerlove and Balestra (1996)'s estimator.
- Haavelmo (1944).

Random Effect model

A simplistic homogeneous model

$$y_{it} = \alpha + X_{it}'\beta + u_{it} \tag{20}$$
$$u_{it} = \mu_i + \nu_{it}$$

heterogeneity given by μ_i. $\mu_i \backsim N\left(0, \sigma_\mu^2\right)$ and $\nu_{it} \backsim N\left(0, \sigma_\nu^2\right)$.

$$u_{it}u_{it}' = (\mu_i + \nu_{it})(\mu_i + \nu_{it})'$$
$$\sigma^2 = \sigma_\mu^2 + \sigma_\nu^2$$

σ^2 is the variance based in the latter we construct $\rho = \frac{\sigma_\mu^2}{\sigma^2}$ (*equicorrelation*).

Random Effect model

the error component for *it*

$$u_{it} = \mu_i + v_{it}$$

the error component for all (matrix form)

$$u = (I_N \otimes \iota_T)\, \mu + v$$

we need to have the total (expected) variance of the error component

$$\begin{aligned} u'u &= \mu' \left(I_N \otimes \iota_T\right)' \left(I_N \otimes \iota_T\right) \mu + v'v \\ &= \mu' \left(I_N \otimes \iota_T' \iota_T\right) \mu + v'v \end{aligned}$$

Taking expectations

$$E(u'u) = T\sigma_\mu^2 + \sigma_v^2 \tag{21}$$

Random Effect model

getting the "verall" matrix variance

$$u = (I_N \otimes \iota_T)\,\mu + \nu$$

So,

$$\underset{NT\times NT}{\Omega} = E(uu')$$
$$\underset{NT\times NT}{\Omega} = \sigma_\mu^2\,(I_N \otimes \mathcal{J}_T) + \sigma_\mu^2\,(I_N \otimes I_T)$$

Random Effect model

$$
\begin{aligned}
\underset{NT\times NT}{\Omega} &= T\sigma_\mu^2\left(I_N \otimes \overline{\mathcal{J}}_T\right) + \sigma_\nu^2\left(I_N \otimes \left(E_T + \overline{\mathcal{J}}_T\right)\right) \\
\underset{NT\times NT}{\Omega} &= \left(T\sigma_\mu^2 + \sigma_\nu^2\right)\left(I_N \otimes \overline{\mathcal{J}}_T\right) + \sigma_\nu^2\left(I_N \otimes E_T\right) \\
\underset{NT\times NT}{\Omega} &= \sigma_1^2 P + \sigma_\nu^2 Q
\end{aligned}
$$

where $P = \left(I_N \otimes \overline{\mathcal{J}}_T\right)$, $Q = I_N \otimes E_T$ and $\sigma_1^2 = \left(T\sigma_\mu^2 + \sigma_\nu^2\right)$

Random Effect model

trick: Replace

$$\mathcal{J}_T \text{ by } T\overline{\mathcal{J}}_T$$
$$I_T \text{ by } E_T + \overline{\mathcal{J}}_T$$

where

$$E_T = I_T - \overline{\mathcal{J}}_T$$
$$\overline{\mathcal{J}}_T = \frac{\mathcal{J}_T}{T}$$

Random Effect model

the variance

$$\underset{NT\times NT}{\Omega} = \sigma_1^2 P + \sigma_\nu^2 Q$$

Fuller and Battese (1974)'s result considers $r = \frac{-1}{2}$

$$\underset{NT\times NT}{\Omega^{-\frac{1}{2}}} = \sigma_1^{-1} P + \sigma_\nu^{-1} Q$$

Remember that is very difficult to inverse a $NT \times NT$ matrix.

Baltagi says...

premultiplying by y

$$\underset{NT\times NT}{\Omega^{-\frac{1}{2}}} y = \sigma_1^{-\frac{1}{2}} Py + \sigma_\nu^{-\frac{1}{2}} Qy \tag{22}$$

so $Py = \underset{NT\times NT}{\overline{y}_i}$ and $Qy = \underset{NT\times NT}{y_{it}} - \overline{y}_i$.

$$\underset{NT\times NT}{\Omega^{-\frac{1}{2}}} y = \sigma_1^{-1}\overline{y}_i + \sigma_\nu^{-1}\left(y_{it} - \overline{y}_i\right)$$

Baltagi says...

finally

$$\underset{NT\times NT}{\Omega^{-\frac{1}{2}}} y = \sigma_1^{-1}\overline{y}_i + \sigma_v^{-1}\left(y_{it} - \overline{y}_i\right)$$

$$\underset{NT\times NT}{\Omega^{-\frac{1}{2}}} y = \sigma_v^{-1}(y_{it} - \left(1 - \frac{\sigma_1^{-1}}{\sigma_v^{-1}}\right)\overline{y}_i)$$

$$\sigma_v \underset{NT\times NT}{\Omega^{-\frac{1}{2}}} y = (y_{it} - \theta\overline{y}_i)$$

being $\theta = 1 - \frac{\sigma_v}{\sigma_1}$ or $\theta = 1 - \sqrt{\frac{\sigma_v^2}{T\sigma_\mu^2 + \sigma_v^2}}$

GLS

The question is: have we got an $i.i.d$ error disturbance with this "*long-step*" transformation?. in the case of the first observation of vector y

$$var\left(y_{i=i,t=0}\right) = var(y_{i0}) + \theta^2 \frac{\sum_j^T var(y_{i0})}{T} + \\ - 2\theta cov\left(y_{i0}, \overline{y}_i\right)$$

so

$$var\left(y_{i=i,t=0}\right) = (1-\theta)^2 \, var(y_{i0}) \\ = \frac{\sigma_\nu^2}{\sigma_1^2} var(y_{i0}) = \sigma_\nu^2$$

Theta's interpretation

- if $\sigma_\mu^2 = 0 \Rightarrow \theta = 0 \Rightarrow$ OLS is enough.

Theta's interpretation

- if $\sigma_\mu^2 = 0 \Rightarrow \theta = 0 \Rightarrow$ OLS is enough.
- if $T \longrightarrow \infty \Rightarrow \theta \longrightarrow 1$. Within effects! (macro data).

Theta's interpretation

- if $\sigma_\mu^2 = 0 \Rightarrow \theta = 0 \Rightarrow$ OLS is enough.
- if $T \longrightarrow \infty \Rightarrow \theta \longrightarrow 1$. Within effects! (macro data).
- So, $0 < \theta < 1$ for random effects.

Theta's interpretation

- if $\sigma_\mu^2 = 0 \Rightarrow \theta = 0 \Rightarrow$ OLS is enough.
- if $T \longrightarrow \infty \Rightarrow \theta \longrightarrow 1$. Within effects! (macro data).
- So, $0 < \theta < 1$ for random effects.
- We need to estimate the variance components in θ. That means 1 parameter instead of $N - 1$ dummies!.

Theta's interpretation

- if $\sigma_\mu^2 = 0 \Rightarrow \theta = 0 \Rightarrow$ OLS is enough.
- if $T \longrightarrow \infty \Rightarrow \theta \longrightarrow 1$. Within effects! (macro data).
- So, $0 < \theta < 1$ for random effects.
- We need to estimate the variance components in θ. That means 1 parameter instead of $N - 1$ dummies!.

Random Effect model: summarizing

- OLS still unbiased and consistent if the true model is RE model.
- But standard errors are wrong $\Rightarrow$ t-statistics are wrong. The estimator is inefficient!.
- GLS is ok (some cases are straightforward to treat) but we need to implement FGLS

FGLS in a Random effect model.

Estimators

- Wallace and Hussain (1969)
- Amemiya (1971)
- Swamy & Arora (1972)
- Breusch (1987) iterative GLS using MLE.
- MINQUE
- MIVQUE

Softwares : STATA, EVIEWS, RATS, GAUSS, OX,R. (*Auditorium: On above estimator, what do they want?*)

Random Effect model

- The variance-covariance matrix

$$\Omega = (I_N \otimes \iota_T) E(\mu\mu') (I_N \otimes \iota_T)' + E(\nu\nu')$$
$$\Omega = \sigma_\mu^2 (I_N \otimes \iota_T)(I_N \otimes I_T)(I_N \otimes \iota_T)' + \sigma_\nu^2 (I_N \otimes I_T) \tag{23}$$

this implies a homoskedastic variance $var(\epsilon_{it}) = \sigma_\mu^2 + \sigma_\nu^2$. for all i and t. Baltagi states that block-diagonal covariance matrix exhibits serial correlation over time

$$\begin{aligned} cov(\epsilon_{it}, \epsilon_{js}) &= \sigma_\mu^2 + \sigma_\nu^2 \quad \text{for} i = j \quad , \quad t = s \\ &= \sigma_\mu^2 \quad \text{for} i = j \quad , \quad t \neq s \end{aligned}$$

Random Effect model

Because of Baltagi's assumption we have that for $i \neq j$ and $t = s$ the following covariance should be zero:

$$\begin{aligned} cov\left(\epsilon_{it}, \epsilon_{jt}\right) =& E\left(\mu_i \mu_j\right) + E\left(\mu_i \nu_{jt}\right) + \\ & E\left(\mu_j \nu_{it}\right) + E\left(\nu_{it} \nu_{jt}\right) \end{aligned}$$

at first thought $E\left(\mu_i \nu_{jt}\right) = E\left(\mu_j \nu_{it}\right) = 0$ and we need to think in "very good reasons"to have other terms not equal to zero.

Random Effect model

Maddala (1971) shows that the GLS estimator can be expressed as follows:

$$\widehat{\beta}_{GLS} = \left[\frac{(X'QX)}{\sigma_\nu^2} + \frac{X'\left(P - \overline{\mathfrak{J}}_{NT}\right)X}{\sigma_1^2} \right]^{-1} \left[\frac{(X'Qy)}{\sigma_\nu^2} + \frac{X'\left(P - \overline{\mathfrak{J}}_{NT}\right)y}{\sigma_1^2} \right]$$

simplifying:

$$\widehat{\beta}_{GLS} = \left[W_{xx} + \phi^2 B_{xx} \right]^{-1} \left[W_{xy} + \phi^2 B_{xy} \right] \tag{24}$$

where $W_{xx} = X'QX, W_{xy} = X'Qy, B_{xx} = X'\left(P - \overline{\mathfrak{J}}_{NT}\right)X, B_{xy} = X'\left(P - \overline{\mathfrak{J}}_{NT}\right)y$ and $\phi^2 = \frac{\sigma_\nu^2}{\sigma_1^2}$.
Expression (24) is the Maddala's GLS estimator

Random Effect model

thus, recall that

$$\widehat{\beta}_{WG} = W_{xx}^{-1} W_{xy}$$
$$\widehat{\beta}_{BE} = B_{xx}^{-1} B_{xy}$$

So,

$$\widehat{\beta}_{GLS} = \vartheta \widehat{\beta}_{WG} + (1 - \vartheta) \widehat{\beta}_{BE} \tag{25}$$

where $\vartheta = \left[W_{xx} + \phi^2 B_{xx} \right]^{-1} W_{xx}$. Expression (25) tells us that GLS estimator is a weighted average between within-group estimator and between-group estimator.

Random Effect model

three cases to be considered:

- if $T \longrightarrow \infty \Rightarrow \phi^2 \longrightarrow 0 \Rightarrow \vartheta \longrightarrow 1$: In this case $\widehat{\beta}_{GLS}$ asymptotically approachs to within-group estimator.

Random Effect model

three cases to be considered:

- if $T \longrightarrow \infty \Rightarrow \phi^2 \longrightarrow 0 \Rightarrow \vartheta \longrightarrow 1$: In this case $\widehat{\beta}_{GLS}$ asymptotically approachs to within-group estimator.
- if $\phi^2 \longrightarrow \infty \Rightarrow \vartheta \longrightarrow 0$: In this case $\widehat{\beta}_{GLS}$ asymptotically approachs to between-group estimator.

Random Effect model

three cases to be considered:

- if $T \longrightarrow \infty \Rightarrow \phi^2 \longrightarrow 0 \Rightarrow \vartheta \longrightarrow 1$: In this case $\widehat{\beta}_{GLS}$ asymptotically approachs to within-group estimator.
- if $\phi^2 \longrightarrow \infty \Rightarrow \vartheta \longrightarrow 0$: In this case $\widehat{\beta}_{GLS}$ asymptotically approachs to between-group estimator.
- if $\sigma_\mu^2 = 0 \Rightarrow \phi^2 = 1$: In this case $\widehat{\beta}_{GLS}$ is equal to $\widehat{\beta}_{OLS}$

FGLS strategy: Amemiya

Considering the total variance

$$\underset{NT\times NT}{\Omega} = \sigma_1^2 P + \sigma_\nu^2 Q$$

So,

$$\widehat{\sigma}_1^2 = \frac{u' P u}{N} = \frac{T \sum_{i=1}^{N} \overline{u}_i}{N}$$

$$\widehat{\sigma}_\nu^2 = \frac{u' Q u}{N(T-1)} = \frac{T \sum_{i=1}^{N} \sum_{t=1}^{T} (u_{it} - \overline{u}_i)^2}{N(T-1)}$$

Random-Effect-model STATA syntax

```
. xtreg I F C, re theta

Random-effects GLS regression                   Number of obs      =       200
Group variable: fn                              Number of groups   =        10

R-sq:  within  = 0.7668                         Obs per group: min =        20
       between = 0.8196                                        avg =      20.0
       overall = 0.8061                                        max =        20

Random effects u_i ~ Gaussian                   Wald chi2(2)       =    657.67
corr(u_i, X)       = 0 (assumed)                Prob > chi2        =    0.0000
theta              = .86122362

------------------------------------------------------------------------------
           I |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           F |   .1097811   .0104927    10.46   0.000     .0892159    .1303464
           C |    .308113   .0171805    17.93   0.000     .2744399    .3417861
       _cons |  -57.83441   28.89893    -2.00   0.045    -114.4753   -1.193537
-------------+----------------------------------------------------------------
     sigma_u |  84.20095
     sigma_e |  52.767964
         rho |  .71800838   (fraction of variance due to u_i)
------------------------------------------------------------------------------

. estimates store re
```

Random-Effect (RE) options

RE_options	description
Model	
re	use random-effects estimator; the default
sa	use Swamy-Arora estimator of the variance components
SE/Robust	
vce(*vcetype***)**	*vcetype* may be **conventional**, **robust**, **cluster** *clustvar*, **bootstrap**, or **jackknife**
Reporting	
level(#)	set confidence level; default is **level(95)**
theta	report theta
display_options	control spacing and display of omitted variables and base and empty cells
+ **coeflegend**	display coefficients' legend instead of coefficient table

+ **coeflegend** does not appear in the dialog box.

FGLS strategy: MLE iterative

Breusch (1987) concentrates the likelihood with respect to:

$$\widehat{\phi}^2 = \frac{d' Q d}{(T-1)\, d' \left(P - \overline{\mathfrak{J}}_{NT}\right) d} = \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} \left(d_{it} - \overline{d}_i\right)^2}{T\,(T-1) \sum_{i=1}^{N} \left(\overline{d_{i\cdot}} - \overline{d}_{\cdot\cdot}\right)^2}$$

where $d = y - X\widehat{\beta}$, and $\phi^2 = \frac{\sigma_\nu^2}{\sigma_1^2}$, in fact $\theta \backsim 1 - \phi$

$$\widehat{\beta}_{mle} = \left[X' \left(Q + \phi^2 \left(P - \overline{\mathfrak{J}}_{NT}\right)\right) X\right]^{-1} \left[X' \left(Q + \phi^2 \left(P - \overline{\mathfrak{J}}_{NT}\right)\right) y\right]^{-1}$$

iteration process in order to get $\widehat{\beta}_{mle}$. be careful of values ϕ this should be positive and finite.

MLE iterative STATA syntax

```
.
. xtreg I F C, mle

Fitting constant-only model:
Iteration 0:   log likelihood = -1387.6302
Iteration 1:   log likelihood = -1291.9897
Iteration 2:   log likelihood = -1254.2888
Iteration 3:   log likelihood = -1243.6309
Iteration 4:   log likelihood = -1242.0548
Iteration 5:   log likelihood = -1241.9709
Iteration 6:   log likelihood = -1241.9696
Iteration 7:   log likelihood = -1241.9696

Fitting full model:
Iteration 0:   log likelihood = -1105.6101
Iteration 1:   log likelihood = -1098.8418
Iteration 2:   log likelihood = -1095.4188
Iteration 3:   log likelihood = -1095.2576
Iteration 4:   log likelihood =  -1095.257

Random-effects ML regression                    Number of obs      =       200
Group variable: fn                              Number of groups   =        10

Random effects u_i ~ Gaussian                   Obs per group: min =        20
                                                               avg =      20.0
                                                               max =        20

                                                LR chi2(2)         =    293.43
Log likelihood  =  -1095.257                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
           I |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           F |   .1097626   .0103389    10.62   0.000     .0894988    .1300265
           C |    .307942   .0171006    18.01   0.000     .2744254    .3414585
       _cons |   -57.7672   27.70004    -2.09   0.037    -112.0583   -3.476114
-------------+----------------------------------------------------------------
    /sigma_u |   80.29729   18.37811                      51.27213    125.7536
    /sigma_e |   52.49255    2.69306                      47.47094    58.04534
         rho |   .7005943   .0985226                      .4881266    .8603709
------------------------------------------------------------------------------
Likelihood-ratio test of sigma_u=0: chibar2(01)=  193.09 Prob>=chibar2 = 0.000
```

Interpretation of Phi

- if $\phi = 0 \Longrightarrow$ within-group estimator
- if $\phi = \infty \Longrightarrow$ between-group estimator

The model

Estimating in STATA

Estimating in STATA

Let's go to STATA!

Summarizing the STATA results

```
.
. estimates table fixed ols areg be, star stats(N r2 r2_a F)
```

Variable	fixed	ols	areg	be
F	.1101238***	.11556216***	.1101238***	.13464609**
C	.31006534***	.23067849***	.31006534***	.03203148
_Ifn_2	172.20251***			
_Ifn_3	-165.27514***			
_Ifn_4	42.487396			
_Ifn_5	-44.320126			
_Ifn_6	47.135393			
_Ifn_7	3.7432122			
_Ifn_8	12.751033			
_Ifn_9	-16.925585			
_Ifn_10	63.728842			
_cons	-70.296685	-42.714369***	-58.743932***	-8.5271142
N	200	200	200	200
r2	.94407252	.81240802	.94407252	.85776823
r2_a	.94080016	.81050353	.94080016	.81713058
F	288.49961	426.57575	309.01419	21.107723

legend: * p<0.05; ** p<0.01; *** p<0.001

Summarizing the STATA results

```
. estimates table ols sw_ar xtgls_iid xtgls_h, star stats(N r2 r2_a F)
```

Variable	ols	sw_ar	xtgls_iid	xtgls_h
F	.11556216***	.10978115***	.11556216***	.11163281***
C	.23067849***	.30811298***	.23067849***	.15377179***
_cons	-42.714369***	-57.834409*	-42.714369***	-21.443484***
N	200	200	200	200
r2	.81240802			
r2_a	.81050353			
F	426.57575			

legend: * p<0.05; ** p<0.01; *** p<0.001

Summarizing the STATA results

```
. estimates table xtgls_iid xtgls_h xtgls_ar1 xtgls_c, star stats(N r2 r2_a F)
```

Variable	xtgls_iid	xtgls_h	xtgls_ar1	xtgls_c
F	.11556216***	.11163281***	.09381622***	.11275148***
C	.23067849***	.15377179***	.30398969***	.22311756***
_cons	-42.714369***	-21.443484***	-36.903476	-39.843818***
N	200	200	200	200
r2				
r2_a				
F				

legend: * p<0.05; ** p<0.01; *** p<0.001

References

References

- Arellano, M. 2003. Panel Data Econometrics.Oxford University Press.
- Baltagi, B. 2001. Econometric Analysis of Panel Data.
- Benavente, J.M. 2005. Guia de microeconometria. Universidad de Chile.

References

- Arellano, M. 2003. Panel Data Econometrics.Oxford University Press.
- Baltagi, B. 2001. Econometric Analysis of Panel Data.
- Benavente, J.M. 2005. Guía de microeconometría. Universidad de Chile.

- NOTAS

Econometría Avanzada

Catedrático: PhD. Freddy Rojas Cama

LaTeX suppor and edition:
Joel Vicencio-Damian
joel.nestor.damian@gmail.com